# The Impact of Light Rail Station Type on Single Family Residential Property Value in the City of Portland, OR

Siman Ning

A thesis

submitted in partial fulfillment of the

requirements for the degree of

Master of Urban Planning

University of Washington

2020

Committee:

Qing Shen

Jan Whittington

Program Authorized to Offer Degree:

College of Built Environments

University of Washington

**Abstract**

The Impact of Light Rail Station Typology on Single Family Residential Property Value in the City of Portland, OR

Siman Ning

Chair of the Supervisory Committee:

Professor Qing Shen

Department of Urban Design and Planning

Light rail projects increase accessibility and light rail stations should result in a land value uplift (LVU) benefit for housing near stations. Extensive research has tested this hypothesis, generally using distance or distance band as a proxy of accessibility. However, the context of transit stations has not gained enough attention in past research. This research divides light rail stations into four different types based on land use ratio, density, sidewalk length, and accessibility; later, the station types are incorporated into a hedonic model. The modeling results confirm the price premium brought by light rail stations to nearby single-family homes. Findings also suggest that, after controlling for accessibility, TOD impacts on single family housing price varies in relation to station type. In other words, station area land use matters when considering land use value capture from light rail development. Future research should further examine how surrounding land uses interact with TOD in affecting housing price.

**Table of Contents**

## List of Figures

## List of Tables

## Chapter 1 Introduction

A good public transport system is often one of the defining features of a city, but even in the case of the world's great public transport systems, fares do not fully cover costs. The construction and operation of the light rail systems need substantial financial subsidies from government. It is always challenging for the city to determine where this money should come from. Land value capture is an often-discussed topic for gathering funding for transit service (C. D. Higgins & Kanaroglou, 2016b; Mulley, Ma, Clifton, Yen, & Burke, 2016; Salon & Shewmake, 2012; Suzuki, Murakami, Hong, & Tamayose, 2015). As a successful public transport system leads to increasing land values, the government should be able to capture part of the value increase along transit corridors and use these funds to subsidize the transit system.

This research doesn't discuss the mechanism of how land value capture works, but it focuses on the question of whether TOD type has an impact on housing price. TOD typology groups areas near light rail stations in a region based on shared sets of characteristics (Kamruzzaman, Baker, Washington, & Turrell, 2014).Urban planners have been promoting TOD to fight urban sprawl for years (Ewing & Cervero, 2010; C. D. Higgins & Kanaroglou, 2016a; Nasri & Zhang, 2014). Planners argue that TOD increases accessibility and such benefits should be captured by the price of housing, and a lot of research confirms that TOD improves accessibility (Ewing & Cervero, 2010; C. Higgins & Kanaroglou, 2018; Mulley et al., 2016; Nasri & Zhang, 2014; Shen, Xu, & Lin, 2018). However, there is not a 'one-size-for-all' TOD solution when we consider the various contexts in cities (Dittmar and Poticha, 2004; Zemp, Stauffacher, Lang, & Scholz, 2011; Kamruzzaman, Baker, Washington, & Turrell, 2014). Belzer and Autler (2002) argue that TOD can take a variety of forms; and there are advocates for understanding the benefits of TOD from a typology perspective (Renne and Wells, 2005).

A lot of studies have been done to understand the TOD housing premium and the built environment (Deng, 2011; Diao, Zhu, & Zhu, 2017; Ewing & Cervero, 2010; Hess & Almeida, 2007; C. D. Higgins & Kanaroglou, 2016a; C. Higgins & Kanaroglou, 2018; Laaly, 2014; Polloni, 2019; Shen et al., 2018; Suzuki, Cervero, & Iuchi, 2013). Past research has shown that built environment matters: being close to a parking lot is a disamenity to TOD housing, while the pedestrian-friendly design is an amenity. Some studies have built typologies of TOD, in order to describe the context itself and to understand TOD performance and relationship with socio-economic characteristics, based on either subjective judgement or built environment characteristics (Atkinson-Palombo, 2010; Atkinson-palombo & Kuby, 2011; Guan & Peiser, 2018; C. D. Higgins & Kanaroglou, 2016a; Kamruzzaman et al., 2014; Ratner & Goetz, 2013; Zemp et al., 2011). However, there aren't many studies about housing price and TOD typology (Atkinson-Palombo, 2010; C. Higgins & Kanaroglou, 2018).

I use hedonic price model to examine the relationship between TOD types, TOD proximity and single-family housing sales price in the City of Portland, Oregon. Hedonic price model is a widely used technique in studies of housing price analysis. The TOD types are identified by cluster analysis method, which is technique to group similar observations into clusters. The control variables consist of housing structure factors, neighborhood factors, and location factors. This research may help planners understand the different types of TOD, and hopefully more policy attention will be paid to the built environment of TOD.

The key interest of this study is the underlying mechanism of interaction between the TOD station typology and surrounding single-family properties, which can be translated into the following three questions:

- What types of TOD exist within the city of Portland, OR?
- How has TOD changed the price of single-family houses in the city of Portland, OR?
- To what extent do the different TOD types impact the capitalization of TOD on single-family houses?

This thesis begins with a literature review to identify influential factors to property value to build a conceptual model, followed by a brief introduction to the study area. Then I give a detailed description of the methodology and regression results of both cross-section and time-series models to measure the premiums of TOD typology and proximity. Finally, conclusions and implications are given for future research.

## Chapter 2 Literature review

### 2.1 What is TOD? How does existing literature define TOD?

The concept of transit-oriented development (TOD) comes from Peter Calthorpe in the early 1990s when he brought up a new urbanism concept. Since then, various definitions of TOD have surfaced in literature with different scopes yet shared certain elements. My thesis uses the definition of TOD as "compact, mixed-use, pedestrian-friendly development organized around a transit station" (Suzuki, Cervero, & Iuchi, 2013, page. xxi). Suzuki et al. (2013) further describe TOD "embraces the idea that locating amenities, employment, retail shops, and housing around transit hubs promotes transit usage and non-motorized travel." (Suzuki, Cervero, & Iuchi, 2013, page. xxi) .

On the other hand, another concept of TAD (Transit Adjacent Development)- which is usually seen as 'the evil brother of TOD' has since emerged. Cervero et al. describe TAD as "development that is physically near transit; it fails to capitalize upon this proximity, however, to promote transit riding. A TAD lacks any functional connectivity to transit, whether in terms of land-use composition, means of station access, or site design" (Cervero, Ferrell, & Murphy, 2002, page 6). Sarah Lovell, the TOD program manager of King County Metro summarized characteristics of TAD as: overprescribed parking, underbuilt site, not well-connected public space, 'island' like development, not-human scale, single usage, and exclusive (Lovell, 2019). Renne et al. (2016), use walkability and housing density targets as criteria to classify TOD and TAD. A Walk Score below 70 or a gross housing density below 8 units per acre serve as the classification criteria for TAD. However, these criteria haven't gained wide application.

Most of the literature studying the TOD housing price hasn't compared TAD to TOD. TOD is assumed to exist within a certain proximity to the light rail stations. A number of US TODs discussed in the literature actually resemble TADs (Cervero, Ferrell, & Murphy, 2002).

### 2.2 TOD Typology and Context

Instead of viewing the TOD and TAD as 'good' and 'evil', a more practical approach might be admitting the multitude of regional contexts and using typology to contain the complexity. Dittmar and Poticha (2004) argue that there are differences between places, and it is hard to find a 'one-size-for-all' TOD solution. They urge to use typology as a tool to identify performance benchmarks for different places. A few researchers have created typologies of TOD using either normative approach or positive approach.

#### 2.2.1 Normative approach

Some literature classifies TOD based on a combination of observation and subjective judgment. Usually such categorization is either to describe the different urban form contexts or to build strategies for future planning and policy. For example, White and McDaniel (1999) identified six types of TOD based on urban form observation: single use corridor development, mixed-use corridor development, neotraditional development, TOD with pedestrian pockets, hamlets and villages, and purlieus.

Dittmar and Poticha (2004) produced a TOD-centric typology of six hypothetical TOD contexts: urban downtown, urban neighborhood, suburban center, suburban neighborhood, neighborhood transit zone, and commuter town center. They created the contexts to discuss how different TOD goals may differ by typology. Ratner and Goetz (2008) analyzed TOD data from Denver and to examine how transit accessibility and TOD impact land use and urban form in the Denver region. The study created a typology of Denver's TOD stations based on desired land use mix, desired housing types, commercial employment types, proposed scale, and transit system function. The study concludes that the regional policy, which is encouraging more transit and higher-density TOD, is having a recognizable impact on Denver's land use and urban form. The normative approach is generally an early-stage exploration to understand TOD contexts, but it fails to take into account the built environmental characteristics in a rigorous way.

### 2.2.2 Positive approach

Later research tries to find ways based on TOD performance or characteristics of TOD to create classification criteria, for the general purpose of guiding future policy and strategic planning. Usually more sophisticated statistic methods are employed in the research, such as factor analysis, cluster analysis and latent class analysis. Atkinson-Palombo and Kuby (2011) used TOD typology to evaluate spatial distribution of overlay zoning policy (an additional, targeted layer of regulations) around light rail stations areas in Metropolitan Phoenix. The station areas were defined as parcels within a 1/2-mile walking distance of light rail stations. Five distinct station-area types were identified using cluster analysis: employment centers, middle-income mixed-use areas, transportation nodes, high population/rental areas, and urban poverty areas. The study shows that overlay zoning differs significantly across the classes (e.g. most in areas of urban poverty, and least in areas with more single family housing). The work suggests that station typology may be useful to policy-makers by providing insights into issues such as income, housing tenure and whether or not existing land uses are compatible with TOD. Kamruzzaman, Baker, Washington, and Turrell (2014) used six indicators (employment density, residential density, land use diversity, intersection density, cul-de-sac density, and public transit accessibility) to classify TOD, in order to understand people's mode choice behavior in Brisbane, Australia. Four clusters were derived, namely the non-TOD, active center TOD, potential TOD and residential TOD. The multinomial logistic regression model reveals that different types of TODs do influence mode choice behavior after controlling for residential self-selection, suggesting different development strategies for future TOD planning. Higgins and Kanaroglou ( 2016a) used a model-based latent class method to classify TOD, to understand whether TOD typology makes a difference in performance measures, such as commute mode share, travel characteristics, and socio-economic characteristics. They created a typology of ten TOD contexts across a number of transit lines based on density, development mix, street connectivity, interaction potential, and land use mix. Later, they applied the typology in further analysis with a hedonic model, in order to understand land value uplift from various station types in Toronto. The hedonic model suggests that the maximum amount and spatial impact area of land value capitalization differ by station contexts (Higgins & Kanaroglou, 2018b).

To develop a typology is to create groups that have shared characteristics. The use of typology in TOD research has a unique advantage: by categorizing shared characteristics, policymakers can enhance the planning, design and operation of TOD in many ways. For example, the typology can help policy makers and stakeholders create common strategies for areas with similarities to improve performance (Kamruzzaman et al., 2014). TOD typology helps answer "what densities and level of transit service are necessary?" or "what kind of mixed land use is effective in TOD planning?" (Belzer and Autler, 2002). TOD typology is helpful in building performance assessments, identifying successful benchmarks and focusing on needs for action (Zemp et al., 2011).

## 2.3 TOD Housing Price Study

Salon and Shewmake (2012) summarize six major quasi-experimental design strategies to evaluate the impact of transportation infrastructure on land values: before/after study, control region comparisons, hedonic regression (cross‑sectional), fixed effects, propensity score matching, and spatial regression. For our study purpose, this literature review only looks into hedonic regression.

Hedonic regression can control for observable differences between treated and untreated areas. Hedonic regression allows researchers to decompose a good into different characteristics and the contribution of value brought by each characteristic. The hedonic regression model can be expressed as following (Rosen 1974):

$$P_i^j = \alpha + x_i \delta + \beta D^j + \varepsilon_i^j$$

Here the $x_i$ are observable factors at neighborhood and parcel level, such as neighborhood socio-economic aspects (e.g., crime, income, school district), physcial structure aspects (e.g., lot size, building age), location aspects (e.g., distance to highway, distance to CBD), environmental aspects (e.g. air pollution and noise), amenity (e.g., parks, shopping). $\beta$ is the parameter that reflects the impact of transportation improvement on property value. The dummy variable $D^j$ stands for treatment (transportation improvement). Some studies model the transportation improvement by a distance of impact rather than a zero/one impact, using $D^j$ as the measure of distance rather than a dummy variable. In either case, hedonic models can combine cross-section and time-series data. (Salon & Shewmake, 2012)

The hedonic regression modeling is popular in evaluations of the impact of TOD on property values (Bartholomew & Ewing, 2011; Dill, 2008; Duncan, 2011; Guan & Peiser, 2018; C. D. Higgins & Kanaroglou, 2016; C. Higgins & Kanaroglou, 2018; Mulley, Ma, Clifton, Yen, & Burke, 2016; Nelson et al., 2015; Seo et al., 2014; Shen, Xu, & Lin, 2018; Wang, 2016; Welch, Gehrke, & Farber, 2018; Yan, Delmelle, & Duncan, 2012). A main concern for the hedonic method is that omitted variables might contribute to the price of goods. In the case of TOD housing price, seeking to value the implicit prices of neighborhood amenities might be problematic as many neighborhood attributes might not be observable or might be highly correlated with other characteristics (de Haan & Diewert, 2013).

## 2.4 Current findings of TOD impact on housing price

The price effects of TOD on surrounding properties consist of several parts:

a. The introduction of transit service increases travel options for residents and employees in the area of central business districts (CBD) and other activity centers (Fejarang, 1993). The accessibility provided by TOD will be perceived by business as location advantage, which further translates into land values (Landis and Huang 1995).

b. The mixed land use associated with TOD increases local access to different kinds of daily activities, like school, shopping, and recreation. Such enhancement is capitalized into land values. Although not all mixtures of land uses near TOD increase housing values, parks are most likely to be a land use welcomed near residential properties. Commercial uses are likely to have both amenities and disamenities (known as the "next door" phenomenon), as the convenience of shopping have a wider geographic benefits, while the visual pollution and noise are likely to be experienced by homes with immediate adjacency (Bartholomew & Ewing, 2011).

c. Pedestrian-friendly design around TOD stations may also increase surrounding land value. Bernick and Cervero (1996) emphasize the role of the "three Ds" (density, diversity and design) as the necessary elements to achieve successful TOD. Song and Knaap (2003) focus on block size and street interconnectivity, and their research finds that home buyers are willing to pay a premium for neighborhoods containing interconnected streets and smaller blocks in Portland, OR. Duncan (2011) looks into TOD in the San Diego, CA, condominium market and suggests that station proximity has a significantly stronger impact when coupled with a pedestrian-oriented environment.

d. Disamenities and amenities both exist in TOD. Seo, Golub and Kuby (2014) find out that the disamenities of noise and pollution should emanate from the rail links as well as the nodes. This brings a layer of disamenity to consider in TOD research. The crime and poverty sometimes associated with transit may lower surrounding property value. (Bowes & Ihlanfeldt, 2001)

e. The length of 'after' period and the selection of control group have impact on the coefficients. Pilgram and West (2018) find that the price effects of TOD are measurable in years following the initiation of light rail service, but faded to a statistical zero as time passed. The premium for station proximity varies substantially depending on period definitions for "after" light rail, but little on the "before" light rail. The work also suggests using homes in the rest of Minneapolis as controls yields higher positive premiums than using homes in neighborhoods similar to those near stations.

f. Neighborhood context matters. After examining the zoning data near stations in LA, Boarnet and Crane (1997) concluded that station areas in suburban and predominantly residential municipalities had more residential zoning than stations in more centrally-located cities did. Atkinson-Palombo (2010) figured that amenity-dominated mixed-use neighborhoods see a capitalization for single family and multifamily houses, while park and ride neighborhoods experienced no capitalization benefits for single family houses. Higgins and Kanaroglou (2018) found that the station areas are heterogenous in capitalization of TOD over time.

## Chapter 3 Study Area

### 3.1 Portland Transit Network

According to the Census estimate, the city of Portland had a population of 583,776 in 2010 and 653,115 in 2018. It is the largest city in Oregon, with proximity to a few other significant cities like Beaverton, Hillsboro, Gresham in OR and Vancouver in Washington. The Portland Metropolitan area is served by three public transit services- the MAX Light Rail, the Portland Streetcar and WES Commuter Rail. This thesis focuses on the MAX Light Rail Stations within the city boundary of

Portland.

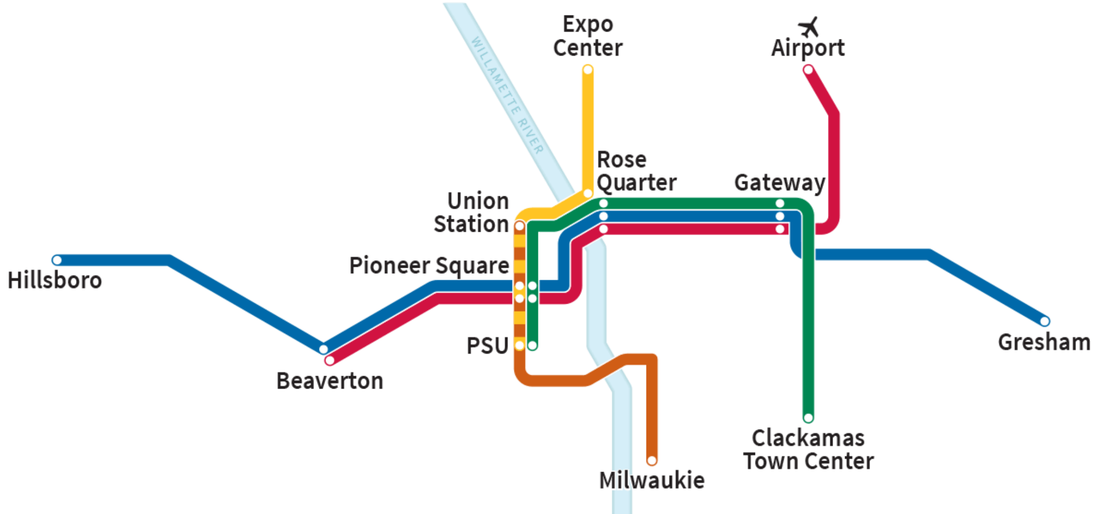



**Figure 1 Portland Light Rail Map (Source: Trimet.org)**

The five lines of MAX Light Rail are Blue, Red, Yellow, Green, and Orange, following the sequence of construction. The detailed construction cost and timeline could be found in Table 1.

