The College Of WILLIAM & MARY

Estimating the Effect of Transit on Residential Property Values: The Case of the Portland MAX System

Keith Klovers
The College of William and Mary

Alfredo Marvão Pereira
The College of William and Mary

College of William and Mary
Department of Economics
Working Paper Number 166

August 2015

COLLEGE OF WILLIAM AND MARY
DEPARTMENT OF ECONOMICS
WORKING PAPER # 166
August 2015

Estimating the Effect of Transit on Residential Property Values: The Case of the Portland MAX System

Abstract

Using a hedonic price model with demographic, spatial, and house characteristic data, this study finds significant property value premiums associated with access to rail mass transit in the Portland, Oregon. Incremental changes to a "benchmark" model—similar to many used in the literature—lead to the adoption of a more appropriate "target" model that incorporates several innovations. This paper finds that assessment data is more appropriate and consistent than sale price data; that the use of a continuous measurement of the distance to transit is more powerful than a binary measure; and that the consideration of "community amenities"—such as schools or parks—in the model specification improves the resulting estimates. The target model finds significantly higher transit access premiums than prior models, suggesting that municipalities may see greater property tax benefits from transit construction than previously estimated.

Keywords: Residential Property Values, Rail Mass Transit, Hedonic Price Model, Portland
JEL Classification: H71, L92, R12, R32, R51

Keith Klovers
Department of Economics,
The College of William and Mary,
Williamsburg, VA 23187

Alfredo Marvão Pereira
Department of Economics,
The College of William and Mary, Williamsburg, USA
PO Box 8795, Williamsburg, VA 23187
ampere@wm.edu

Estimating the Effect of Transit on Residential Property Values: The Case of the Portland MAX System

1. Introduction

The issue of how public transportation affects property values continues to garner a lot of attention in both academic and policy circles (Cervero et al., 2002; Wardrip, 2011; Becker, 2013). Classic bid-rent theory holds that residential property rents and values rise in relation to the proximity to employment centers. The central aspect of bid-rent theory is that individuals, particularly those commuting to employment centers, have a positive value of time. If a transportation improvement reduces their travel time to work, this "additional" time can be freely spent by the individual on greater work, which carries wage benefits, or leisure, which carries less explicit personal utility benefits, or some optimal combination of the two. Since transportation amenities such as highways and rail transit systems can be seen as utilities that reduce the time-distance to employment, it holds that properties near these amenities should see high demand; as long as supply is sufficiently scarce, this demand will bid up prices and generate rent and value premiums (Alonso, 1974). The transportation amenity has also been described as an extension of traditional "capitalization" theory, whereby any relevant amenities are incorporated into underlying property values (Landis et al, 1994). This expansion of bid-rent theory allows for all manner of other amenities to be priced into property values. In this manner other factors—such as neighborhood characteristics, access to community institutions, such as parks, libraries, and the quality of a given structure—can have an impact on residential property values.

A wide array of studies has been conducted which tends to support the idea that there is a significant effect of mass transit access on surrounding property values. Landis et al (1994) found significant premiums for residential properties associated with Californian transit systems such as San Francisco's CalTrain commuter rail and BART heavy rail systems, as did Voith (1993) in his study of the Lindenwold section of Philadelphia's SEPTA commuter line. Similarly significant effects have been estimated for the Atlanta, Boston, Chicago, Portland, and Washington, D.C. areas (Baum-Snow and Kahn, 2000). The effect of dis-amenities—such as noise and crime—that often accompany transit stops have also been examined, primarily for Californian transit systems (Landis et al, 1994) and Atlanta's MARTA (Bowes and Ihlanfeldt, 2001). Perhaps reflecting a mix of amenity and dis-amenity effects, one influential study found that transit development exerted a minimal impact, if any, on property values in the Miami area (Gatzlaff and Smith, 1993).

The predominant sampling method uses the market sales price of a given property as a proxy for that property's underlying value (e.g. Voith, 1993; Cervero, 2004; and many others); often, the most recent available year is selected. Although the use of such data is generated in the open market—as opposed to assessment data—and thus could more accurately reflect a property's value, such samples are also more susceptible to bias. Cervero (2004), among others, readily noted that sales data includes non-"arms-length" transactions, such as below-market-price sales between relatives; as a result researchers often filter the data to exclude records for which the sale price is significantly different from the assessed value.

Several approaches have been taken to measure a property's access to mass transit. Early analysis, such as Voith (1993), considered all properties in a U.S. Census Tract that included a station as accessible; if the tracts were sufficiently small, sometimes adjacent tracts were also considered transit-accessible. Weinstein and Clower (2002) and Fejarang (1994) used a system of matched pairs, so that a group of properties near transit were matched to a similar non-transit-accessible group. Among the most comprehensive of these rudimentary approaches, Dewees (1976) considers the walking time to transit and time on transit, early forms of "cost" distance, for his study of a subway improvement in Toronto.

More recently, there has also been concern that the endogenous variable, underlying property value, and the variable of interest, access to mass transit, may not be properly measured by the traditional proxy variables. In response, some authors have used county assessment value instead of recorded sale price (e.g. Weinstein and Clower, 2002), while other studies have used continuous measures of transit access, such as straight-line distance, in place of more traditional binary variables (e.g. Landis et al, 1994; Bowes and Ihlanfeldt, 2001).

The existence of a positive externality of mass transit on property values and, even more so, the absolute size of such externality is of great interest to the policy community. If the size of the transit access premium is large, then municipalities can expect the rise in property values to translate into significantly higher property tax revenue, which can be used to build and maintain the system. Such estimates also enter into cost-benefits analyses, so that projects with large estimated property value and, thereby, large tax benefits are more likely to be undertaken.

Unfortunately for policymakers, it is difficult to compare previous studies and come to general conclusions about the presence and size of the transit access premium. Indeed, although each study is loosely based on the hedonic price model, models vary substantially in terms of specification, location, and time frame. Given the many levels of variation between studies, it is virtually impossible to disentangle the relative effect that differences in sample (e.g. location, time-frame), measurement (e.g. real sale data or county assessment data, straight-line distance or distance dummies) or specification have on the resulting estimates. Thus, for both academic and policy reasons, it would be important to examine how the effect of each one of the above changes, in isolation, affects the estimated effect of mass transit on property values.

This paper is an attempt to bridge previous studies and promote direct comparison within the literature. Using data from 2001 to 2006 for single-family homes in the Portland, Oregon area, this study begins with a "benchmark" model derived from the work of Bowes and Ihlanfeldt (2001), Cervero (2004), and Voith (1993) to identify the effects of mass transit on residential property values. From this model, incremental changes are made until the most adequate mix of sample, variable measurement, and specification is determined. This approach ensures that sample problems related to location or time heterogeneity are controlled—that is, it allows the aforementioned variations to be examined holding location and time-frame constant. Therefore, in the process of identifying the effects of mass transit on residential property values in the Portland area, this paper addresses three important methodological questions: (1) Is the traditional measure of underlying property value, real estate sales data, more appropriate than government assessment data or a hybrid average of several years' sales data?; (2) Is the continuous distance measurement superior to the constant-distance (also known as the "distance ring") binary variable approach?; and (3), Does the existing literature omit important explanatory variables, particularly "community amenities"?

Portland represents an excellent setting for testing these hypotheses. The city pioneered light rail mass transit construction and anti-sprawl legislation in the 1970s and 1980s, and is often cited as a model of regional transportation planning. The first rail mass transit line, the Eastside Blue Line, began operation September 5, 1986; the most recent addition, the Interstate Yellow Line, has operated since May 1, 2004 (Tri-Met, 2006). This long history of rail mass transit operation, along with the absence of service changes over the past few years, suggests that land values near transit stops fully reflect the amenity effects that such stops are reputed to provide. Portland is also ideal from a policymaker's perspective, as it mirrors the many mid-sized American cities which are most interested in expanding their rail mass transit systems; for example, in the last decade systems have been initiated or expanded in Dallas, Denver, Houston, Minneapolis, and St. Louis, among others.