The Blue Line is the oldest yet longest line in the network, connecting the cities of Gresham, Portland, Beaverton and Hillsboro. The funding for the first light rail in Portland comes from the reallocation of federal assistance funds for freeway projects after 1970s freeway revolts. The Eastside MAX (Blue Line) opened in 1986, while the Westside MAX was delayed for almost a decade due to funding disagreement and opened in 1998. The two segments were unified in 2000 as the Blue Line (Trimet, 2016a; Trimet, 2016b).

The Red Line was envisioned to connect the Portland International Airport to the city, due to the steady growth of airport traffic. In 1997, a unique public/private partnership was formed between Bechtel Enterprises and the region. Bechtel contributed about a quarter of funding to build the light rail extension in exchange for development rights to a piece of mixed-use land near the airport. The construction took five years, and the Redline opened in 2001 (Trimet, 2016c).

Considering aligning the growing need of growing Clackamas County and Downtown Portland, the Green Line was planned to provide a multitude of destinations. The construction was very efficient and took less than three years. The Green Line came into use in September 2009 (Trimet, 2016d).

The Yellow Line was originally part of the South-North light rail project, which would link Milwaukie, Portland, and Vancouver. The plan was later modified to link the PSU and the Expo Center in the north. About three-quarters of the project cost was funded by federal money. The Yellow Line opened in May 2004, four months ahead of schedule (Trimet, 2016e).

With a projection of 400,000 new residents in the Portland Metropolitan area by 2035, it is expected that the Portland-Milwaukie corridor would go through significant growth. The Orange Line was conceived to align with growth in the corridor and started service in September 2015 (Trimet, 2016f).

**Table 1 Summary of MAX Lines**

| | |
|---|---|
| MAX Blue Line Eastside | Opened: September 5, 1986<br>Length: 15 miles<br>Stations: 48<br>Construction: March 1982–September 1986<br>Cost: $214 million |
| MAX Blue Line Westside | Opened: September 12, 1998<br>Length: 18 miles<br>Stations: 20<br>Construction: August 1993–September 1998<br>Cost: $963.5 million |
| MAX Red Line | Opened: September 10, 2001<br>Length: 5.5 miles<br>Stations: 26<br>Construction: May 1999–September 2001<br>Cost: $125 million |
| MAX Yellow line | Opened: May 1, 2004<br>Length: 5.8 miles<br>Stations: 17 |

| | |
|---|---|
| | Construction: February 2001 – April 2004<br><br>Cost: $350 million |
| MAX Green Line | Opened: September 12, 2009<br><br>Length: 8.3 miles<br><br>Stations: 30<br><br>Construction: February 2007–September 2009<br><br>Cost: $575.7 million |
| MAX Orange Line | Opened: Sept 12, 2015<br>Length: 7.3 miles<br>Stations: 17<br>Construction: May 2011–Sept. 2015<br>Cost: $1.49 billion |

3.2 The Evolution of TOD in Portland

According to the Transportation Research Board (2003), the Portland region has a very aggressive TOD program, which has evolved over three decades.

The idea to integrate land use and transportation in the region can be traced back to the 1973 Downtown Plan. A transit mall was envisioned as the centerpiece of downtown, which would revitalize the community. After the Transit Mall opened in 1978, the strategy of using transit investment to achieve community goals became a signature.

The promotion of TOD strategy came from the afterthought of Portland's first rail line opening. A brief history of how TOD evolved can be found in the Trimet fact sheets:

- Informed by the Eastside light-rail line (now part of the Blue Line), the Westside stations were purposefully aligned with future development (Transportation Research Board, 2003; Trimet, 2016a; Trimet, 2016b).

- In the 1990s, the Streetcar project focused on promoting housing in Central City and served as a tool for both transportation and housing (Transportation Research Board, 2003).

- TOD was the central idea to the financing of the Red Line, as Bechtel Enterprise asked for the development right of the 120-acre mixed-use at the airport's entrance. However, Bechtel eventually sold the development rights in 2006. The 9/11 event and economy recession were cited as a reason why development didn't occur as expected. Today, the 120-acre site is a retail center with 44 stores and restaurants and three hotels (Transportation Research Board, 2003; Trimet, 2016c).

- The Yellow Line project placed more focus on revitalization opportunities and pedestrian scale. However, not all the transit corridors lead to compact urban development, which is a central feature of TOD (Trimet, 2016d).

- Not all the transit corridors in Portland led to compact and walkable urban development. Part of the Green Line corridor was aligned with the highway, which is more cost-effective for lower costs on Right-of-Way. Such alignment created some pedestrian access challenges, according to the Transit-Oriented Development Strategic Plan (Center for Transit-Oriented Development, 2011).

Overall, we can see an evolving effort in the Portland metropolitan area to use TOD as a tool to align land use with transit planning, to create compact development near the light rail station, to revitalize communities, and to guide growth in the region.

To stimulate the construction of transit villages, Metro (the Regional Government) uses federal transportation funds to create the TOD Implementation Program. The focus of the TOD Implementation program is to help acquire, design, and construct transit amenities as part of TOD. Due to the high cost of development, a "highest and best transit use" appraisal is used to establish sale price. (Transportation Research Board, 2003) For this study, it also makes sense to check whether the heavy investment on TOD has an expected return.

## Chapter 4 Data and Methodology

In the literature, hedonic pricing is the most popular method to measure the utility of TOD brought to residential housing. The repeat-sales method, which assesses how house valuations change over time by focusing on the different sale prices of the same piece of real estate, is sometimes also used for TOD studies.

The hedonic pricing model has the advantage of estimating values based on real choices, mainly when applied to property markets with readily available, accurate data. Hedonic pricing is flexible enough to be adapted to relationships, among other market goods and external factors. The method assumes that regressing the utility-bearing attributes to the housing price can reveal the inherent value at market equilibrium.

The main drawback of hedonic pricing model is that omitted independent variables could seriously bias the estimation of the dependent variables. Since in the practice many variables are hard to measure or simply are not available, omitted variables bias will be almost inevitable. Another shortage of hedonic model is self-selection. If there is heterogeneity across individuals in tastes for TOD, then the estimates of the hedonic model may only reflect the preferences of subgroup population (de Haan & Diewert, 2013).

In this study, a two-staged approach is adopted to reveal the TOD impact on land values. First, I measure each parcel's proximity to the nearest light rail station as a main key variable, like most of other previous studies. Next, a second key variable that controls for the TOD context is included in the model, to isolate the value of different types of TOD. A cluster analysis method is applied to determine the typology of TOD based on density, accessibility, sidewalk length and land use diversity. Then the classification of TOD is applied to the hedonic pricing model as a key variable. The model could be described as shown in the following Figure 2.

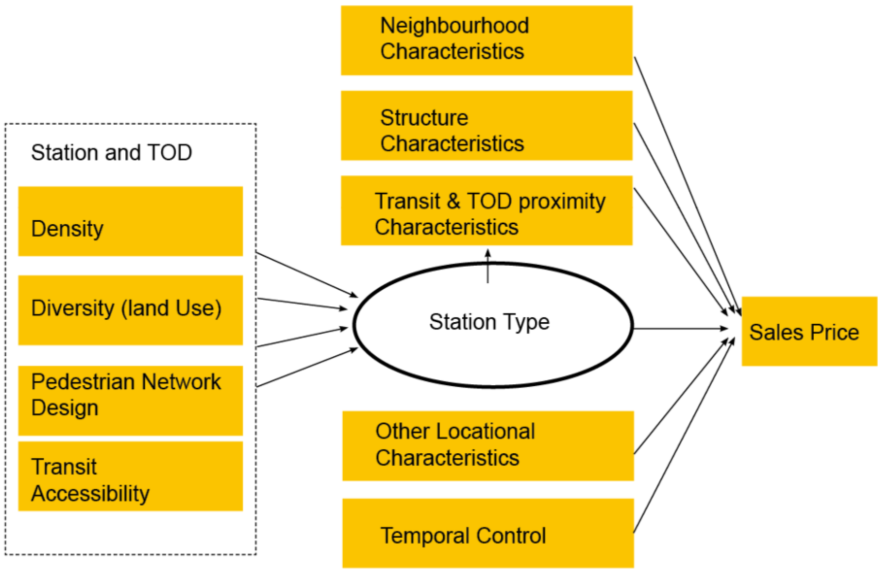



**Figure 2 the Model Structure**

To evaluate the impact of TOD on single family housing, the following equation shows how the variables reflecting structural, location, TOD and time of sale characteristics of the home are regressed on its sale price.

$$P = c_0 \sum \alpha_i A_i + \sum \beta_j B_j + \sum \mu_k U_k + \sum v_l V_l + \sum \omega_m W_m + \varepsilon \quad (1)$$

Where

P = the real sales price for single family property located near light rail stations within City of Portland;

$A_i$ = the physical characteristics of the property;

$B_j$ = the neighborhood socio-economic characteristics;

$U_k$ = the TOD characteristics, including job accessibility, land use mix entropy, TOD typology;

$V_l$ = the distance to light rail stations;

$W_m$ = the other locational variables;

$c_0$ = a constant.

In this paper I am going to test the following hypotheses:

- H1: The accessibility and the built environment characteristics within a rapid transit station's catchment area (0.5 mile ) combined can be categorized into two or more types.
- H2: Proximity to TOD will create a house price premium in single family houses.
- H3: Different categories of TOD are associated with differential land value uplift price. Locations with higher TOD characteristics- mixed uses, more walkable, higher density, should have higher land value uplift.

## 4.1 Data by Types

### 4.1.1 Housing Transaction Data.

The housing transaction data and property characteristics are available through a government-sponsored web portal PortlandMaps.com. For housing data, I first obtain a list of 61,381 single family building addresses from the city's open data website. This list includes the property address, unique property ID, usage, and geographic coordinates. Then I programmatically query the web portal with Scrapy (a framework for web scrapping) and Selenium (a framework for the automated testing of web applications) using each address in the previous list. When a query matches the residential property address and property ID, the sales data and housing structure information such as bathroom, land area, building area, rooftop type, etc. are recorded. However, it should be noted that a significant amount of sales data recorded are deleted due to implausible information, such as '0' and extremely low values (below 20,000 in 2019 dollars) for sales value. In the end, I have a total of 48,237 sales transactions.

### 4.1.2 Home Price Index Data

To account for inflation and market volatility, the S&P/Case-Shiller Portland Home Price Index is introduced to convert all sales price data into present dollars. Case-Schiller index was developed specifically to measure the average change of residential property value. Therefore the Case-Schiller index serves our purpose better than the Consumer Price Index. Fig X shows the index.

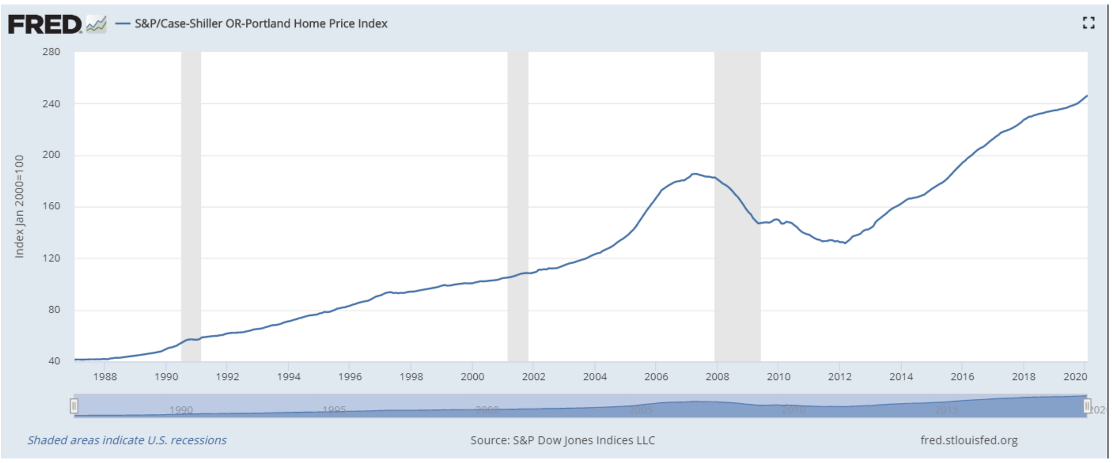



**Figure 3 S&P/Case-Shiller OR-Portland Home Price Index**

(source: [https://fred.stlouisfed.org/series/POXRSA](https://fred.stlouisfed.org/series/POXRSA))

4.1.3 Housing structure data

Housing structure data are extracted from the PortlandMaps.com portal website, which is the same portal as that used for the housing transaction data. The record includes the roof cover type, heating type, plumbing (how many bathrooms), land area, building area, year built, surface elevation, etc. Since the record is about current condition of properties, using this data for historical transactions may create inaccuracy. I have 21,968 addresses out of the total 61,381 single family home addresses. It should be noted that one address can have multiple transaction records.

4.1.4 School District and Performance

The school district and school performance data are obtained from different sources: the school attendance area is available from Portland Open Data website; the school accountability data can be downloaded from the Oregon Department of Education; the accountability data includes regular attendance, ELA (English language arts) and math achievement and growth, etc. The graduation rate of each school can be obtained from the same website as well. Although I intended to use the GreatSchool.org rating as an indicator of school quality, many schools in Portland do not have a rating at all. Thus turning to the accountability data is a natural choice. However, this creates extra complexity as there are many dimensions of school performance measurement, and school performance is not our primary focus. There are a total of 131 school districts in Portland, and the housing data overlapped with 70 of them.

### 4.1.5 Socio-economic Data

The socioeconomic data are from the 2017 Census and American Community Survey (ACS) 5-year Estimate (2013-2017). The census GDB data are obtained from TIGER/Line Geodatabases. The data of focus include median household income, race ratio for white, total population etc.

### 4.1.6 Neighborhood Crime Data

The neighborhood crime data is gathered from the website of the Police Bureau. Crime incidents are averaged from 2015- 2019 for calculation. The neighborhood population is gathered from the civic website. 78 out of 96 neighborhoods were included in the final address list.

### 4.1.7 Employment-household dynamic

Employment-household dynamic data is used to calculate location accessibility. Location of workers' jobs and residence can be generated from OnTheMap, a web-based mapping tool developed by U.S. Census Bureau.

### 4.1.8 Other GIS Data

Most of the GIS data in this category is to provide location information such as waterbody, open space, road network, rail network, etc. Some of the data can be used to calculate land use entropy (land uses, zoning). Location factors of housing properties are to be measured in the ArcMap Desktop 10.7. The locational data I am using here includes ,distance to river, distance to parks, distance to nearest school, distance to arterial road, distance to nearest bus stop, distance to railroad, distance to nearest light rail station.

## 4.2 Data by Category

### 4.2.1 Dependent Variable- Adjusted Transaction Price

The transaction data is gathered from Portland open data website. Considering that inflation will reduce the purchasing power of money, transaction price in previous years are adjusted to 2019 dollars using the annual S&P/Case-Shiller Portland Home Price Index. The function is as follows:

$$V_n = V_p \times (I_n / I_p) \qquad (2)$$

Where $V_n$ is the value of dollar in the present year. Vp is the value of dollar in the previous year. $I_n$ is the S&P/Case-Shiller Portland home price index in the present year. $I_p$ s the S&P/Case-Shiller Portland home price index in the present year.

4.2.2 Controlling Independent Variable-Housing Structure

The physical structure controlling variables entering in the model are square footage of the parcel, square footage of building area, number of bathrooms, age of the structure at sale time, and number of stories. The building structure data is directly obtained from the open data website of portlandmaps.com.

4.2.3 Controlling Independent Variable-Locational Characteristics

The locational characteristics include distance to highway, distance to arterial road, distance to river, distance to nearest open space, distance to nearest school, and distance to other bus stops. All of these locational variables are obtained by using ArcMAP near function. It should be noted that all are direct distances, not network distances.

4.2.4 Controlling Independent Variable – Social-Economic Factors

- White ratio: White ratio is derived from the population of white divided by total population in the census block group. Population data is gathered from 2014- 2018 ACS 5 year estimates.

- Median Household Income: Median household income at census block group is derived from 2018 ACS 1 year estimate.

-School Quality: Housing location layer is overlaid with the school district layer and then linked to the school accountability data. For the accountability data, only current ELA and Math achievement ratings are used for elementary school, middle school and high school in each school district. The ratings are ordered from level 1 to level 5. Level 5 means the performance meets state target for 2025. Level 4 means the performance is on track to meet state target for 2025. Level 3 means the performance is at or above state average but not on track. Level 2 means the performance is below state average. Level 1 means the bottom 10 percent statewide.

Crime Rate: The crime rate is generated from 5 year average incidents and the neighborhood population. The function is as follows:

$$R_c = I_c \times 100{,}000 / P_n \qquad (3)$$

Where $R_c$ is the crime incidents per 100,000 people. $I_c$ is the 5 year average crime incidents in each neighborhood. $P_n$ is the total population in each neighborhood.

4.2.5 Explanatory Independent Variables – TOD-Related Variables

Ewing and Cervero (2010) defined the five 'D' characteristics of TOD as density, diversity, design distance to transit, and destination access. In the density, the category includes household/population density, job density, and commercial floor area ratio. The diversity category includes variables like land use mix, job-housing balance, and distance to a store. In the Design, categories include intersection/street density and percent of four-way intersections. Destination accessibility includes jobs within 1 mile. Higgins and Kanaroglou (2018) utilized the 5 'D' theory and constructed the variables as follows: distance to transit (distance); population and employment per hectare (density); ratio of employment to the sum of population and jobs, and land use mix (diversity), the interaction between stations (accessibility). Here I use a similar approach to conceptualize variables.