The development of powerful Geographical Information Systems ("GIS") software has allowed a refinement of these spatial methods. GIS-based methods have quickly become the standard in transit studies, and can be categorized as one of two basic approaches. One method, similar to past categorization methods, uses distance circles centered on transit stops to separate properties within a certain distance to transit from those farther away; such methods express the transit access premium as the coefficient of the relevant dummy variable (e.g. Cervero, 2004). The other method involves the calculation of exact distances from properties to the nearest transit stop; although this distance is usually a straight-line distance (e.g. Koutsopoulis, 1977; Landis et al, 1994) it has also been expressed as a "cost distance" that takes street layout and obstructions—such as rivers, highways, and so forth—into consideration (Bowes and Ihlanfeldt, 2001). This cost distance has more broadly been measured as the commute time for a given property (Dewees, 1976) or a property's census tract (Cervero, 2004).

In addition to quantifying transit access, GIS can and has been used to measure spatial accessibility to other commonly recognized amenities, particularly highways and community parks. Cervero (2004) and Landis et al (1994) incorporate highway accessibility effects into their models, while the effect of parks on property values has been explored by environmental economists (e.g. Weigher and Zerbst, 1973; Wu et al, 2004). Considering that there are other community amenities—such as schools, libraries, and hospitals—that individuals value having access to, these factors have been included as well in the model.

This paper is divided into four sections. The first section examines the study's theoretical and empirical foundations and introduces the empirical models. The second discusses the study area, data sources, and derived variables. The third section provides summary statistics and multiple regression results. Finally, the fourth section provides a summary, concluding remarks as well as suggestions for suggest future research.

2. The Empirical Models

Rosen (1974), among others, discusses how a product's price is in fact the sum of the vectors of the product's attributes; as such, any product's price can be decomposed into implicit, or "hedonic", prices attributed to each attribute vector. This hedonic approach is widely used in transportation economics, with parameter estimates most often found using Ordinary Least Squares multiple regressions. This approach is used in many of the more recent transit studies, including Landis et al (1994), Cervero (2004), Bowes and Ihlanfeldt (2001), Baum-Snow and Kahn (2000), and Voith (1993), to name a few. Building upon this lineage, the models in this paper are also expressed as hedonic price models.

This paper's research is based on several models, starting with an approximation of the "typical" model found in current literature; from this base the model is gradually modified until the "target" model is defined. This section presents the starting model and our target model; the other models used in this paper are variants of these approaches. Note that, because the majority of these models are log-log, the parameter estimates must be transformed from elasticities or semi-elasticities into meaningful nominal numbers; these transformations are noted when necessary.

The target model has grown out of a number of incremental changes from the models seen in the current transit scholarship. To capture how a typical model would estimate the transit access premium with our dataset, a "benchmark" model was created that integrates several of the most common features in previous research. This benchmark model can be expressed as:

(1) $$P_i = f(C, N, T)$$

where P_i is the sale price of property *i* in a particular year;

C is a vector of property characteristics;

N is a vector of neighborhood characteristics;

and T is a vector of transportation access

P_i—the sale price of the property in a given year—ensures that each model is only for properties that sell in that particular year. Note that this applies Cervero's filter criterion, so only those properties for which the sale price is within 10% of the assessed value are considered. C—the vector of house structure characteristics—includes structure size, structure age, and the property's lot size. N—the vector of neighborhood characteristics—is expressed for the respective Census Tract that the property is part of, and includes the median age, median household income, percent of the population that is not white, and the percentage of households that are single-family detached houses, that are vacant, that are renter-occupied, and that have an income exceeding \$100,000. T—the vector of transportation access—includes the straight-line distance to the nearest highway, the average commute time of the Census Tract that the property lies in, and whether or not the property is within one-half mile of a transit station.

Although these methods can be found throughout the literature, these particular variables can collectively be found in Voith (1993), Cervero (2001), and Bowes and Ihlanfeldt (2001).

All nominal variables have been scaled by taking the natural log, so that the model is estimated in log-log form. As a result, these relationships can be interpreted as elasticities, so that a one percent change in the exogenous variable x_1 induces a B_{x1} percent change in the endogenous variable y. As such, a change in relative distance from 1 mile to 2 miles is analogous to a change from 5 miles to 10 miles (a 100% increase), so that the marginal effect of a fixed-unit change—such as a 10 foot increase in the distance to the nearest transit station—decreases as that variable increases in absolute size. In contrast, a linear relationship would treat these changes equally, as the marginal effect is constant regardless of the initial distance.

The target model incorporates these basic features while introducing non-linearities, expanding the sample, enhancing factor measurement, and adding an additional exogenous factor, the straight-line distance to the nearest "community amenities"—schools, libraries, hospitals, and parks. Thus, our model can be expressed in the form:

(2) $$V_i = f(C, N, A, T)$$

where V_i is the estimated total value of property i for 2006

C is a vector of property characteristics;

N is a vector of neighborhood characteristics;

A is a vector of the distance to certain community amenities

and T is a vector of transportation access;

Although (2) is similar in specification to (1), there are two significant changes. First, equation (2) is estimated as a log-log model in which value, spatial distances, demographic data (e.g. median income), commute time, property age, and property square footage have been scaled. Secondly, this model includes the community amenities vector, reflecting an attempt to meld environmental economic research into transit analysis.

Assessment data, rather than sales data, were chosen due to concerns that sales data may suffer from particularly severe selection bias. Sales conditions may vary widely across properties, as some homeowners may seek a quick sale while others hold out for higher prices, or as skilled realtors drive up prices on their properties relative to less-skilled competitors; such unobserved heterogeneity may introduce bias. Additionally, demographic or locational factors may affect whether particular types of properties are sold in a given year, complicating the selection of a representative sample. Furthermore, the data suggest that those properties without sale records tend to be significantly older than those properties sold recently; considering that areas tend to be built around the same time, there may be geographic biases introduced by using sales data. Finally, as discussed in the Data section, sales data include many non-“arms-length” transactions (e.g. transactions at non-market prices, such as when a house is “sold” from one family member to another for $1); the filtering techniques used to clean this data are often unduly stringent and themselves produce distortions in the data.

Assessment values, although not perfect, appear to be a better measure of underlying property value. Such values are assessed universally and thus available for all properties. In addition, assessments come from central county assessors’ offices, so that valuations are consistent across like properties; this process eliminates potential biases arising out of seller, buyer, or realtor behavior heterogeneity. Weinstein and Clower (2002), noting similar concerns with sales data, use assessment data in their study of the Dallas transit system, DART.

3. Data

3.1 Data Sources

Data comes from two sources, the Regional Land Information System (RLIS) “Lite” dataset and the U.S. Census Bureau’s 2000 Decennial Census. The RLIS Lite dataset provides basic property characteristics as well as substantial GIS spatial data. The dataset includes over 560,000 individual residential, commercial, industrial, and special-use property records for the three-county area governed by Metro, the three-count regional government, and details assessed land value, assessed total value, most recent sale price and sale date, structure square footage, and structure age for each property. The GIS data include spatial information and attributes for the three county area’s properties, census tracts, environment, public buildings, and transportation infrastructure, as well as the square footage of each property’s lot. The data are updated regularly by Metro, and all data are described as “current”, "annually updated", or "updated as needed" in the associated Metadata; as such, it appears that all data layers are current as of purchase (November 2006). The US Census Bureau's Census 2000 Summary Files 1 and 3 contain demographic information for the Portland area, including median age, median income, median household size, owner occupancy rates, housing vacancy rates, the percentage of housing units that are single-family detached units, the percentage of households with an income exceeding $100,000 per year, and race structure. Finally, the assessed value data was provided to Metro by the three constituent counties, Clackamas, Multnomah, and Washington. All assessments are updated yearly, based on the last year's "market conditions"—primarily sales prices. Although such a method is "adjusted" (upward) so that assessed values are based on predictions for the future year, it is unclear whether such a method implies that assessed values lag or lead market sales prices.