- Diversity: Land use entropy

Land use entropy is a measurement which reflects the evenness of distribution of land use types within the region (Frank et al, 2004, 2006). The land use entropy is expressed as follows:

$$\text{Land Use Entropy} = \frac{-\Sigma(P_{ij} \times \ln(P_{ij})}{\ln(N_j)} \quad (4)$$

Where $P_{ij}$ is the proportion of total land area of a particular land-use category i within area j. $N_j$ is the total land uses in the study area. Since the entropy is normalized, the value of entropy is between 0 and 1. Where 1 represents evenly distributed land uses across all categories, and 0 represents homogeneous land use. In this study, j is the area within a 0.5 mile radius of each TOD station, and Nj includes seven land use types: single-family, multi-family, office, commercial, institution, open space and industrial use.

- Design: Sidewalk Length

The total sidewalk length is calculated within the 0.5 mile radius of TOD station. The sidewalk data is obtained from Portland GIS open data website.

- Destination: Accessibility

The destination accessibility is a way to measure the interaction potential between TOD stations using a gravity equation. In this study two different accessibility measures were used. The job accessibility applied in the hedonic model is provided by Hansen. Another more complicated accessibility measure provided by Shen (1998) is used in the cluster analysis, which will be explained later. The job accessibility equation is as follows:

$$A_i = \sum_j O_j f(C_{ij}) \qquad (5)$$

Where

$A_i$ = the accessibility of people living within 0.5 mile of TOD station i;

$O_j$ = the number of employment opportunities within 0.5 mile of TOD station j;

$f(C_{ij})$ = the impedance function measuring the spatial separation between TOD station i and j. The impedance function is specified as below:

$$f(C_{ij}) = e^{-\beta\, C_{ij}} \qquad (6)$$

β = model coefficients which is large than 0.

$C_{ij}$ = network distance from light rail station i to light rail station j;

$e$ = base of the natural logarithms.

- Destination: Land Use Development Mix

Development mix is a way to measure whether the light rail station area is more employment centered or more residential centered. It is calculated as follows:

$$Development\ Mix = \frac{O_i}{O_i + P_i} \qquad (7)$$

Where

$O_i$ = job opportunities within 0.5 mile radius of light rail station i;

$P_i$ = population within 0.5 mile radius of light rail station i.

$O_i$ = obtained from the OnTheMap application, a census mapping tool.

$P_i$ = generated from the simple area weighting of census block groups intersecting the 0.5 mile radius of light rail stations.

- Distance to LRT Station

The distance to light rail stations are measured in two ways. One way is to measure the direct distance from each single family house to the closest light rail stations. The other way is to create dummy variable of distance bands with 0 – 0.25 mile, 0.25- 0.5 mile, 0.5- 0.75 mile, 0.75 -1 mile, and beyond 1 mile. The distance dummy variables are called 1/4_mi, 1/2_mi, 3/4_mi, 1_mi, outside_1_mi respectively.

4.2.6 Explanatory Independent Variable –TOD Typology (sub model)

The core of my research question is whether TOD context makes a difference for the value of distance. Here I adopt a similar approach to that proposed by Higgins and Kanaroglou (2018) to distill the light rail station area accessibility and built environment characteristics into a TOD typology.

Cluster analysis is conducted to generate the typology of TOD stations due to the difficulty of creating intuitive types of TOD stations. Cluster analysis categorizes a set of objects in a way that objects within the same group (cluster) are in a sense more similar to each other than to those in other groups. It is very useful when there is no preconceived idea of classification. I used the R Package factoextra and mclust to conduct the cluster analysis. According to Smith and Saito (2001) , the major challenges for cluster analysis are the proper selection of distance measure and linkage method, and standardization procedure. They recommended using a combination of the Z-score standardization of variables, the Ward's linkage method, and the Squared Euclidean distance measure.

The variables that are included in this model are accessibility, land use mix entropy, density, sidewalk length, development mix and ratio of the seven categories of land uses (office, institution, commercial, open space, industry, single-family and multifamily). The accessibility is calculated as the following equation (7). Previously, Zhang, Shen and Sussman (1998) categorize three basic types of accessibility measures: performance-based measures, potential-based measures and preference-based measures. Based on the criteria of transparency, comprehensiveness and intuitiveness, the gravity model accessibility measure makes great sense to measure employment accessibility. The destination accessibility of TOD station is measured as follows:

$$A_i = \sum_j \frac{O_j f(C_{ij})}{D_j}, \quad D_j = \sum_k P_k f(C_{kj}) \qquad (8)$$

Where

$A_i$ = the accessibility of people living within 0.5 mile of TOD station i;

$O_j$ = the number of employment opportunities within 0.5 mile of TOD station j;

$P_k$ = the number of people within 0.5 mile of TOD station k seeking to employment opportunities, k = 1, 2, ..., N;

$f(C_{kj})$ = the impedance function measuring the spatial separation between TOD station k and j. The impedance function is specified as below:

$$f(C_{kj}) = e^{-\beta C_{kj}} \qquad (9)$$

β = model coefficients which is large than 0.

$C_{kj}$ = network distance from light rail station k to light rail station j;

$e$ = base of the natural logarithms.

When this impedance function was applied, the coefficient b was set to 0.1, following the example given by Shen (1998).

All of the variables are transformed to a Z score in order to carry out cluster analysis. The function is as follows:

$$Z = \frac{x - \mu}{\sigma} \qquad (10)$$

However, after implementing the recommended the Ward's linkage method, and the Squared Euclidean distance measure by Smith and Saito (2001), the resulting groups didn't have meaningful interpretations in terms of land uses. So I instead used the Manhattan distance measure and the McQuitty/WPGMA linkage.

The following silhouette plot suggests that the category of four types is the best option for TOD typology across the 62 light rail stations within Portland City Boundary. This is due to housing data and zoning data availability.

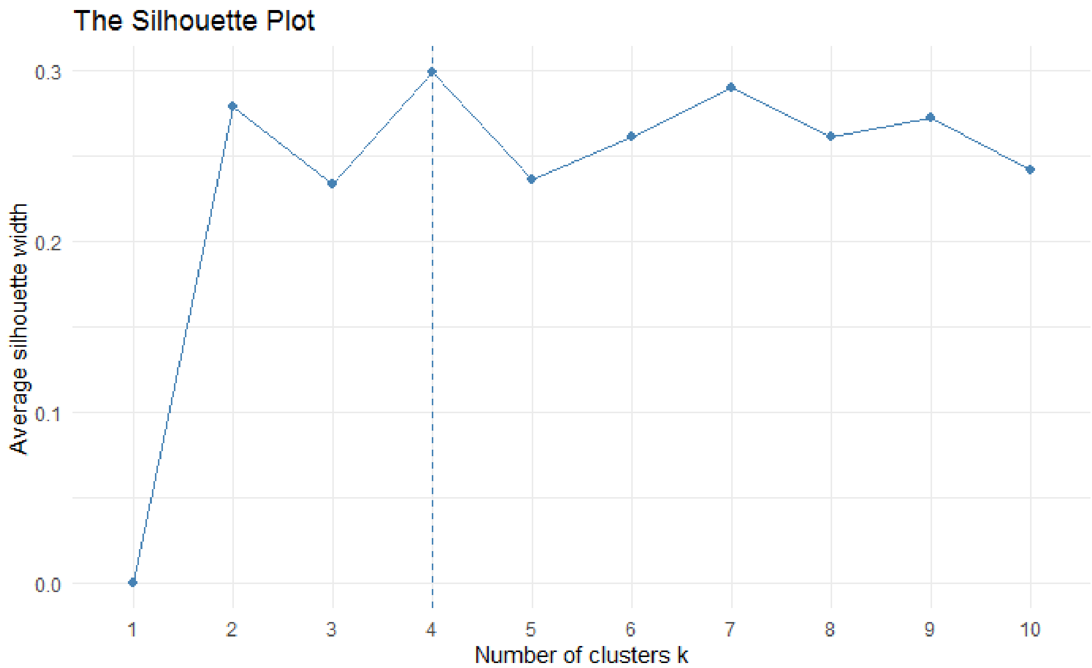



**Figure 4 The Silhuouette Plot of cluster analysis**

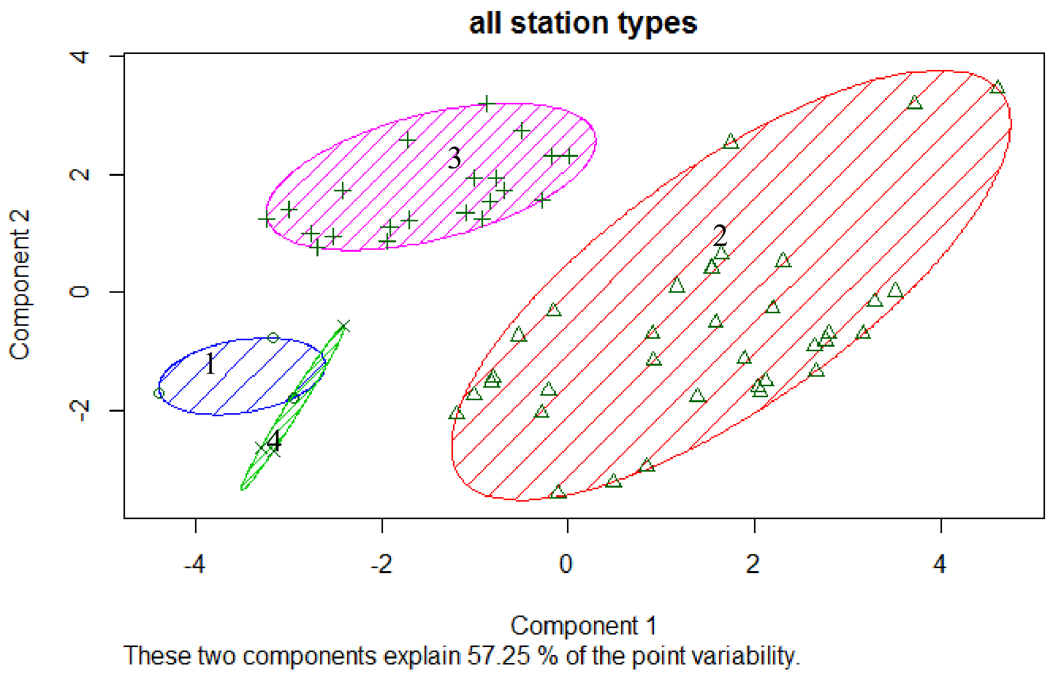



**Figure 5 Cluster Plot of Light Rail Station Typology**

Figure 5 shows the clustering of station types, which clearly suggest separation of different types. The following map in figure 6 shows the station types on map. The grey dots located along the riverside are type 1. The red dots near downtown are type 2. The yellow dots located in the middle of the city are type 3. The green dots located near the outskirt of city boundary are station type 4.

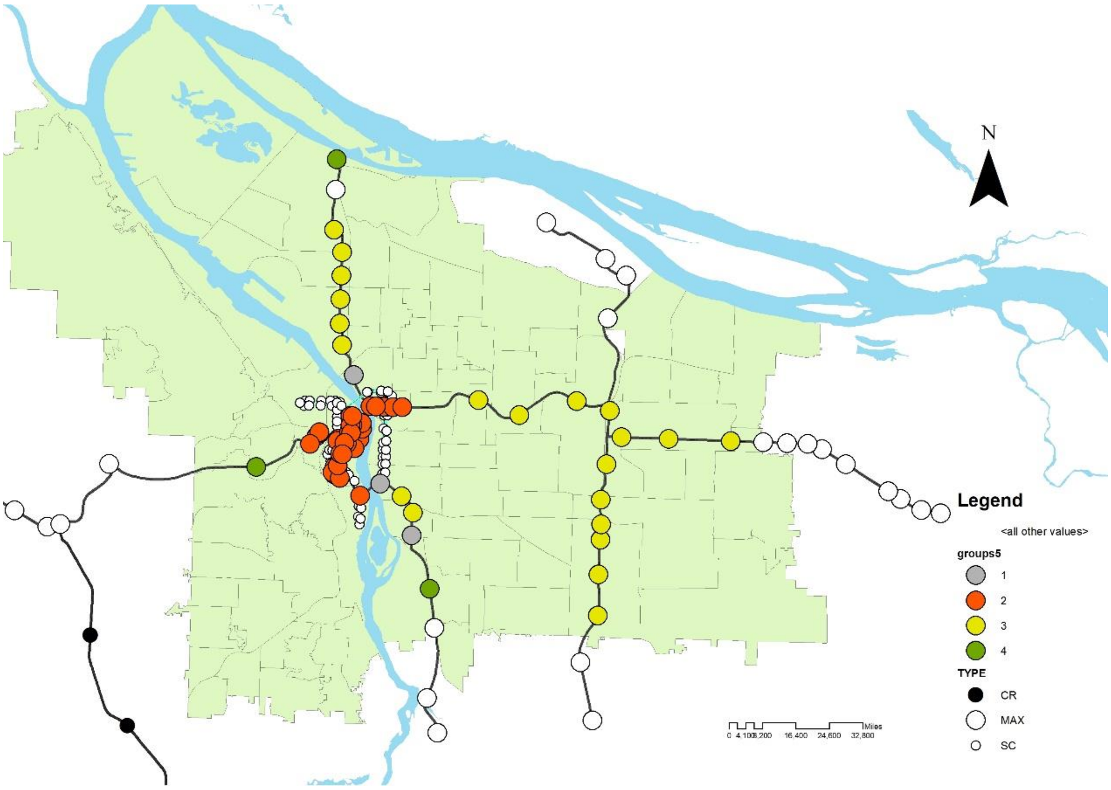



**Figure 6 Station Type Map**

4.2.7 Explanatory Independent Variable – Time-Series Variables

To account for changes of willingness to pay over time, I create time series dummy variables to capture different years of sale. Since there are multiple light rail lines constructed through the past 34 years, it is difficult to create a before-during-after model without isolating a particular rail line.

The following table 2 summarize the variables used in the regression model.

**Table 2 Summary of Variables**

| Variable Name | Description | Source |
|---|---|---|
| Neighborhood Characteristics | | |
| HSELA.Achievement | High School English language arts achievement. Score level 1-5. Level 5 means the performance meets state target for 2025. Level 4 means the performance is on track to meet state target for 2025. Level 3 means the performance is at or above state average but not on track. Level 2 means the performance is below state average. Level 1 means the bottom 10 percent statewide. | School accountability data from Oregon Government. |
| HSMath.Achievement | High School mathematics achievement. | Same as above |
| MSELA.Achievement | Middle School English language arts achievement | Same as above |
| MSMath.Achievement | Middle School mathematics achievement | Same as above |
| ESELA.Achievement | Elementary School English language arts achievement | Same as above |
| ESMath.Achievement | Elementary School mathematics achievement | Same as above |
| white_ratio | Ratio of white people in the census block group the house is located | 2014-2018 5-year American Community Survey |
| M_Hshd_income | Median Household Income in the past 12 months (2018) | 2018 1-year American Community Survey |
| crime_rate | Average number of crime rate from 2015-2019 （incidents per 100,000 population） | Portland government police data |
| Physical Characteristics | | |
| Full Bath | Number of full bathrooms | Extracted from Portland GIS open data |
| Half Bath | Number of half bathrooms | Extracted from Portland GIS open data |
| land_sf | Area of the parcel in square footage | Portland GIS open data |
| Building_area | Area of the building in square footage | Portland GIS open data |
| NUM_STORY | Number of stories of the house | Portland GIS open data |
| Elevation | Elevation of house in footage | Portland GIS open data |
| age | The age of building at sale time | Calculated by the year of sale minus the year built |
| Location Variable | | |
| Distance_frwy | Distance to freeway in footage | Measured in GIS using transportation data from Trimet |
| Distance_bus_stop | Distance to bus stop in footage | Measured in GIS using bus stop data from Trimet |
| Distance_open_space | Distance to open space in footage | Measured in GIS using Outdoor Recreation and Conservation Areas data from Oregon Metro |
| Distance_river | Distance to river in footage | Measured in GIS using river data from Oregon Metro |

| | | |
|---|---|---|
| Distance_school | Distance to school in footage | Measured in GIS using school site data from Oregon Metro |
| Distance_CBD (ft) | Euclidean distance from house to CBD area | Measured from each home to the Pioneer Square in Euclidean distance |
| Station Characteristics | | |
| Accessibility_shen (Cluster Analysis) | $A_i = \sum_j \frac{O_j f(C_{ij})}{D_j}, \quad D_j = \sum_k P_k f(C_{kj})$ <br> $A_i$ is the accessibility of people living within 0.5 mile of TOD station i; $O_j$ is the number of employment opportunities within 0.5 mile of TOD station j;$P_k$ is the number of people within 0.5 mile of TOD station k seeking to employment opportunities, k = 1, 2, ..., N; $f(C_{kj})$ is the impedance function measuring the spatial separation between TOD station k and j. The impedance function is specified as below: <br> $f(C_{kj}) = e^{-\beta C_{kj}}$ <br> β is model coefficient which is large than 0. $C_{kj}$ is network distance from light rail station k to light rail station j; $e$ is base of the natural logarithms. | Employment data is gathered from Census application OnTheMap. Population data is calculated by area weighted method of census block groups that intersected with the 0.5 mile radius of light rail station. Census block group population data is from 2014-2018 5-year American Community Survey. |
| Job Accessibility (Hedonic Model) | $A_i = \sum_j O_j f(C_{ij})$ <br> Where <br> $A_i$ is the accessibility of people living within 0.5 mile of TOD station i; $O_j$ is the number of employment opportunities within 0.5 mile of TOD station j; $f(C_{ij})$ is the impedance function measuring the spatial separation between TOD station i and j. The impedance function is specified as below: <br> $f(C_{ij}) = e^{-\beta C_{ij}}$ <br> β is model coefficient which is large than 0. $C_{ij}$ is network distance from light rail station i to light rail station j; $e$ is base of the natural logarithms. | |
| mix_entropy | $\text{Land Use Entropy} = \frac{-\sum(P_{ij} \times \ln(P_{ij}))}{\ln(N_j)}$ <br> Where Lij is the percent of area j allocated to land use type i, and Nj is the number of land use types in area j. The mix entropy is only calculated at TOD station level within 0.5 mile radius. | Land use data is obtained from zoning map from Portland Open Data website. It is a substitution for existing land use map, which is not available online. Some adjustment are made according to the comprehensive plan description in the zoning data. Land use ratios are calculated only within the 0.5 mile radius of light rail station. |
| sidewalk_length (ft) | Sidewalk length within 0.5 mile radius of the TOD station | Measured by sidewalk data from Oregon Metro |
| density_acre (jobs + population/ acre) | Total of population and jobs divided by the TOD area (0.5 mile radius) | Employment data is gathered from Census application OnTheMap. Population data is calculated by area weighted method of census block groups that intersected with the 0.5 mile radius of light rail station. Census block group population data is from 2014-2018 5-year American |

| | | |
|---|---|---|
| | | Community Survey. |
| development_mix | A ratio of Employment to Population + Employment, which reflects the balance between population and employment in a station area. | Same as above |
| ST_type | Dummy variable for station type with four levels-ST_type1, ST_type2, ST_type3, and ST_type4. | Identified through cluster analysis of accessibility, land use ratios, sidewalk length, development mix, density and land use entropy. |
| TOD Distance | | |
| Distance_LRT_stop | Distance to light rail stations. | Measured from each home to the nearest light rail station in Euclidean distance |
| TOD dummy | Dummy variable for house located within 0.25- 0.5 mile of TOD station. A total of 5 levels are included: distance bands with 0 – 0.25 mile, 0.25- 0.5 mile, 0.5- 0.75 mile, 0.75 -1 mile, and beyond 1 mile; each are called 1/4_mi, 1/2_mi, 3/4_mi, 1_mi, outside_1_mi, respectively. | Measured the distance of building data and light rail station data. Building data is from Portland Open Data. Light rail station data is from Trimet |