3.2 Generating Spatial Data

Spatial data was generated using ArcGIS. Distances are expressed as straight-line distances rather than cost distances. Choosing straight-line distances should yield superior measurements of distance for commuters who walk to transit and inferior measurements for commuters who drive to transit, as cost distances calculate a faster movement on major streets than on side-streets (by supposing a higher speed limit); since walking commuter travel times are not affected by road speed limits, a straight-line distance should provide a better estimate of the actual distance facing these commuters. Because Portland is a leader in so-called Transit-Oriented Development (TOD), which emphasizes dense residential development within walking distance of transit, and—according to an October 2006 Tri-Met (Portland transportation authority) publication (Tri-Met, 2006a)—has only 8,112 parking spots at MAX Park-and-Ride facilities versus an weekday average of 82,500 MAX riders (Tri-Met, 2006b), it is expected that more commuters walk to transit than drive to transit. Thus, straight-line distances should provide a better over-all estimate of commuters' time-distance to transit than a cost distance approach would. Future studies, however, might wish to examine how the results of a cost distance model differ from those of a straight-line distance model, particularly for more auto-oriented transit systems such as commuter rail. GIS was also useful for assigning Census data to each property, as the Census tract footprints could be overlaid onto the property map with ArcGIS and then merged to Census Bureau tables using STATA.

Since a portion of the three-county area surrounding Portland is rural, all properties located outside the Urban Growth Boundary ("UGB") were excluded. The Boundary is the statutory line between urban and rural land in the Portland area, and in effect marks the limit of urban expansion. Although this is particularly the case in

Oregon given the legal environment, urban and rural property (and employment) markets are in general sufficiently different to warrant such a distinction between them. A property was determined to be inside or outside the UGB by digitally overlaying the UGB layer onto the property parcels layer, and performing a spatial join similar to that used to add Census Tract identifiers.

A typical ArcGIS layer overlay looks similar to the map presented in Figure 1, which shows the UGB Boundary (solid red line) and various distance rings (in this case demarcating one-half-, one-, three-, and five-mile distances from the nearest MAX transit stop, in Beige, Yellow, Blue, and Green, respectively) overlaid onto a property map of Portland's three-county region.

3.3 Deriving the Full Data Sample

Although the target model will utilize assessment data, for comparative purposes some models will be estimated using sales price data. As past studies have selected all properties sold in a given year, sub-samples were created for those properties sold (at a non-zero price) in the years 2001, 2002, 2003, 2004, and 2005. In addition, sub-samples were created which included all properties last sold in the intervals 1981-2005, 1991-2005, 1996-2005, and 2001-2005. Because RLIS Lite records the most recent sale price for a property, which ranges from 1942 to 2006 in our dataset (and is absent altogether for properties who are owned by the original owner), the sales price mean is not consistent, and is omitted from the summary statistics for the full sample.

The presence of some obvious assessment errors (such as a house that recently sold for $300,000 registering an assessed value of $3,000) required the imposition of a "floor" assessment value. This lower bound was set at $40,000 after consulting a density plot of the total assessed value of properties, which is reproduced in Figure 2. It is worth noting that the density plot has an upper bound of $100,000 so that the lower-assessed values are easily distinguished; the sample itself includes many properties above the $100,000 level.

Because some authors have noted with concern that such samples may include non-arms-length transactions, the common practice of filtering these yearly samples was also followed. One approach was to use the filtering criteria set out in Cervero (2004), which keeps all properties with a sale price that is within 10% of that property's assessed total value. Since the sale price data is collected yearly and trends upward over time, while the assessments are only available for 2006 (based on 2005 sales data), the assessment values were deflated for each year using an index based on the average property appreciation rate from 2001 to 2005, which appears to be about 37%. Table 1 confirms that the index is appropriate, as the mean proportion of the sale price to indexed assessed value is close to unity for 2001-2004 (and for 2005 no adjustment was made, as the assessments were in part based on 2005 sales data).

The Cervero criteria filter out an inordinately large percent of the total sample; Table 2 suggests that, if assessment values are indexed, greater than 50% of all observations would be filtered out for any give year. Thus, rather than using the Cervero criteria to filter out non-arms-length transactions, a less stringent filter was used that did not have an upper bound. The lower bound was set so that properties with a sale price of less than 50% of the indexed assessed value were excluded. As Table 2 indicates, this resulted in a much lower exclusion rate, of between 3.57% and 14.86%. To ensure that this filtered sample is approximately normally distributed, and thus meets Gauss-Markov assumptions, a density plot of assessed value was generated. Figure 3 confirms a general normal distribution, although with a slightly elongated tail.

Finally, an examination of the dataset revealed that properties built in 2005 or 2006 did not have the resulting increase in value factored into the recorded assessments. Table 1 demonstrates this relationship for those properties built in 2005 vis-à-vis those built in the previous four years; the proportion for 2006 is substantially worse, which can be expected since those properties built in 2006 had not been assessed at the time the data were purchased. Given the relationship presented in Table 1, all properties built in 2005 or 2006 were also omitted from the dataset.

3.4 Summary Statistics

After omitting errors, non-arms-length sales, and properties built in 2005 or 2006, the full sample consists of 340,364 single-family residential properties. Table 3 presents summary statistics for this full sample. As the efficiency of Cervero's stringent filter will be evaluated against this more moderate one, Table 4 and Table 5 present summary statistics for the Cervero filter and the filter used for the full sample; they present these statistics for sub-samples encompassing properties sold in each year from 2001 to 2005, along with one "pooled" sub-sample that contains properties last sold in any of those five years.

4. Empirical Results

Our changes in the endogenous factor and sample, how transit access is defined, and model specification require that the results be presented in three parts, each detailing how a particular change has affected the parameter estimates vis-à-vis our "benchmark" model. Thus, section 4.1 explores how the results react to the use of Cervero's strict filter or a more mild "modified" filter; section 4.2 details how the results obtained using assessment data compare to those estimated using sale price data; section 4.3 evaluates how different measures of transit access can affect the results; and section 4.4 examines how the community amenity vector affects the model estimates. After each section, the more "efficient" method is chosen and integrated into the benchmark model, so that the expanded-specification model presented in 4.4 matches our target model. Finally, section 4.5 presents an overview of the estimated transit access premium in nominal and relative terms. To facilitate things, details of the different modelling assumptions are presented in Table 6.

4.1 Sample Selection – Cervero Filter vs. "Lenient" Filter

To test how different the results would be if the Cervero filter had been applied, as opposed to the filter ultimately chosen, the benchmark model was estimated using each selection technique. Table 7 presents the results for Model 1, which uses Cervero's selection technique, while Table 8 presents results for Model 2, which uses the filter adopted in Section 3.2. Note that the parameter estimates are generally very similar, as are the standard robustness measures.

The direction of the parameter estimates is in keeping with what theory and previous empirical studies suggest. Thus, a property's price will tend to be higher for newer, larger properties that sit on larger lots. Properties tend to be more valuable if they are in areas with wealthier and comparatively younger populations. Contrary to traditional theory, this finds a positive relationship between property value and areas which are less racially white. This is in large part due to Portland's unique demographics, as Hispanics and Asians are the top two minority groups, rather than Black and Hispanic as in the U.S. as a whole. As Table 4 indicates, furthermore, the non-population weighted percentage of the population which is not white is 17.5% for Portland, which is significantly less than the 24.9% figure the 2000 Census estimated for the nation as a whole. Other potential surprises, such as higher values for properties in areas with more rental and vacant properties, are not statistically significant.