## Chapter 5 Modeling Results

### 5.1 Descriptive Statistics

Table 3 describes the context factors relevant for the classification of LRT station typology. Included are factors describing light rail accessibility of catchment area, factors describing land use characteristics of the catchment area, and factors describing development intensity. The rail stop data within Portland includes 138 stations, with 62 LRT stops and 76 streetcar stops. Although the two transit types look similar in many ways, there are differences in service area, station spacing, right-of-way, capacity and speed. This study only investigates the LRT stations and houses near LRT stations.

I generate four LRT station types from cluster analysis. The first station type is the urban manufacturing. This station type is featured with major land use in industrial land use, with some office and commercial land uses. The accessibility is generally low but the land use mixture entropy is high. The station area is somewhat walkable with relatively low density.

The second station type is the urban commercial. The station type is featured with major land use of commercial land, followed by multifamily and open space. The station area is highly accessible and highly walkable. The jobs near the station generally outnumber residents. The station area generally has a high density.

The third station type is the residential neighborhood. The station type has a high ratio of residential uses including single family and multifamily, with some commercial and open space uses. The area is highly accessible, but the walkability is moderate. It is usually of medium density.

The fourth station type is the park TOD. The main land use near the station is open space. The surrounding area has low accessibility, low walkability and low density.

**Table 3 Summary Statistics of LRT Station Types**

| | ST_Type1 (N=3) | ST_Type2 (N=35) | ST_Type3 (N=21) | ST_Type4 (N=3) | Overall (N=62) |
|---|---|---|---|---|---|
| **accessibility shen** | | | | | |
| Mean (SD) | 0.500 (0.236) | 1.76 (0.736) | 1.67 (0.408) | 0.443 (0.500) | 1.60 (0.713) |
| Median [Min, Max] | 0.436 [0.302, 0.761] | 1.67 [0.617, 3.92] | 1.72 [1.00, 2.33] | 0.161 [0.146, 1.02] | 1.62 [0.146, 3.92] |
| **mix entropy** | | | | | |
| Mean (SD) | 0.735 (0.0791) | 0.600 (0.0557) | 0.744 (0.0840) | 0.483 (0.111) | 0.650 (0.104) |
| Median [Min, Max] | 0.742 [0.653, 0.811] | 0.604 [0.493, 0.700] | 0.750 [0.587, 0.870] | 0.533 [0.356, 0.561] | 0.644 [0.356, 0.870] |
| **office ratio** | | | | | |
| Mean (SD) | 0.174 (0.0862) | 0.0756 (0.0843) | 0.0458 (0.0560) | 0 (0) | 0.0666 (0.0788) |
| Median [Min, Max] | 0.145 [0.107, 0.272] | 0.0606 [0, 0.349] | 0.0203 [0, 0.162] | 0 [0, 0] | 0.0404 [0, 0.349] |
| **institution ratio** | | | | | |
| Mean (SD) | 0.0309 (0.0536) | 0.000871 (0.00344) | 0.0226 (0.0328) | 0.00633 (0.0110) | 0.00995 (0.0241) |
| Median [Min, Max] | 0 [0, 0.0928] | 0 [0, 0.0160] | 0.00700 [0, 0.133] | 0 [0, 0.0190] | 0 [0, 0.133] |
| **open ratio** | | | | | |
| Mean (SD) | 0.0913 (0.0566) | 0.164 (0.0885) | 0.112 (0.0718) | 0.421 (0.0484) | 0.155 (0.103) |
| Median [Min, Max] | 0.0875 [0.0367, 0.150] | 0.150 [0.0265, 0.368] | 0.107 [0.00730, 0.224] | 0.397 [0.390, 0.477] | 0.140 [0.00730, 0.477] |
| **commercial ratio** | | | | | |
| Mean (SD) | 0.129 (0.0686) | 0.523 (0.0979) | 0.155 (0.0746) | 0.118 (0.167) | 0.360 (0.209) |
| Median [Min, Max] | 0.107 [0.0746, 0.206] | 0.563 [0.194, 0.634] | 0.137 [0.0242, 0.307] | 0.0455 [0, 0.309] | 0.417 [0, 0.634] |
| **multifamily ratio** | | | | | |
| Mean (SD) | 0.0329 (0.00951) | 0.169 (0.117) | 0.220 (0.0977) | 0.00810 (0.0140) | 0.172 (0.118) |
| Median [Min, Max] | 0.0368 [0.0221, 0.0399] | 0.154 [0.00630, 0.498] | 0.209 [0.0576, 0.372] | 0 [0, 0.0243] | 0.162 [0, 0.498] |
| **single family ratio** | | | | | |
| Mean (SD) | 0.0771 (0.105) | 0.0292 (0.0595) | 0.375 (0.0983) | 0.345 (0.298) | 0.164 (0.190) |
| Median [Min, Max] | 0.0279 [0.00620, 0.197] | 0 [0, 0.225] | 0.393 [0.196, 0.551] | 0.511 [0, 0.523] | 0.0589 [0, 0.551] |
| **industrial ratio (see** | | | | | |
| Mean (SD) | 0.464 (0.125) | 0.0379 (0.0547) | 0.0698 (0.112) | 0.102 (0.173) | 0.0724 (0.125) |
| Median [Min, Max] | 0.399 [0.385, 0.608] | 0 [0, 0.159] | 0.0148 [0, 0.386] | 0.00360 [0, 0.302] | 0.00770 [0, 0.608] |
| **sidewalk length (ft)** | | | | | |
| Mean (SD) | 94600 (20200) | 391000 (136000) | 81800 (37600) | 30400 (45300) | 254000 (188000) |
| Median [Min, Max] | 90000 [77100, 117000] | 417000 [78800, 595000] | 70400 [25800, 161000] | 5950 [2570, 82800] | 217000 [2570, 595000] |
| **development mix (jobs/ (population + jobs))** | | | | | |
| Mean (SD) | 0.443 (0.0951) | 0.624 (0.0828) | 0.656 (0.0603) | 0.697 (0.208) | 0.629 (0.0944) |
| Median [Min, Max] | 0.478 [0.336, 0.516] | 0.625 [0.478, 0.824] | 0.657 [0.504, 0.783] | 0.711 [0.483, 0.898] | 0.638 [0.336, 0.898] |
| **density (jobs + population/ acre)** | | | | | |
| Mean (SD) | 8.50 (2.26) | 24.6 (9.27) | 16.2 (4.55) | 4.86 (4.99) | 20.0 (9.51) |
| Median [Min, Max] | 7.46 [6.94, 11.1] | 25.5 [7.56, 47.6] | 14.1 [9.42, 25.0] | 2.73 [1.29, 10.6] | 18.5 [1.29, 47.6] |

*** Ratio is the particular land use category area within 0.5 mile of the light rail station divided by the total land area of the 0.5 mile radius.

The following table 4 provides a descriptive result of the independent and dependent variables, including means, standard deviation, minimum, and maximum.

**Table 4 Descriptive Statistics of Hedonic Model**

| | ST_Type 1 (N=824) | ST_Type 2 (N=524) | ST_Type 3 (N=42504) | ST_Type 4 (N=4385) | Overall (N=48237) |
|---|---|---|---|---|---|
| **price_real** | | | | | |
| Mean (SD) | 414984 (159753) | 956932 (660906) | 402547 (191118) | 605410 (327704) | 427223 (231496) |
| Median [Min, Max] | 409468 [21807, 1916708] | 792519 [20654, 6381407] | 363034 [20125, 4925902] | 531972 [21438, 4506313] | 376748 [20125, 6381407] |
| **ESELA.Achievement** | | | | | |
| Level 1 | 58 (7.0%) | 0 (0%) | 8296 (19.5%) | 79 (1.8%) | 8433 (17.5%) |
| Level 2 | 548 (66.5%) | 0 (0%) | 19928 (46.9%) | 147 (3.4%) | 20623 (42.8%) |
| Level 3 | 148 (18.0%) | 182 (34.7%) | 6136 (14.4%) | 590 (13.5%) | 7056 (14.6%) |
| Level 4 | 9 (1.1%) | 8 (1.5%) | 4173 (9.8%) | 2454 (56.0%) | 6644 (13.8%) |
| Level 5 | 61 (7.4%) | 334 (63.7%) | 3971 (9.3%) | 1115 (25.4%) | 5481 (11.4%) |
| **HSMath.Achievement** | | | | | |
| Level 1 | 0 (0%) | 0 (0%) | 7085 (16.7%) | 927 (21.1%) | 8012 (16.6%) |
| Level 2 | 824 (100%) | 291 (55.5%) | 35382 (83.2%) | 2531 (57.7%) | 39028 (80.9%) |
| Level 3 | 0 (0%) | 233 (44.5%) | 37 (0.1%) | 927 (21.1%) | 1197 (2.5%) |
| **white_ratio** | | | | | |
| Mean (SD) | 0.829 (0.0560) | 0.870 (0.0715) | 0.767 (0.115) | 0.879 (0.0609) | 0.779 (0.115) |
| Median [Min, Max] | 0.835 [0.595, 0.908] | 0.881 [0.609, 0.972] | 0.785 [0.401, 0.997] | 0.880 [0.519, 0.988] | 0.799 [0.401, 0.997] |
| **M_Hshd_income (dollars)** | | | | | |
| Mean (SD) | 54155 (16014) | 95153 (43123) | 66634 (24194) | 102811 (46551) | 70020 (29262) |

| | | | | | |
|---|---|---|---|---|---|
| Median [Min, Max] | 55391 [34085, 144844] | 92829 [17772, 214333] | 60735 [15174, 176125] | 83696 [40377, 220903] | 62303 [15174, 220903] |
| **crime_rate (incidents per 100,000 population)** | | | | | |
| Mean (SD) | 8656 (5396) | 7691 (6481) | 7681 (2782) | 4437 (2980) | 7403 (3078) |
| Median [Min, Max] | 7621 [5192, 27444] | 4576 [2289, 22066] | 7611 [2597, 29199] | 5391 [1587, 28643] | 7585 [1587, 29199] |
| **full_bath** | | | | | |
| Mean (SD) | 1.58 (0.689) | 2.21 (0.936) | 1.58 (0.665) | 1.93 (0.781) | 1.62 (0.690) |
| Median [Min, Max] | 1.00 [0, 6.00] | 2.00 [0, 8.00] | 1.00 [0, 5.00] | 2.00 [0, 6.00] | 2.00 [0, 8.00] |
| **half_bath** | | | | | |
| Mean (SD) | 0.152 (0.366) | 0.565 (0.653) | 0.202 (0.426) | 0.309 (0.500) | 0.215 (0.438) |
| Median [Min, Max] | 0 [0, 2.00] | 0 [0, 3.00] | 0 [0, 6.00] | 0 [0, 4.00] | 0 [0, 6.00] |
| **Land area (sq ft)** | | | | | |
| Mean (SD) | 5227 (1422) | 6778 (4203) | 6113 (2902) | 8146 (5149) | 6290 (3228) |
| Median [Min, Max] | 5000 [2400, 17850] | 5000 [960, 32100] | 5000 [900, 43560] | 6200 [1487, 43560] | 5000 [900, 43560] |
| **Building area (sq ft)** | | | | | |
| Mean (SD) | 1471 (717) | 2516 (1253) | 1526 (626) | 1915 (877) | 1572 (681) |
| Median [Min, Max] | 1355 [576, 10648] | 2254 [666, 8117] | 1416 [360, 8584] | 1726 [528, 9308] | 1442 [360, 10648] |
| **Elevation (ft)** | | | | | |
| Mean (SD) | 207 (251) | 268 (234) | 177 (135) | 341 (256) | 193 (161) |
| Median [Min, Max] | 103 [1.00, 903] | 101 [9.00, 906] | 160 [1.00, 903] | 238 [1.00, 909] | 162 [1.00, 909] |
| **NUM_STORY** | | | | | |
| Mean (SD) | 1.27 (0.484) | 1.77 (0.569) | 1.23 (0.433) | 1.30 (0.484) | 1.25 (0.444) |
| Median [Min, Max] | 1.00 [1.00, 3.00] | 2.00 [1.00, 3.00] | 1.00 [1.00, 4.00] | 1.00 [1.00, 4.00] | 1.00 [1.00, 4.00] |
| **Year_built** | | | | | |
| Mean (SD) | 1931 (23.0) | 1925 (27.0) | 1939 (24.8) | 1946 (23.7) | 1939 (24.8) |
| Median [Min, Max] | 1925 [1884, 2014] | 1914 [1880, 2005] | 1938 [1870, 2017] | 1947 [1884, 2009] | 1938 [1870, 2017] |
| **age** | | | | | |
| Mean (SD) | 77.2 (23.8) | 83.0 (27.3) | 69.4 (25.1) | 62.6 (24.1) | 69.1 (25.2) |
| Median [Min, Max] | 83.0 [0, 127] | 92.0 [1.00, 136] | 73.0 [0, 148] | 64.0 [0, 123] | 72.0 [0, 148] |
| **Distance_frwy (ft)** | | | | | |
| Mean (SD) | 9577 (3209) | 2355 (1102) | 5344 (3507) | 8678 (3720) | 5687 (3685) |
| Median [Min, Max] | 10597 [1775, 13777] | 2357 [175, 5220] | 4907 [72.7, 25369] | 8653 [186, 31116] | 5211 [72.7, 31116] |
| **Distance_arterial (ft)** | | | | | |
| Mean (SD) | 496 (312) | 630 (572) | 645 (463) | 672 (576) | 645 (474) |
| Median [Min, Max] | 433 [40.7, 1358] | 467 [48.7, 3149] | 542 [37.6, 3444] | 516 [45.6, 3101] | 538 [37.6, 3444] |
| **Distance_bus_stop (ft)** | | | | | |
| Mean (SD) | 510 (256) | 598 (397) | 703 (573) | 645 (485) | 694 (561) |
| Median [Min, Max] | 462 [53.9, 1150] | 476 [62.9, 1917] | 593 [32.0, 8765] | 515 [44.2, 3127] | 581 [32.0, 8765] |
| **Distance_open_space (ft)** | | | | | |
| Mean (SD) | 961 (470) | 1103 (875) | 964 (572) | 816 (538) | 952 (573) |
| Median [Min, Max] | 896 [93.0, 2329] | 822 [20.7, 3175] | 887 [4.09, 2986] | 711 [0, 2414] | 873 [0, 3175] |
| **Distance_river (ft)** | | | | | |
| Mean (SD) | 5598 (2947) | 6507 (1180) | 8571 (4354) | 7970 (4775) | 8443 (4377) |
| Median [Min, Max] | 5541 [484, 11061] | 6581 [2625, 8996] | 8508 [58.7, 18519] | 7756 [32.9, 18461] | 8264 [32.9, 18519] |
| **Distance_school (ft)** | | | | | |
| Mean (SD) | 974 (473) | 1003 (654) | 1013 (634) | 1317 (838) | 1040 (659) |
| Median [Min, Max] | 962 [56.5, 2140] | 910 [52.9, 3294] | 935 [0, 7169] | 1239 [31.2, 12866] | 956 [0, 12866] |
| **Distance_CBD (ft)** | | | | | |
| Mean (SD) | 15655 (2803) | 8985 (2746) | 24937 (8582) | 18588 (4491) | 24028 (8603) |
| Median [Min, Max] | 15774 [9636, 20661] | 9594 [3869, 13618] | 23132 [7738, 51084] | 19215 [6723, 48147] | 22177 [3869, 51084] |
| **LRT_dummy** | | | | | |
| 1/4_mi | 47 (5.7%) | 12 (2.3%) | 2289 (5.4%) | 113 (2.6%) | 2461 (5.1%) |
| 1/2_mi | 114 (13.8%) | 76 (14.5%) | 7248 (17.1%) | 579 (13.2%) | 8017 (16.6%) |
| 3/4_mi | 209 (25.4%) | 192 (36.6%) | 6907 (16.3%) | 726 (16.6%) | 8034 (16.7%) |
| 1_mi | 148 (18.0%) | 180 (34.4%) | 6710 (15.8%) | 468 (10.7%) | 7506 (15.6%) |
| outside_1_mi | 306 (37.1%) | 64 (12.2%) | 19350 (45.5%) | 2499 (57.0%) | 22219 (46.1%) |
| **Distance_LRT_stop (ft)** | | | | | |
| Mean (SD) | 4798 (2508) | 3789 (1232) | 5340 (3081) | 6499 (3713) | 5419 (3144) |
| Median [Min, Max] | 4368 [123, 10528] | 3859 [642, 7023] | 4875 [169, 24982] | 6281 [677, 31186] | 4916 [123, 31186] |
| **mix_entropy** | | | | | |
| Mean (SD) | 0.799 (0.0413) | 0.623 (0.0250) | 0.721 (0.0869) | 0.462 (0.0872) | 0.697 (0.115) |
| Median [Min, Max] | 0.811 [0.653, 0.811] | 0.604 [0.604, 0.674] | 0.742 [0.587, 0.870] | 0.533 [0.356, 0.561] | 0.727 [0.356, 0.870] |
| **Job_accessibility** | | | | | |
| Mean (SD) | 59276 (9885) | 159923 (61322) | 64835 (48439) | 43987 (28721) | 63878 (48189) |
| Median [Min, Max] | 62069 [24336, 62069] | 112043 [112043, 686553] | 58861 [8887, 142956] | 68400 [6226, 68400] | 62069 [6226, 686553] |

The average inflated value of a single family house within the sample is $427,223, with a fair degree of variation. The homes near type 4 (open space TOD) has highest mean price while the single family houses near station type 2 (urban commercial) has the highest median value. The average crime rate is 7,403 incidents per 100,000 population. The type 1 station (urban manufacturing) seems to have the highest mean and median crime rate. About 78% of the population is white, within block groups where the single family houses are located. The median household income is $70,020, about $6,000 higher than the Portland median household income. Generally, the median household income is high near station type 2 and station type 4.

The average land area for single family houses is 6,290 square ft. The mean building area is 1,572 square footage. The average number of stories is 1.25. The mean number of full bathrooms is 1.62 and the mean number of half bathroom is 0.22. The mean age at sale is 69 years old. The average year built is 1939. Single family houses near station type 4 tend to have larger land parcels, while single family houses near station type 2 tend to have more stories and larger building area.

Distance to natural amenities and other transportation facilities varies a lot. The mean distance from single family house to arterial road is 645 ft. The average distance to freeway is 5,687 ft. The mean distance to bus stop is 694 ft. The mean distance to light rail stations is 5,419 ft, much higher than distance to bus stops. The mean distance to open space is 952 ft. The mean distance to river is 8,443 ft. The average distance to school is 1,040 ft. The mean distance to CBD area is 24,028 ft.

The number of observations situated within 0.25 mile, 0.25-0.5 mile, 0.5 – 0.75 mile, 0.75-1 mile, and beyond 1 mile distance of light rail stations are 2461, 8017, 8034, 7506 and 22219 respectively. The average land use mix entropy of the various TOD stations is about 0.697, which means a relatively diverse land use near light rail stations. The job accessibility measure has an average number of 63878, with much variation among light rail stations.

## 5.2 Cross-Section Regression Results

I created 8 pool models as illustrated in Table 5. Pool models 1-6 are cross-section models, and pool models 7-8 are longitudinal models. The models followed an order from simple to complicated measures. Model 1 and 2 are the simplest models with the only difference being measure of distance to light rail stations. Model 3 and model 4 are based on model 1 and model 2, with station types.

Additional interaction terms between station types and distance to light rail stations are added to model 5 and model 6, in addition to variables in model 3 and 4. Lastly, time series dummy variables are included in model 7 and model 8 to examine the impact of time.

**Table 5 Comparison of Pool Models 1-8**

| Models | Distance (continuous vs dummy) | Station Type (Yes vs No) | Interaction Term (Yes vs No) | Time Series (Yes vs No) |
|---|---|---|---|---|
| Model 1 | Continuous | No | No | No |
| Model 2 | Dummy | No | No | No |
| Model 3 | Continuous | Yes | No | No |
| Model 4 | Dummy | Yes | No | No |
| Model 5 | Continuous | Yes | Yes | No |
| Model 6 | Dummy | Yes | Yes | No |
| Model 7 | Continuous | Yes | Yes | Yes |
| Model 8 | Dummy | Yes | Yes | Yes |

The results of four cross-sectional pool models 1-4 are presented in Table 6. The adjusted R-squares show that all the four models can explain over 61% of variation in the dependent variable.

**Table 6 Regression Models 1-4**

| | Dependent variable: | | | |
|---|---|---|---|---|
| | price_real | | | |
| | (1) | (2) | (3) | (4) |
| ESELA.AchievementLevel 2 | 14,907.550*** | 12,234.240*** | 13,929.340*** | 11,059.860*** |
| | (2,085.336) | (2,083.168) | (2,083.080) | (2,080.900) |

| | | | | |
|---|---|---|---|---|
| ESELA.AchievementLevel 3 | 20,462.530*** | 22,512.450*** | 14,270.110*** | 15,397.660*** |
| | (2,834.556) | (2,856.453) | (2,855.799) | (2,885.326) |
| ESELA.AchievementLevel 4 | 42,525.930*** | 44,695.160*** | 49,264.940*** | 49,576.600*** |
| | (3,075.292) | (3,080.141) | (3,198.139) | (3,201.946) |
| ESELA.AchievementLevel 5 | 111,403.300*** | 112,581.800*** | 103,827.000*** | 105,031.300*** |
| | (3,376.007) | (3,389.558) | (3,371.293) | (3,387.711) |
| HSMath.AchievementLevel 2 | 15,195.560*** | 16,796.150*** | 10,285.250*** | 13,429.540*** |
| | (2,154.627) | (2,163.730) | (2,167.769) | (2,177.838) |
| HSMath.AchievementLevel 3 | 155,611.700*** | 154,139.100*** | 111,569.300*** | 112,499.400*** |
| | (5,823.908) | (5,834.333) | (6,098.578) | (6,116.004) |
| white_ratio | 99,983.560*** | 95,327.250*** | 101,160.500*** | 94,473.770*** |
| | (7,626.783) | (7,617.065) | (7,660.694) | (7,651.471) |
| M_Hshd_income | 1.039*** | 0.993*** | 1.096*** | 1.055*** |
| | (0.033) | (0.033) | (0.033) | (0.033) |
| crime_rate | -1.643*** | -1.426*** | -2.581*** | -2.225*** |
| | (0.269) | (0.274) | (0.274) | (0.279) |
| full_bath | 42,487.070*** | 42,536.230*** | 42,463.360*** | 42,522.200*** |
| | (1,299.849) | (1,299.176) | (1,291.791) | (1,292.229) |
| half_bath | 41,362.970*** | 41,388.470*** | 40,045.150*** | 40,180.330*** |
| | (1,736.813) | (1,736.055) | (1,726.293) | (1,726.982) |
| land_area | 6.050*** | 6.075*** | 5.991*** | 5.999*** |
| | (0.259) | (0.259) | (0.258) | (0.258) |

| | | | | |
|---|---|---|---|---|
| Building area | 108.352*** | 107.360*** | 106.915*** | 106.089*** |
| | (1.477) | (1.478) | (1.469) | (1.471) |
| Elevation | -4.535 | 0.207 | 1.074 | 4.407 |
| | (4.755) | (4.763) | (4.759) | (4.767) |
| NUM_STORY | 23,746.750*** | 22,846.370*** | 22,995.950*** | 22,163.630*** |
| | (1,825.024) | (1,823.309) | (1,813.340) | (1,813.267) |
| age | 478.347*** | 494.164*** | 479.156*** | 498.202*** |
| | (33.057) | (33.026) | (32.878) | (32.865) |
| Distance_frwy | 2.193*** | 1.008*** | 3.531*** | 1.916*** |
| | (0.278) | (0.250) | (0.305) | (0.272) |
| Distance_arterial | 15.779*** | 13.976*** | 15.702*** | 14.646*** |
| | (1.566) | (1.590) | (1.558) | (1.586) |
| Distance_bus_stop | 14.596*** | 13.858*** | 13.409*** | 12.670*** |
| | (1.385) | (1.395) | (1.383) | (1.398) |
| Distance_open_space | 2.987** | 2.863** | -0.507 | -0.140 |
| | (1.290) | (1.295) | (1.289) | (1.294) |
| Distance_river | 0.607*** | 0.369* | 0.849*** | 0.702*** |
| | (0.198) | (0.199) | (0.207) | (0.208) |
| Distance_school | -3.250*** | -3.725*** | -1.853* | -2.621** |
| | (1.125) | (1.123) | (1.120) | (1.118) |
| Distance_CBD | -4.111*** | -4.089*** | -4.730*** | -4.622*** |
| | (0.157) | (0.157) | (0.170) | (0.172) |

| | | | | |
|---|---|---|---|---|
| Distance_LRT_stop | -3.282*** | | -4.502*** | |
| | (0.343) | | (0.362) | |
| LRT_dummy1/4_mi | | -1,038.049 | | 7,734.382** |
| | | (3,676.995) | | (3,760.226) |
| LRT_dummy1/2_mi | | 6,454.443** | | 12,921.250*** |
| | | (2,521.966) | | (2,637.655) |
| LRT_dummy3/4_mi | | 18,107.160*** | | 20,586.170*** |
| | | (2,313.351) | | (2,383.011) |
| LRT_dummy1_mi | | 23,412.850*** | | 23,059.530*** |
| | | (2,208.630) | | (2,230.568) |
| mix_entropy | -16,403.950** | -19,982.270*** | -58,214.570*** | -51,946.030*** |
| | (7,453.541) | (7,453.198) | (9,383.592) | (9,430.981) |
| Job_accessibility | 0.219*** | 0.240*** | 0.016 | 0.061** |
| | (0.022) | (0.022) | (0.025) | (0.025) |
| ST_type1 | | | 15,173.980** | 14,030.440** |
| | | | (6,719.135) | (6,734.000) |
| ST_type2 | | | 203,822.700*** | 188,633.800*** |
| | | | (8,228.301) | (8,253.631) |
| ST_type3 | | | 31,858.720*** | 24,086.550*** |
| | | | (4,398.532) | (4,418.955) |
| Constant | -26,526.690** | -35,493.500*** | 10,770.920 | -9,924.011 |
| | (11,196.950) | (11,282.710) | (11,374.120) | (11,338.660) |

| | | | | |
|---|---|---|---|---|
| Observations | 48,237 | 48,237 | 48,237 | 48,237 |
| $R^2$ | 0.587 | 0.587 | 0.592 | 0.592 |
| Adjusted $R^2$ | 0.587 | 0.587 | 0.592 | 0.592 |
| Residual Std. Error | 148,826.300 (df = 48210) | 148,733.000 (df = 48207) | 147,847.000 (df = 48207) | 147,881.200 (df = 48204) |
| F Statistic | 2,634.513*** (df = 26; 48210) | 2,367.129*** (df = 29; 48207) | 2,415.567*** (df = 29; 48207) | 2,187.495*** (df = 32; 48204) |
| Note: | *p<0.1; **p<0.05; ***p<0.01 | | | |

Most of the coefficients for physical structure characteristics are highly significant at 0.01 level. In model 1, people are willing to pay for $6.05 for each additional square foot of parcel land, $108.35 for one square foot increase of building area. Controlling for the building area, adding an additional full bath corresponds to a $42,487 increase in market value while having an extra half bath results in a $41,363 rise in housing price. The reason a full bath is worth more is because it could be that a full-bath provides more functions. Each additional building story translates into a $ 23,747 increase in house price. One foot higher elevation of building surface is associated with -$4.35 in model, although the relationship is not significant across model 1 to model 8. After controlling all other structure variables, the age is positively associated with single family home price with a coefficient of $478.

All the neighborhood socio-economic factors are significant at 0.01 level across the four models. All else being equal, a 1% increase of white population in the census block translates into $999.83 in the average home sales price. A $1,000 rise in median household income of the census block group increases the sales price by $1,039. Each additional crime incident per 100,000 people will decrease the average home price by $1.64. All others being equal, locating at the school zone with elementary school ELA achievement at level 2, level 3, level 4 and level 5 will increase the average single-family home price by $14908, $20463, $42526 and $111403 respectively, with level 1 as reference.

On the other hand, school zones with high school math achievement of level 2 and level 3 add $ 14,474 and $223,714, respectively, compared to level 1 However, the coefficient of math achievement at level 3 seems to be unreasonably high and varies across the four models. It might require a closer examination. It should be noted that of the 131 school districts in the city of Portland, only one has high school math achievement of level 4 and none has achieved level 5. The fact that no housing data of level 4 is found in the sample could be due to sample bias, as I only searched for single family housing within 6 miles of light rail stations.

Most of the locational factors are significantly associated with single-family housing price. In model 1, extra foot away from the arterial, freeway, bus stop, and river all translate into an increase in single-family home price. Distance from the arterials seems to have the largest marginal effect. Distance to open space is significant and positive in model 1 and 2, but loses the significance and has various signs in other models. The positive sign of distance to open space is surprising, as Bartholomew and Ewing (2011) state that the there is general willingness to pay for closer location to open space in their literature review. It is possible that how people value proximity to open space is dependent on the type of open space.

After controlling for other factors, we can look at the TOD variables and station variables. Both the distance to light rail station variable and dummy variables are significant at 0.01 level. The negative sign for continuous distance variable suggests a larger distance from the light rail station is a drop in price. Every additional 1000 feet away from light rail station appears to decrease house price by $3282. The dummy variables in model 2 suggest that compared to single family houses outside a 1 mile radius of light rail stations, those located within 1/4 mile, 1/4-1/2 mile, 1/2- 3/4 mile and 3/4 – 1 mile distance bands have marginal effects of -$1,038, $6,454, $18,107, and $23,412 respectively, all others being equal. Model 4 has similar results, with the only difference that the marginal effects of a 1/4 mile band being positive after controlling for station types. Both model 2 and 4 suggest that single family homes located within 3/4 – 1 mile tend to have highest TOD premium when other factors are the same.

When the accessibility value increases by 1, housing value, on average, increases $0.22 while controlling for the effects of other variables. The mix entropy of land uses near 0.5 mile radius of the light rail station shows a negative coefficient at -$16,403 with 0.01 significance level across model 1 to model 4. This suggests single family housing price is more likely to increase with more homogenous land uses.

The coefficients of light rail station types are also consistent across models 3 and 4. Compared to type 4, of which primary land use is green space, type 1, type 2, and type 3 have a relatively higher premium of $15,174, $203,822, and $31,858 after controlling for job accessibility and other factors in model 3. The type 2 is the most walkable and dense type among the four and scattered around the downtown Portland. After controlling for distance between home and CBD, the station type 2 still have a surprisingly high premium. It is possible that the high accessibility, dense environment, mixed land uses, and walkable environment create a synergy effect on housing price.

5.3 Longitudinal Regression Result

Results of the longitudinal case study can be found in Model 6 and Model 7 from Table 10.

**Table 7 Models 5-8**

| | Dependent variable: | | | |
|---|---|---|---|---|
| | price_real | | | |
| | (5) | (6) | (7) | (8) |
| ESELA.AchievementLevel 2 | 9,944.894*** | 8,359.763*** | 8,774.051*** | 7,415.855*** |
| | (2,089.561) | (2,071.360) | (2,000.449) | (1,982.213) |
| ESELA.AchievementLevel 3 | 9,520.084*** | 12,193.100*** | 10,118.000*** | 12,654.120*** |
| | (2,858.760) | (2,869.944) | (2,736.334) | (2,746.123) |
| ESELA.AchievementLevel 4 | 51,786.300*** | 53,365.400*** | 47,830.960*** | 49,295.680*** |
| | (3,189.963) | (3,184.993) | (3,054.573) | (3,048.721) |
| ESELA.AchievementLevel 5 | 94,951.130*** | 94,168.740*** | 95,684.630*** | 94,855.220*** |
| | (3,394.120) | (3,414.518) | (3,248.592) | (3,266.867) |
| HSMath.AchievementLevel 2 | -1,528.286 | 2,962.585 | -180.270 | 4,381.960** |
| | (2,273.255) | (2,217.083) | (2,176.081) | (2,121.593) |
| HSMath.AchievementLevel 3 | 68,116.050*** | 79,894.960*** | 56,399.740*** | 69,519.420*** |
| | (6,450.715) | (6,299.341) | (6,176.673) | (6,028.923) |

| | | | | |
|---|---|---|---|---|
| white_ratio | 95,574.810*** | 95,189.240*** | 94,914.660*** | 95,002.730*** |
| | (7,635.332) | (7,606.535) | (7,309.186) | (7,278.926) |
| M_Hshd_income | 1.040*** | 0.934*** | 1.029*** | 0.925*** |
| | (0.033) | (0.034) | (0.032) | (0.032) |
| crime_rate | -2.485*** | -2.305*** | -2.212*** | -2.014*** |
| | (0.273) | (0.279) | (0.261) | (0.267) |
| land_sf | 6.259*** | 6.174*** | 5.989*** | 5.896*** |
| | (0.258) | (0.256) | (0.247) | (0.245) |
| BUILDING_AREA | 105.081*** | 104.380*** | 112.397*** | 111.787*** |
| | (1.465) | (1.462) | (1.408) | (1.404) |
| NUM_STORY | 24,243.010*** | 22,656.040*** | 29,064.800*** | 27,451.250*** |
| | (1,806.838) | (1,801.554) | (1,731.047) | (1,725.277) |
| full_bath | 42,147.490*** | 42,177.400*** | 32,015.630*** | 32,005.220*** |
| | (1,286.670) | (1,282.409) | (1,242.780) | (1,238.196) |
| half_bath | 39,949.090*** | 40,204.170*** | 32,527.820*** | 32,776.710*** |
| | (1,719.091) | (1,713.711) | (1,649.176) | (1,643.378) |
| Elevation | -0.225 | -0.412 | -3.931 | -4.575 |
| | (4.753) | (4.854) | (4.549) | (4.644) |
| age | 461.526*** | 454.780*** | -104.186*** | -113.582*** |
| | (32.767) | (32.727) | (32.522) | (32.473) |
| Distance_frwy | 1.621*** | 0.916*** | 1.603*** | 0.984*** |
| | (0.323) | (0.275) | (0.309) | (0.263) |

| | | | | |
|---|---|---|---|---|
| Distance_arterial | 17.188*** | 16.121*** | 15.332*** | 14.338*** |
| | (1.556) | (1.578) | (1.490) | (1.510) |
| Distance_bus_stop | 12.163*** | 11.702*** | 9.912*** | 9.520*** |
| | (1.379) | (1.391) | (1.320) | (1.331) |
| Distance_open_space | -0.989 | 0.976 | 0.924 | 2.968** |
| | (1.287) | (1.291) | (1.232) | (1.236) |
| Distance_river | 2.075*** | 1.772*** | 2.120*** | 1.814*** |
| | (0.219) | (0.213) | (0.210) | (0.204) |
| Distance_school | -0.619 | -0.105 | 0.180 | 0.755 |
| | (1.117) | (1.119) | (1.069) | (1.070) |
| Distance_CBD | -4.890*** | -4.744*** | -5.681*** | -5.530*** |
| | (0.171) | (0.172) | (0.164) | (0.165) |
| mix_entropy | -87,208.980*** | -83,125.370*** | -82,295.360*** | -77,530.630*** |
| | (9,477.411) | (9,460.452) | (9,071.308) | (9,051.941) |
| Job_accessibility | -0.018 | 0.058** | 0.032 | 0.109*** |
| | (0.025) | (0.025) | (0.024) | (0.024) |
| ST_type11 | -55,710.470*** | 75,596.040*** | -54,113.870*** | 68,708.830*** |
| | (12,637.580) | (9,818.251) | (12,095.440) | (9,394.366) |
| ST_type12 | 395,499.600*** | 213,133.600*** | 384,092.200*** | 206,904.800*** |
| | (22,390.350) | (19,214.160) | (21,430.190) | (18,385.150) |
| ST_type13 | -57,682.900*** | 73,526.040*** | -62,882.430*** | 76,474.090*** |
| | (6,965.377) | (4,961.305) | (6,667.870) | (4,747.220) |

| | | | |
|---|---|---|---|
| Distance_LRT_stop | -14.638*** | -15.178*** | |
| | (0.718) | (0.688) | |
| year_sale2001 | | 13,624.520*** | 13,844.950*** |
| | | (3,731.975) | (3,717.681) |
| year_sale2002 | | 27,341.520*** | 27,401.030*** |
| | | (3,685.855) | (3,671.976) |
| year_sale2003 | | 37,731.160*** | 37,252.570*** |
| | | (3,640.757) | (3,626.911) |
| year_sale2004 | | 41,877.710*** | 41,702.950*** |
| | | (3,649.366) | (3,635.424) |
| year_sale2005 | | 48,073.030*** | 48,248.880*** |
| | | (3,563.530) | (3,549.685) |
| year_sale2006 | | 43,365.860*** | 43,648.930*** |
| | | (3,729.358) | (3,714.898) |
| year_sale2007 | | 47,936.490*** | 48,821.320*** |
| | | (3,862.744) | (3,848.355) |
| year_sale2008 | | 58,229.350*** | 58,695.970*** |
| | | (4,312.605) | (4,296.248) |
| year_sale2009 | | 85,292.510*** | 85,560.980*** |
| | | (4,270.911) | (4,254.630) |
| year_sale2010 | | 89,253.620*** | 88,902.820*** |
| | | (4,253.346) | (4,237.692) |

| | | | | |
|---|---|---|---|---|
| year_sale2011 | | | 87,508.740$^{***}$ | 87,674.140$^{***}$ |
| | | | (4,413.957) | (4,397.231) |
| year_sale2012 | | | 121,152.600$^{***}$ | 121,275.500$^{***}$ |
| | | | (4,180.208) | (4,163.808) |
| year_sale2013 | | | 124,174.500$^{***}$ | 124,413.900$^{***}$ |
| | | | (4,098.610) | (4,082.834) |
| year_sale2014 | | | 118,310.100$^{***}$ | 118,537.100$^{***}$ |
| | | | (4,064.234) | (4,048.559) |
| year_sale2015 | | | 129,392.300$^{***}$ | 130,123.300$^{***}$ |
| | | | (3,954.659) | (3,940.057) |
| year_sale2016 | | | 137,404.400$^{***}$ | 137,542.800$^{***}$ |
| | | | (4,049.228) | (4,034.089) |
| year_sale2017 | | | 138,325.400$^{***}$ | 138,298.600$^{***}$ |
| | | | (4,047.496) | (4,032.134) |
| year_sale2018 | | | 128,181.100$^{***}$ | 127,887.200$^{***}$ |
| | | | (4,124.698) | (4,109.346) |
| year_sale2019 | | | 94,542.630$^{***}$ | 92,080.140$^{***}$ |
| | | | (13,012.130) | (12,960.250) |
| ST_type1:Distance_LRT_stop | 13.240$^{***}$ | | 12.296$^{***}$ | |
| | (2.195) | | (2.101) | |
| ST_type2:Distance_LRT_stop | -54.350$^{***}$ | | -53.313$^{***}$ | |
| | (5.356) | | (5.126) | |

| | | | | |
|---|---|---|---|---|
| ST_type3:Distance_LRT_stop | 13.588$^{***}$ | | 14.464$^{***}$ | |
| | (0.818) | | (0.784) | |
| LRT_dummy1/4_mi | | 105,635.600$^{***}$ | | 110,381.000$^{***}$ |
| | | (14,659.370) | | (14,024.690) |
| LRT_dummy1/2_mi | | 103,630.000$^{***}$ | | 109,196.900$^{***}$ |
| | | (7,485.091) | | (7,161.706) |
| LRT_dummy3/4_mi | | 153,788.400$^{***}$ | | 158,146.100$^{***}$ |
| | | (6,809.556) | | (6,515.081) |
| LRT_dummy1_mi | | 93,118.850$^{***}$ | | 94,080.000$^{***}$ |
| | | (7,700.942) | | (7,368.872) |
| ST_type1:LRT_dummy1/4_mi | | -87,169.600$^{***}$ | | -85,571.900$^{***}$ |
| | | (27,293.960) | | (26,111.590) |
| ST_type2:LRT_dummy1/4_mi | | 31,450.970 | | 8,241.010 |
| | | (48,758.330) | | (46,647.650) |
| ST_type3:LRT_dummy1/4_mi | | -114,897.400$^{***}$ | | -120,869.900$^{***}$ |
| | | (15,110.330) | | (14,456.500) |
| ST_type1:LRT_dummy1/2_mi | | -124,425.200$^{***}$ | | -114,429.200$^{***}$ |
| | | (17,730.950) | | (16,963.590) |
| ST_type2:LRT_dummy1/2_mi | | 87,887.090$^{***}$ | | 84,906.030$^{***}$ |
| | | (26,371.670) | | (25,233.350) |

| | | | | |
|---|---|---|---|---|
| ST_type3:LRT_dummy1/2_mi | | -106,129.200$^{***}$ | | -113,004.000$^{***}$ |
| | | (7,820.348) | | (7,482.732) |
| ST_type1:LRT_dummy3/4_mi | | -155,558.900$^{***}$ | | -150,239.700$^{***}$ |
| | | (14,832.840) | | (14,190.350) |
| ST_type2:LRT_dummy3/4_mi | | -29,577.520 | | -32,864.580 |
| | | (22,313.430) | | (21,350.140) |
| ST_type3:LRT_dummy3/4_mi | | -152,445.600$^{***}$ | | -159,660.900$^{***}$ |
| | | (7,115.822) | | (6,809.082) |
| ST_type1:LRT_dummy1_mi | | -68,423.100$^{***}$ | | -61,462.280$^{***}$ |
| | | (16,635.210) | | (15,916.330) |
| ST_type2:LRT_dummy1_mi | | -163,708.000$^{***}$ | | -160,426.400$^{***}$ |
| | | (22,718.690) | | (21,736.670) |
| ST_type3:LRT_dummy1_mi | | -76,520.730$^{***}$ | | -78,893.280$^{***}$ |
| | | (7,982.341) | | (7,637.909) |
| Constant | 129,627.100$^{***}$ | -9,122.433 | 115,459.600$^{***}$ | -30,754.030$^{***}$ |
| | (13,352.620) | (11,291.480) | (13,017.440) | (11,096.920) |
| Observations | 48,237 | 48,237 | 48,237 | 48,237 |
| $R^2$ | 0.596 | 0.599 | 0.630 | 0.633 |
| Adjusted $R^2$ | 0.596 | 0.598 | 0.630 | 0.633 |

| | | | | |
|---|---|---|---|---|
| Residual Std. Error | 147,223.500 (df = 48204) | 146,693.900 (df = 48192) | 140,875.900 (df = 48185) | 140,315.000 (df = 48173) |
| F Statistic | 2,220.572*** (df = 32; 48204) | 1,634.839*** (df = 44; 48192) | 1,609.153*** (df = 51; 48185) | 1,319.402*** (df = 63; 48173) |
| Note: | *p<0.1; **p<0.05; ***p<0.01 | | | |

Compared with Model 1-4, most physical structure characteristic variables are still significant and have consistent coefficients in model 5. However, the willingness to pay for a full bath and half bath drop by about $8,000- $10,000 in Model 7 and Model 8 after controlling for time. The marginal effect of half bath and full bath are much closer after controlling for time. Additionally, age at the time of sale also has a negative effect on single-family house prices after controlling for time dummy variables. It makes sense, as newer houses might be easier to maintain. The elevation of homes has negative signs but is not significant across Model 5-8.

Among location characteristics, distance to open space is not significant in model 5-7, and has a varying sign in Model 5-8. Distance to school, which is significant in Model 1 -4 and has a negative sign, is positively related to housing prices across Model 7-8 with no significance. Other location variables like distance to bus stop, distance to freeway, distance to arterial, distance to open space, and distance to school all have the same signs and similar coefficients among Model 1-8.

All the neighborhood socio-economic variables are still significant and with the expected signs in Model 5-8. However, the coefficients for higher achievement levels in elementary school ELA achievement and high school math achievement are much smaller, compared to Model 1-4.

Land use mix entropy is significant and have same signs across model 1- 8. Job accessibility (Job_accessibility) is significant and positive in model 6 and 8, but not in model 5 and 7. The distance to light rail station has the expected negative signs and is significant in Model 5 and 7. The distance to light rail station has a higher marginal effect in Modeling 5 and 7, after controlling for interaction between light rail station type and distance to light rail station. As for the distance to light rail station dummy variable, after controlling for the interaction between distance dummy and station type, the most preferred distance band changes from 3/4 – 1mile to 1/2 – 3/4 mile.

After controlling for interaction terms between distance to light rail station (both the continuous and dummy variables) and station type, the station types in model 5 -8 are still significant at 0.01 level but some have changed signs. What is consistent across all models is that station type 2 (urban commercial), which is the most walkable, highly mixed use, and most dense station type, always has the highest marginal effect on housing price, all others being equal.

The interaction terms between distance to light rail station (both continuous and dummy variables) and station types are all significant. Model 5 and 7 tell us that people are most willing to pay for being closer to light rail stations when the station type is the type 2 (urban commercial). Model 6 and 8 have a more interesting result: for station type 1, all distance dummy variables have a negative coefficient compared to the reference level (which is more than 1 mile away from the light rail station), after controlling for the distance dummies. This suggests that people are most willing to pay to live further away from the light rail station, which makes sense as the main land use around the station area is industry. For station type 2, people are willing to pay higher prices for houses closer to the light rail station (within 0- 1/2 mile) compared to living outside 1 mile radius of the light rail station, after controlling for the distance dummies. It is very likely that people are willing to live close to the amenities around urban commercial stations. For station type 3, people are also willing to pay more for houses further away from the light rail stations, after controlling for the global distance dummies. Interaction terms for station type 1 and 3 have the lowest marginal effect at the 1/2- 3/4 mile distance band. The reason could be that such a distance might be too long to walk yet too short to drive.

All the time dummy variables in Model 7 and 8 are significant and have positive signs. The coefficients of these dummy variables show a sudden increase from 2008 - 2009 and 2011-2012, and a dramatic drop from 2018-2019. The dramatic rise and fall might be related to the population. From Figure 6, we can see that the population in Oregon has continuous growth since the mid-1980s and the growth rate keeps increasing from 2010 to 2018. Although the growth of Portland and Oregon is not likely to be proportional, I have reason to believe that as the biggest city in Oregon, Portland should follow a similar trend of population growth. As more people move in and the land supply is somewhat limited due to the urban growth boundary, the housing price is likely to go higher.

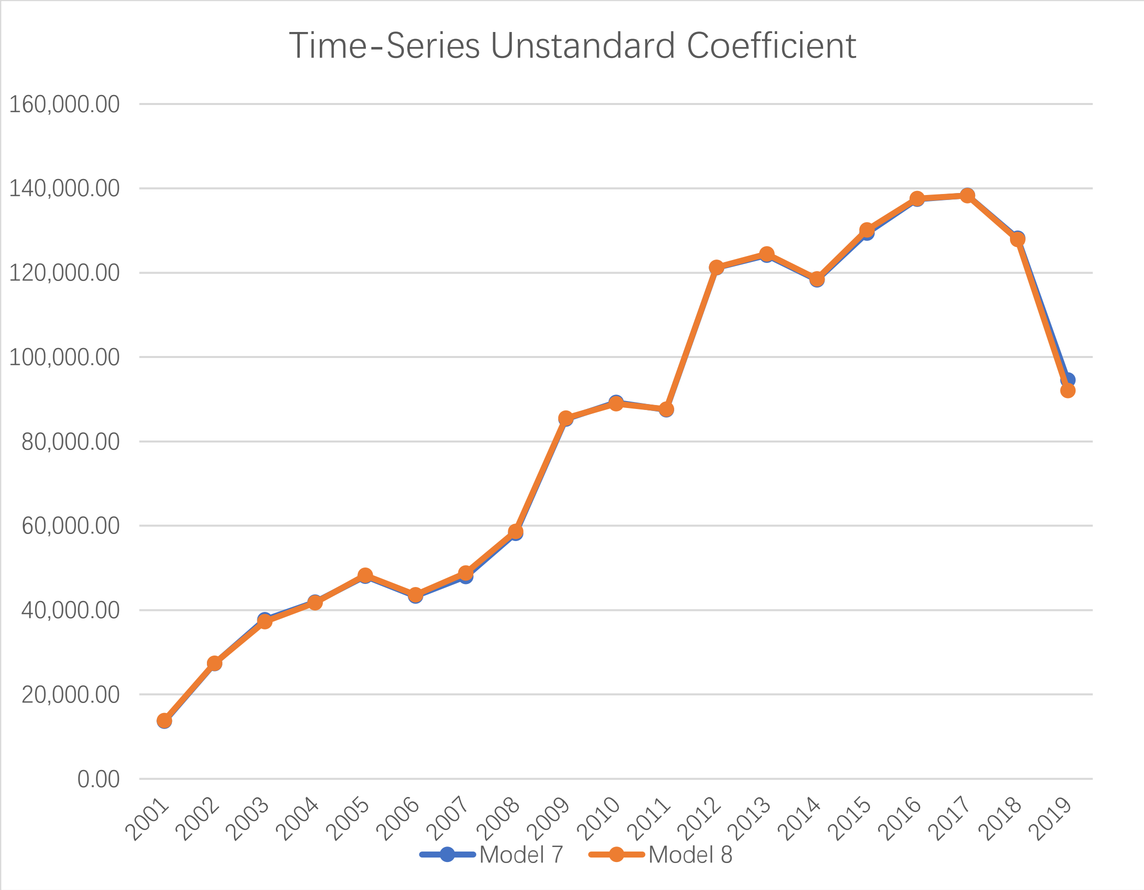



**Figure 5 Unstandardized coefficients of time-series dummy variables**

Another explanation could be the network effect brought by the larger transit network. As the transit network is getting more stations, people will have higher accessibility and probably value more of the light rail transit. The dramatic increase 2008- 2009, which is during economic recession, is coincidental with the end of construction of Green Line. I also suspect that after seeing the price increase of single family homes near light rail stations in the 2008-2009 Green Line completion, people have more land speculation around homes near light rail stations. However, it is also very likely that the increase from 2011- 2012 is related to economic recovery over the nation.

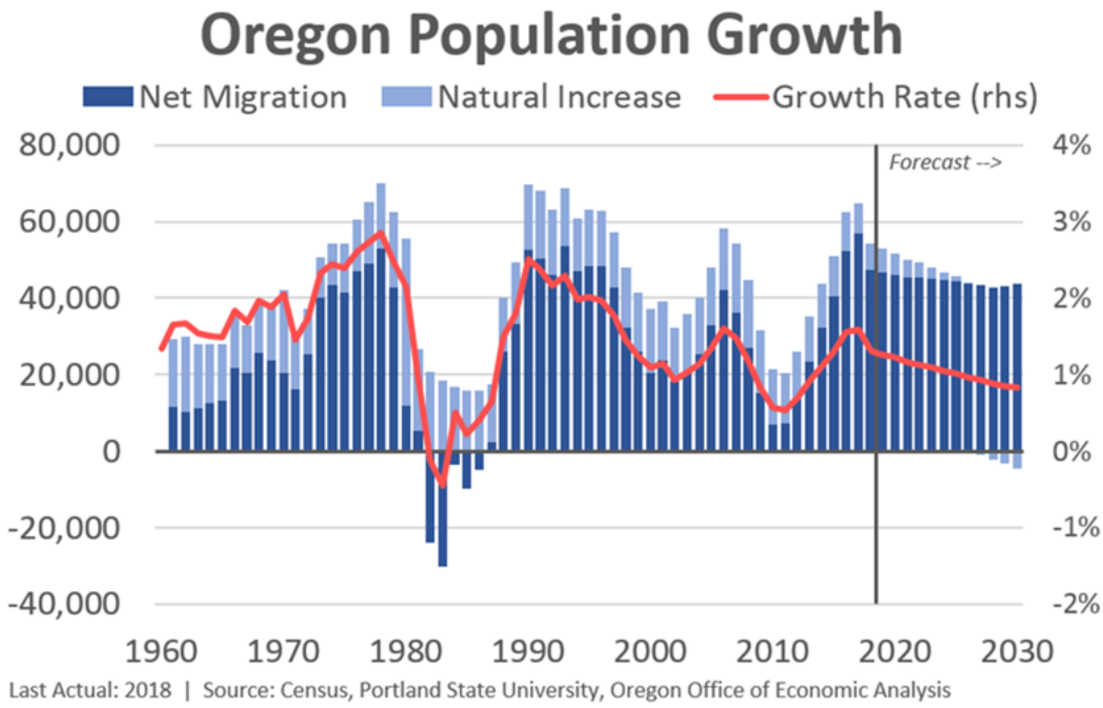



**Figure 6 Population Growth Rate of Oregon (Source: Oregon economic analysis)**

## Chapter 6 Discussion

The result confirms my hypotheses as they relate to land value uplift from TOD in the Portland city area. First, the cluster model results reveal that the characteristics of station areas vary within the study area. Furthermore, different TOD types are priced into the single-family homes at different rates. This particularly confirms the hypothesis that the land value uplift effect is heterogeneous across different TOD contexts. The results show access to TOD transit station is statistically significant for housing value by both continuous and dummy-distance variables. Model 1 suggests a higher valuation for a closer location to the light rail station. Model 2 and model 4 indicate that single-family houses located within a 1-mile radius of the station are valued higher than those outside a 1-mile radius. The 0.75-1 mile dummy has the highest positive impact on the home price, compared to the beyond 1 mile dummy. The 0.25-mile distance dummy is valued much lower compared to the 0.75-1 mile dummy, which might be related to a nuisance effect from the light rail station, such as crime and noise. However, after controlling for the interaction between station type and distance dummy, model 6 and 8 suggest the 0.5- 0.75 mile dummy has the highest positive impact on the home price instead.

Few studies estimated the effects of station typology on single family housing. This thesis fills the gap by assessing the land use mix entropy and the station typology. Results from the model 3-8 suggest that type 2 (urban commercial) has the highest value among the four station types, after controlling for job accessibility and distance to the CBD. After controlling for time, the station type 3 has a higher premium over station type 1.

The interaction terms between station type and distance to light rail station confirm my hypothesis that station types do impact the willingness to pay for proximity to TOD. For station type 1 and type 3, the interaction suggests that consumers may be willing to pay much less for the closer proximity to the light rail stations than station type 2 (urban commercial TOD) and station type 4 (open space TOD). The interaction term between station type and distance dummy suggests that people in station type 1 and 3 are willing to pay additional money to live further away from the light rail stations, after controlling for structural, socio-economic, locational, accessibility and the station type factors. On the other hand, the interaction term suggests that people are willing to pay more to be within the 0-0.5 mile radius of station type 2, in addition to the premium brought by station type and distance dummy. Station type 2 resembles more of the original TOD idea- mixed-use, longer sidewalk length, higher density, and accessibility, while station type 4 is more like the TAD (Transit Adjacent Development) by having lower density, low land use mix and low accessibility. What is surprising is that people near station type 4 value the proximity to light rail stations higher than people near station type 3. One explanation is that as the development near station type 4 is quite low, amenities might concentrate near the station area, and people may have a higher need for these location benefits. The development near station type 3 could have a more dispersed distribution of amenities, giving people more choices and leaving them less likely to value the areas near light rail stations.

The land use mix entropy has a negative sign in the regression model, which is opposite of my initial expectation. In the literature, land use mix has been positively related to walking trips (Ewing and Cervero, 2010), which translates to a more walkable environment. An explanation of the negative sign is that land use entropy as a standalone indicator is not sufficient as it cannot tell the different impacts of various land uses on housing prices. A study of mixed land use impact on housing value in Portland gained a similar negative result by using this entropy. Two different types of mixed use entropy were applied for comparison, one with all land uses (which is what I used in this study) and the other excluding single family land use. The authors conclude that single family home price increases if the nearby dominant land use is single family; and single family homes have a higher price premium if other land uses are evenly distributed (Song & Knaap, 2004). A previous BTOD study in Seattle by Shen et al. (2017) didn't find a significant relationship between land use entropy and single-family house price and suggested that land use entropy as an indicator might need some further examination.

## Chapter 7 Conclusion

With a trend to integrate transportation and land use planning for public transit, it seems reasonable to expect the promotion of TOD to result in significant price increases for locations surrounding light rail. While plenty of previous literature used hedonic price models to assess the impact of TOD proximity and plenty described typology of TOD, few had tried to bring typology into the hedonic model. This thesis tries to bridge the gap by using cluster analysis to create typology of TOD and using the typology in a hedonic model.

The model 8 confirms that TOD do have a significant positive impact on single family residential property located within 1 mile, compared to single family property outside of the 1-mile range. The peak premium is within 0.5- 0.75 mile, followed by 0.25 – 0.5 mile, after controlling for time, station type, interaction between distance dummy and station type and other variables.

The key finding reported in this thesis -that certain TOD station types create a substantial premium for single-family houses nearby has some policy implications. Model 7 and 8 suggest that highly walkable, mixed-use stations do have a higher premium over single family housing price. Future development of station areas should not just focus on the station, but also consider creating walkable, rich-amenity transit station areas. On the other hand, station types with major surrounding uses as single family housing and multifamily housing do not see a high premium for proximity to light rail stations. Future development near such stations may promote more density by infill development and enhance local services and amenities. As for station type 4, the future development might need to avoid falling into the spatial pattern of station type 3. More investment on sidewalks and thoughtful land use planning on greenfields is needed.

While this study is trying to fill the gap between built environment variables and TOD single-family house price, there are some limitations:

1. Because the multifamily data I collected tends to have a lot of recording errors and blank sales prices, I was not able to include multifamily sales or thus analyze how TOD affects multifamily prices. It is possible that multifamily residents have a stronger preference for TOD.
2. This study does not account for spatial autocorrelation, which is likely to bias the model results. Such a violation in the independence of residuals might lead to an overestimate of the impact of public transportation on land values by as high as 16% (Kim & Zhang, 2005). Future research

should use a spatial interaction model to give a more accurate estimate of TOD variables.

3. The data used for recent years might not state the true situation for earlier times. The socio-economic data are extracted from 2014-2018 ACS data, which could be quite different from decades ago. On the other hand, some of the physical structure data and price data extracted from the Portland open data website are not guaranteed to be accurate. This study removes some of the sales data with price '0' or unreasonable sale price, and those with missing key information. However, this removal of invalid data might not specify those with less obvious mistakes.
4. The distance measured in the model are Euclidean distance, which is much easier to calculate in a large dataset but might not represent the true distance for driving and walking.
5. The inclusion of the station type and some station variables in the model correlation might inflate the estimates. The categorization of station type is an unsupervised clustering based on accessibility, population and job density, land use mixture and sidewalk length. While in the final model, most of the variables with high VIF value are removed, the station type was not removed due to inconsistent VIF among different types. Generally, station types are mostly likely to be entangled with land use mixture entropy. Future research should try better means to capture land use impacts.
6. Typology as a variable also has its own limitation. While it is convenient to use typology to avoid some level of multicollinearity among highly correlated variables, it is difficult to tell which one of the correlated variables is the key to future planning.

While most of the study findings corroborate the mainstream study of TOD, some of the results are not intuitive, especially the negative association between mixed land use and single-family home prices. Future research might investigate how the land use mix could impact people's willingness to pay for a single-family house near light rail stations.

## References


Anselin, L. (2006). Spatial Econometrics. In T.C. Mills and K. Patterson (Eds.), Palgrave Handbook of Econometrics: Volume 1, Econometric Theory. Basingstoke, Palgrave Macmillan, pp. 901 - 969.

Atkinson-Palombo, C. (2010). Comparing the Capitalisation Benefits of Light-rail Transit and Overlay Zoning for Single-family Houses and Condos by Neighbourhood Type in Metropolitan Phoenix, Arizona. Urban Studies, 47(11), 2409–2426. https://doi.org/10.1177/0042098009357963

Atkinson-palombo, C., & Kuby, M. J. (2011). The geography of advance transit-oriented development in metropolitan. Journal of Transport Geography, 19(2), 189–199. https://doi.org/10.1016/j.jtrangeo.2010.03.014

Belzer, D., & Autler, G. (2002) Transit Oriented Development: Moving from Rhetoric to Reality (Washington, DC: Brookings Institution and The Great American Station Foundation).

Bernick, M, Cervero, R. (1996) High-speed Rail and Development of California's Central Valley: Comparative Lessons and Public Policy Considerations. University of Berkeley Publication, California, (1996)

Bartholomew, K., & Ewing, R. (2011). Hedonic Price Effects of Pedestrian- and Transit-Designed Development. Journal of Planning Literature, 1–36. https://doi.org/10.1177/0885412210386540

Boarnet, M., & Crane, R. (1997). L.A. Story: A reality check for Transit-Based housing. Journal of the American Planning Association, 63(2), 189–204. https://doi.org/10.1080/01944369708975914

Center for Transit-Oriented Development, & Nelson\Nygaard Consulting Associates. (2011). Transit-Oriented Development Strategic Plan/. Metro TOD Program. Retrieved from https://todresources.org/app/uploads/sites/2/2016/06/2011-portland-tod-final-web.pdf

Cervero, R., Ferrell, C., & Murphy, S. (2002). Research Results Digest. Minerva Medica, 63(14), 838–840. Retrieved from http://onlinepubs.trb.org/onlinepubs/tcrp/tcrp_rrd_52.pdf

Cervero, R. (n.d.). The 'Rail plus Property' model: Hong Kong's successful self-financing formula|McKinsey. Retrieved March 11, 2019, from https://www.mckinsey.com/industries/capital-projects-and-infrastructure/our-insights/the-rail-plus-property-model

De Haan, J., & Diewert, E. (2013). Hedonic Regression Methods. In Handbook on Residential Property Price Indices (pp. 49–64). https://doi.org/10.1787/9789264197183-7-en

Deng, L. (2011). The external neighborhood effects of low-income housing tax credit projects built by three sectors. Journal of Urban Affairs, 33(2), 143–166. https://doi.org/10.1111/j.1467-9906.2010.00536.x

Diao, M., Zhu, Y., & Zhu, J. (2017). Intra-city access to inter-city transport nodes: The implications of high-speed-rail station locations for the urban development of Chinese cities. Urban Studies, 54(10), 2249–2267. https://doi.org/10.1177/0042098016646686

Dill, J. (2008). Transit Use at Transit-Oriented Developments in Portland, Oregon, Area. Transportation Research Record: Journal of the Transportation Research Board, 2063(1), 159–167. https://doi.org/10.3141/2063-19

Dittmar, H. and Poticha, S. (2003) Defining transit-oriented development: the new regional building block, in: H. Dittmar and G. Ohland (Eds) The New Transit Town: Best Practices in Transit-oriented Development, pp. 19–40. Washington, DC: Island Press.

Duncan, M. (2008). Comparing rail transit capitalization benefits for single-family and condominium units in San Diego, California. Transportation Research Record, (2067), 120–130. https://doi.org/10.3141/2067-14

Duncan, M. (2011). The impact of transit-oriented development on housing prices in San diego, CA. Urban Studies, 48(1), 101–127. https://doi.org/10.1177/0042098009359958

Ewing, R. (2015). Research you can use. Planning, 81(9), 56–57.

Ewing, R., & Cervero, R. (2010). Travel and the built environment. Journal of the American Planning Association, 76(3), 265–294. https://doi.org/10.1080/01944361003766766

Gadziński, J., & Radzimski, A. (2016). The first rapid tram line in Poland: How has it affected travel behaviours, housing choices and satisfaction, and apartment prices? Journal of Transport Geography, 54, 451–463. https://doi.org/10.1016/j.jtrangeo.2015.11.001

Guan, C., & Peiser, R. (2018). Accessibility, urban form, and property value: A study of Pudong, Shanghai. Journal of Transport and Land Use, 11(1), 1057–1080. https://doi.org/10.5198/jtlu.2018.1318

Haining, R.P. in International Encyclopedia of the Social & Behavioral Sciences, 2001

Hale, C. (2014). TOD Versus TAD: The Great Debate Resolved…(?). Planning Practice and Research, 29(5), 492–507. https://doi.org/10.1080/02697459.2012.749056

Hess, D. B., & Almeida, T. M. (2007). Impact of proximity to light rail rapid transit on station-area property values in Buffalo, New York. Urban Studies, 44(5–6), 1041–1068. https://doi.org/10.1080/00420980701256005

Hibberd, R., Currans, K., & Nelson, A. C. (2019). Transportation Research Record FUNCTIONAL FORM IN HEDONIC REGRESSION : DETERMINING THE END OF SIGNIFICANCE OF TRANSIT PROXIMITY EFFECTS ON PROPERTY VALUE. In TRB 2019 Annual Meeting.

Higgins, C. D., & Kanaroglou, P. S. (2016a). A latent class method for classifying and evaluating the performance of station area transit-oriented development in the Toronto region. Journal of Transport Geography, 52, 61–72. https://doi.org/10.1016/j.jtrangeo.2016.02.012

Higgins, C. D., & Kanaroglou, P. S. (2016b). Forty years of modelling rapid transit's land value uplift in North America: moving beyond the tip of the iceberg. Transport Reviews, 36(5), 610–634. https://doi.org/10.1080/01441647.2016.1174748

Higgins, C., & Kanaroglou, P. (2018). Rapid transit, transit-oriented development, and the contextual sensitivity of land value uplift in Toronto. Urban Studies, 55(10), 2197–2225. https://doi.org/10.1177/0042098017712680

Kamruzzaman, M., Baker, D., Washington, S., & Turrell, G. (2014). Advance transit oriented development typology: Case study in brisbane, australia. Journal of Transport Geography, 34, 54–70. https://doi.org/10.1016/j.jtrangeo.2013.11.002

Kim, J. and M. Zhang. (2005). Determining transit's impact on Seoul commercial land values: An application of spatial econometrics. International Real Estate Review, 8(1) pp 1 - 26.

Laaly, S. (2014). New definition of Transit Oriented Development (TOD) based on context sensitive paradigm. ProQuest Dissertations and Theses, 171. Retrieved from http://search.proquest.com.ezaccess.library.uitm.edu.my/docview/1551496943?accountid=42518

LEHD Origin-Destination Employment Statistics (2002-2017) [computer file]. Washington, DC: U.S. Census Bureau, Longitudinal-Employer Household Dynamics Program [distributor], retrieved from https://onthemap.ces.census.gov.

Lovell, S. 2019. Transit Oriented Development. Guest Lecture Slide.

Mulley, C., Ma, L., Clifton, G., Yen, B., & Burke, M. (2016). Residential property value impacts of proximity to transport infrastructure: An investigation of bus rapid transit and heavy rail networks in Brisbane, Australia. Journal of Transport Geography, 54, 41–52. https://doi.org/10.1016/j.jtrangeo.2016.05.010

Nasri, A., & Zhang, L. (2014). The analysis of transit-oriented development (TOD) in Washington, D.C. and Baltimore metropolitan areas. Transport Policy, 32, 172–179. https://doi.org/10.1016/j.tranpol.2013.12.009

Nelson, A. C., Eskic, D., Hamidi, S., Petheram, S. J., Ewing, R., & Liu, J. H. (2015). Office rent premiums with respect to light rail transit stations: Case study of Dallas, Texas, with implications for planning of transit-oriented development. Transportation Research Record, 2500(2500), 110–115. https://doi.org/10.3141/2500-13

Nelson, A. C., & Hibberd, R. (2019). The Link Between Transit Stations and Commercial Rents,Jobs, People and Housing. National Institue for Transportation and Communities.

Polloni, S. (2019). Traffic calming and neighborhood livability: Evidence from housing prices in Portland. Regional Science and Urban Economics, 74, 18–37. https://doi.org/10.1016/j.regsciurbeco.2018.11.004

Pilgram, C. A., & West, S. E. (2018). Fading premiums: The effect of light rail on residential property values in Minneapolis, Minnesota. Regional Science and Urban Economics, 69(December 2017), 1–10. https://doi.org/10.1016/j.regsciurbeco.2017.12.008

Portland Streetcar. (2018). Portland Streetcar HB 2017 Advisory Committee Meeting Memorandum. Retrieved from https://trimet.org/meetings/hb2017/pdfs/2018-06-22/streetcar-presentation.pdf

Ratner, K., & Goetz, A. (2013). Transit Use at Transit-Oriented Developments in Portland, Oregon, Area. *Cities, 2063*(1), 159–167. https://doi.org/10.3141/2063-19

Rayle, L. (2015). Investigating the Connection Between Transit-Oriented Development and Displacement: Four Hypotheses. Housing Policy Debate, 25(3), 531–548. https://doi.org/10.1080/10511482.2014.951674

Renne, J. L. and Wells, J. S. (2005) Transit-oriented development: developing a strategy to measure success. National Cooperative Highway Research Project No. 2065(5), Transportation Research Board, Washington, DC.

Renne, J. L., Tolford, T., Hamidi, S., & Ewing, R. (2016). The Cost and Affordability Paradox of Transit-Oriented Development: A Comparison of Housing and Transportation Costs Across Transit-Oriented Development, Hybrid and Transit-Adjacent Development Station Typologies. Housing Policy Debate, 26(4–5), 819–834. https://doi.org/10.1080/10511482.2016.1193038

Rosen, S. (1974) Hedonic prices and implicit markets: product differentiation in pure competition, Journal of Political Economy, 82(1), pp. 34–55.

Salon, D., & Shewmake, S. (2012). Opportunities for Value Capture to Fund Public Transport: A Comprehensive Review of the Literature with a Focus on East Asia. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.1753302

Seo, K., Golub, A., & Kuby, M. (2014). Combined impacts of highways and light rail transit on residential property values: A spatial hedonic price model for Phoenix, Arizona. Journal of Transport Geography, 41, 53–62. https://doi.org/10.1016/j.jtrangeo.2014.08.003

Shen, Q. (1998). Location Characteristics of Inner-City Neighborhoods and Employment Accessibility of Low-Wage Workers. Environment and Planning B: Planning and Design, 25(3), 345-365.

Shen, Q., Xu, S., & Lin, J. (2018). Effects of bus transit-oriented development (BTOD) on single-family property value in Seattle metropolitan area. Urban Studies, 55(13), 2960–2979. https://doi.org/10.1177/0042098017729078

Song, Y., & Knaap, G. J. (2004). Measuring the effects of mixed land uses on housing values.

Regional Science and Urban Economics, 34(6), 663–680. https://doi.org/10.1016/j.regsciurbeco.2004.02.003

Suzuki, H., Cervero, R., & Iuchi, K. (2013). Transforming Cities with Transit. Вестник Казнму (Vol. №3).

Suzuki, H., Murakami, J., Hong, Y.-H., & Tamayose, B. (2015). Financing Transit-Oriented Development with Land Values: Adapting Land Value Capture in Developing Countries. https://doi.org/10.1596/978-1-4648-0149-5_ch7

Transportation Research Board. (2003). TCRP REPORT 102: Transit-Oriented Development in the United States: Experiences, Challenges, and Prospects. https://doi.org/10.17226/23360

Trimet. (2016a). Eastside MAX Blue Line Project Fact Sheet. Retrieved from https://trimet.org/publications/pdf/factsheets/max-blueline-eastside.pdf

Trimet. (2016b). Westside MAX Blue Line Project Fact Sheet. Retrieved from https://trimet.org/publications/pdf/factsheets/max-blueline-westside.pdf

Trimet. (2016c). Airport MAX Red Line Project Fact Sheet.Retrieved from https://trimet.org/publications/pdf/factsheets/max-redline.pdf

Trimet. (2016d). Interstate MAX Yellow Line Project Fact Sheet. Retrieved from https://trimet.org/publications/pdf/factsheets/max-yellowline.pdf

Trimet. (2016e). I-205/Portland Mall MAX Green Line Project Fact Sheet. Retrieved from https://trimet.org/publications/pdf/factsheets/max-greenline.pdf

Trimet. (2016f). Portland-Milwaukie Light Rail Project: A Vital Transportation Link. Retrieved from https://trimet.org/publications/pdf/factsheets/max-orangeline.pdf

Welch, T. F., Gehrke, S. R., & Farber, S. (2018). Rail station access and housing market resilience: Case studies of Atlanta, Baltimore and Portland. Urban Studies, 55(16), 3615–3630. https://doi.org/10.1177/0042098018760708

Wheeler, D., & Tiefelsdorf, M. (2005). Multicollinearity and correlation among local regression coefficients in geographically weighted regression. Journal of Geographical Systems, 7(2), 161–187.

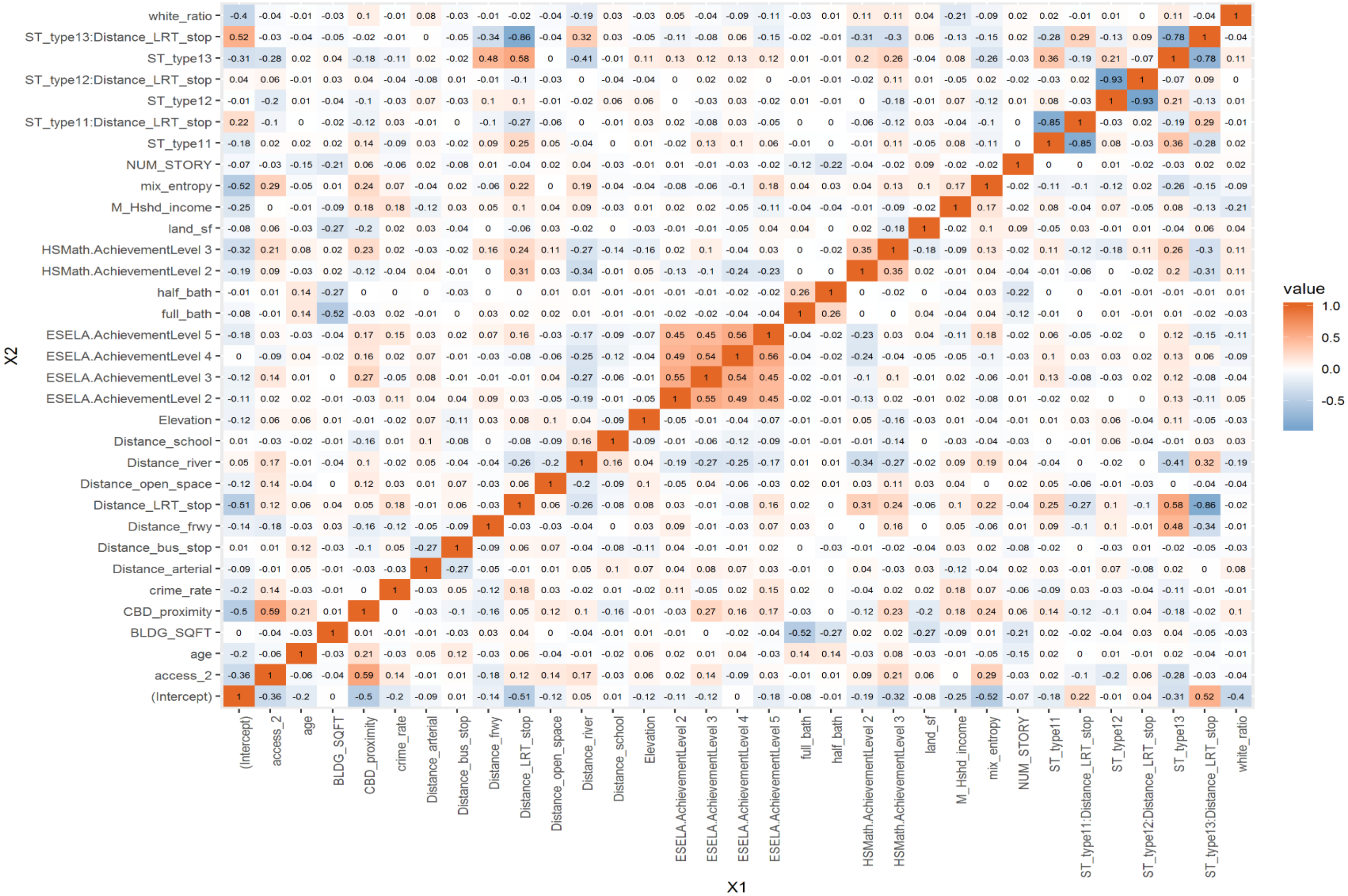
https://doi.org/10.1007/s10109-005-0155-6

White, S. M., & McDaniel, J. B. (1999). The Zoning and Real Estate Implications of Transit-Oriented Development. TCRP Legal Research Digest, 12(January), 1–51.

Yan, S., Delmelle, E., & Duncan, M. (2012). The impact of a new light rail system on single-family property values in Charlotte, North Carolina. Journal of Transport and Land Use, 5(2), 60–67. https://doi.org/10.5198/jtlu.v5i2.261

Yue, Y., Zhuang, Y., Yeh, A. G. O., Xie, J. Y., Ma, C. L., & Li, Q. Q. (2017). Measurements of POI-based mixed use and their relationships with neighborhood vibrancy. International Journal of Geographical Information Science, 31(4), 658–675. https://doi.org/10.1080/13658816.2016.1220561

Zemp, S., Stauffacher, M., Lang, D. J., & Scholz, R. W. (2011). Classifying railway stations for strategic transport and land use planning : Context matters ! Journal of Transport Geography, 19(4), 670–679. https://doi.org/10.1016/j.jtrangeo.2010.08.008




**Appendix A Partial Correlation Table**

## Appendix B The VIF Table of Model 1-8

| Variables | VIFModel 1 | VIFModel 2 | VIFModel 3 | VIFModel 4 | VIFModel 5 | VIFModel 6 | VIFModel 7 | VIFModel 8 |
|---|---|---|---|---|---|---|---|---|
| Job_accessibility | 2.36 | 2.38 | 3.24 | 3.30 | 3.29 | 3.36 | 3.29 | 3.37 |
| age | 1.51 | 1.51 | 1.51 | 1.51 | 1.51 | 1.52 | 1.63 | 1.64 |
| Building_area | 2.20 | 2.21 | 2.20 | 2.21 | 2.21 | 2.22 | 2.23 | 2.24 |
| Distance_CBD | 3.96 | 3.96 | 4.72 | 4.81 | 4.80 | 4.88 | 4.83 | 4.91 |
| crime_rate | 1.50 | 1.55 | 1.57 | 1.63 | 1.57 | 1.65 | 1.57 | 1.65 |
| Distance_arterial | 1.20 | 1.24 | 1.20 | 1.24 | 1.21 | 1.25 | 1.21 | 1.25 |
| Distance_bus_stop | 1.32 | 1.34 | 1.33 | 1.36 | 1.33 | 1.37 | 1.33 | 1.37 |
| Distance_frwy | 2.29 | 1.85 | 2.79 | 2.22 | 3.15 | 2.31 | 3.16 | 2.31 |
| Distance_LRT_stop | 2.53 | NA | 2.86 | NA | 11.35 | NA | 11.36 | NA |
| Distance_open_space | 1.19 | 1.20 | 1.20 | 1.21 | 1.21 | 1.23 | 1.21 | 1.23 |
| Distance_river | 1.64 | 1.65 | 1.81 | 1.83 | 2.04 | 1.95 | 2.05 | 1.95 |
| Distance_school | 1.20 | 1.19 | 1.20 | 1.20 | 1.20 | 1.22 | 1.21 | 1.22 |
| Elevation | 1.28 | 1.29 | 1.30 | 1.31 | 1.31 | 1.38 | 1.31 | 1.38 |
| ESELA.AchievementLevel 2 | 2.32 | 2.32 | 2.34 | 2.34 | 2.38 | 2.35 | 2.38 | 2.36 |

| Variables | VIFModel 1 | VIFModel 2 | VIFModel 3 | VIFModel 4 | VIFModel 5 | VIFModel 6 | VIFModel 7 | VIFModel 8 |
|---|---|---|---|---|---|---|---|---|
| ESELA.AchievementLevel 3 | 2.19 | 2.22 | 2.25 | 2.29 | 2.27 | 2.31 | 2.27 | 2.31 |
| ESELA.AchievementLevel 4 | 2.45 | 2.46 | 2.68 | 2.69 | 2.69 | 2.70 | 2.69 | 2.70 |
| ESELA.AchievementLevel 5 | 2.50 | 2.52 | 2.53 | 2.55 | 2.58 | 2.63 | 2.58 | 2.63 |
| full_bath | 1.75 | 1.75 | 1.75 | 1.75 | 1.75 | 1.76 | 1.79 | 1.79 |
| half_bath | 1.26 | 1.26 | 1.26 | 1.26 | 1.26 | 1.26 | 1.27 | 1.27 |
| HSMath.AchievementLevel 2 | 1.56 | 1.58 | 1.60 | 1.62 | 1.78 | 1.70 | 1.78 | 1.70 |
| HSMath.AchievementLevel 3 | 1.79 | 1.80 | 1.99 | 2.00 | 2.24 | 2.15 | 2.24 | 2.16 |
| land_sf | 1.53 | 1.53 | 1.53 | 1.53 | 1.54 | 1.53 | 1.54 | 1.54 |
| M_Hshd_income | 2.00 | 2.02 | 2.05 | 2.07 | 2.08 | 2.19 | 2.08 | 2.19 |
| mix_entropy | 1.59 | 1.59 | 2.55 | 2.57 | 2.62 | 2.63 | 2.62 | 2.63 |
| NUM_STORY | 1.43 | 1.43 | 1.43 | 1.43 | 1.43 | 1.43 | 1.44 | 1.44 |
| white_ratio | 1.69 | 1.69 | 1.73 | 1.72 | 1.73 | 1.73 | 1.73 | 1.73 |
| LRT_dummy1/2_mi | NA | 1.92 | NA | 2.13 | NA | 17.40 | NA | 17.41 |
| LRT_dummy1/4_mi | NA | 1.43 | NA | 1.51 | NA | 23.32 | NA | 23.33 |
| LRT_dummy1_mi | NA | 1.40 | NA | 1.44 | NA | 17.47 | NA | 17.48 |

| Variables | VIFModel 1 | VIFModel 2 | VIFModel 3 | VIFModel 4 | VIFModel 5 | VIFModel 6 | VIFModel 7 | VIFModel 8 |
|---|---|---|---|---|---|---|---|---|
| LRT_dummy3/4_mi | NA | 1.62 | NA | 1.74 | NA | 14.43 | NA | 14.44 |
| ST_type11 | NA | NA | 1.67 | 1.68 | 5.97 | 3.63 | 5.97 | 3.63 |
| ST_type12 | NA | NA | 1.61 | 1.61 | 11.99 | 8.89 | 11.99 | 8.90 |
| ST_type13 | NA | NA | 4.47 | 4.51 | 11.31 | 5.78 | 11.32 | 5.78 |
| ST_type11:Distance_LRT_stop | NA | NA | NA | NA | 5.30 | NA | 5.30 | NA |
| ST_type12:Distance_LRT_stop | NA | NA | NA | NA | 10.90 | NA | 10.90 | NA |
| ST_type13:Distance_LRT_stop | NA | NA | NA | NA | 16.92 | NA | 16.94 | NA |
| year_sale2001 | NA | NA | NA | NA | NA | NA | 2.01 | 2.01 |
| year_sale2002 | NA | NA | NA | NA | NA | NA | 2.06 | 2.06 |
| year_sale2003 | NA | NA | NA | NA | NA | NA | 2.12 | 2.12 |
| year_sale2004 | NA | NA | NA | NA | NA | NA | 2.11 | 2.11 |
| year_sale2005 | NA | NA | NA | NA | NA | NA | 2.23 | 2.23 |
| year_sale2006 | NA | NA | NA | NA | NA | NA | 2.02 | 2.02 |
| year_sale2007 | NA | NA | NA | NA | NA | NA | 1.89 | 1.90 |
| year_sale2008 | NA | NA | NA | NA | NA | NA | 1.61 | 1.61 |

| Variables | VIFModel 1 | VIFModel 2 | VIFModel 3 | VIFModel 4 | VIFModel 5 | VIFModel 6 | VIFModel 7 | VIFModel 8 |
|---|---|---|---|---|---|---|---|---|
| year_sale2009 | NA | NA | NA | NA | NA | NA | 1.64 | 1.64 |
| year_sale2010 | NA | NA | NA | NA | NA | NA | 1.64 | 1.65 |
| year_sale2011 | NA | NA | NA | NA | NA | NA | 1.58 | 1.58 |
| year_sale2012 | NA | NA | NA | NA | NA | NA | 1.70 | 1.70 |
| year_sale2013 | NA | NA | NA | NA | NA | NA | 1.75 | 1.75 |
| year_sale2014 | NA | NA | NA | NA | NA | NA | 1.78 | 1.78 |
| year_sale2015 | NA | NA | NA | NA | NA | NA | 1.88 | 1.88 |
| year_sale2016 | NA | NA | NA | NA | NA | NA | 1.80 | 1.80 |
| year_sale2017 | NA | NA | NA | NA | NA | NA | 1.81 | 1.81 |
| year_sale2018 | NA | NA | NA | NA | NA | NA | 1.75 | 1.75 |
| year_sale2019 | NA | NA | NA | NA | NA | NA | 1.05 | 1.05 |
| ST_type11:LRT_dummy1/2_mi | NA | NA | NA | NA | NA | 1.66 | NA | 1.66 |
| ST_type11:LRT_dummy1/4_mi | NA | NA | NA | NA | NA | 1.63 | NA | 1.63 |
| ST_type11:LRT_dummy1_mi | NA | NA | NA | NA | NA | 1.90 | NA | 1.90 |
| ST_type11:LRT_dummy3/4_mi | NA | NA | NA | NA | NA | 2.13 | NA | 2.13 |

| Variables | VIFModel 1 | VIFModel 2 | VIFModel 3 | VIFModel 4 | VIFModel 5 | VIFModel 6 | VIFModel 7 | VIFModel 8 |
|---|---|---|---|---|---|---|---|---|
| ST_type12:LRT_dummy1/2_mi | NA | NA | NA | NA | NA | 2.45 | NA | 2.45 |
| ST_type12:LRT_dummy1/4_mi | NA | NA | NA | NA | NA | 1.33 | NA | 1.33 |
| ST_type12:LRT_dummy1_mi | NA | NA | NA | NA | NA | 4.30 | NA | 4.30 |
| ST_type12:LRT_dummy3/4_mi | NA | NA | NA | NA | NA | 4.42 | NA | 4.43 |
| ST_type13:LRT_dummy1/2_mi | NA | NA | NA | NA | NA | 17.50 | NA | 17.52 |
| ST_type13:LRT_dummy1/4_mi | NA | NA | NA | NA | NA | 23.13 | NA | 23.14 |
| ST_type13:LRT_dummy1_mi | NA | NA | NA | NA | NA | 17.10 | NA | 17.12 |
| ST_type13:LRT_dummy3/4_mi | NA | NA | NA | NA | NA | 13.93 | NA | 13.94 |