To properly interpret the parameter estimate on the variable that measures transit access, however, the appropriate transformation must be made. Recall that for Model 1 through Model 4, transit access is measured as a binary variable that takes on a value of 1 if a property is located within one-half mile of a transit stop. For these models, which exhibit a semi-log relationship between the continuous variable property value and the binary variable transit access, the appropriate transformation is:

(3) $$\ln(\text{Premium}) = \beta_{\text{Transit}} \text{ (when } x=1) - \beta_{\text{Transit}} \text{ (when } x=0)$$

(4) $$\text{Premium} = e^{\beta} - e^{0} = e^{\beta} - 1$$

Using this transformation, nominal transit access premiums can be derived; they are presented in Table 9. Note in particular how volatile these estimates are, as the estimated premium moves directly from a high of $7,055 for the 2004 sales sample to a low of $1,759 for the 2005 sales sub-sample. Such movements suggest that estimates for models that use only one years' sales data, which comprise the majority of the current literature, suffer from acute year-to-year volatility. Such volatility is largely incompatible with the bid-rent theory's assertion that the premium is derived from the value of reduced commute times, since both transit travel time savings and the marginal value of time to commuters should remain relatively stable over time.

Finally, despite the relatively poor results that both samples produce, conceptual qualms about the Cervero filter require the adoption of the less stringent filter used for Model 2. If the larger sample used in Model 2 is considered less biased, it appears that the bias introduced by the Cervero filter biases the estimated premiums upward, as all six estimates are higher for Model 1 than for Model 2. Thus the sample used for Model 2 is adopted.

4.2 Sample Selection – Sale Price vs. Assessment Data

As suggested in Section 3, the problem of selection bias in property sales data makes assessment data an attractive alternative. Although Table 3 provides some indication of potential selection bias, for comparative purposes the assessment data model was estimated using one the assessment model with three distinct samples. Three samples are used because, once assessment data is used, the 50% sale-to-assessment ratio used to omit dubious sales is no longer necessary (as the assessment values are not in question). Thus, four models are estimated: (1) a model using the Cervero selection criteria, but with assessment data instead of sales data (Model 3); (2) the "pooled" (2001-2005) sales sub-sample using the original "Modified" filter (Model 3.5); (3) the "pooled" (2001-2005) sales sub-samples (as in Model 3.5) without the 50% assessment "floor" (Model 4A); and (4) using the entire dataset (all years) without the 50% "floor" (Model 4B). Using three models, the results of each model can be directly compared.

Examining the model with the Cervero sub-sample, the parameter estimates are stronger for all variables in the Transportation Vector *T* for the model using assessment data (Model 3, Table 10) than they are for the model using sales data (Model 1, Table 7). Traditional robustness measures—such as the Adjusted R-Squared and Root

Mean Squared Error—suggest, at least on the surface, that these models are all similarly efficient. Note that all variables except the "Percent of the Census tract housing units which are vacant" variable are statistically significant for all sub-samples, including the single-year sales sub-samples (see Tables 10, 11, and 12).

If the "modified" filter is examined, the impact of moving from a sales-based measure of property value to an assessment-based measure is similar. It appears that the assessment-based model (Model 3.5, Table 11) produces much more consistent estimates of all the parameters from year to year than does the sales-based model (Model 2, Table 8). This is particularly important for the parameter estimate for the transit access variable, which has a range of 0.0038 to 0.0213 for Model 2 but a range of only 0.0112 to 0.024. Although mere numbers now, this difference is stark in nominal terms, as Table 13 notes.

Translating the semi-elasticities from Tables 8 and 11 into nominal transit access premiums provides a more complete comparison of the two models' results. The estimated premiums displayed in Table 13 appear, quite simply, to be more consistent for the assessment data version of the model (Model 3) than they are for the sales data version (Model 1). In addition, the estimates are uniformly higher. Although such discrepancies between the two estimates may be the result of fundamental differences between the way markets (e.g. sales prices) and local governments (e.g. assessment values) value each component of a property, it may also be indicative of an underlying bias in the sales data.

There is significant variation in the assessment dataset as well, variation that, according to Table 14, reduces the estimated transit access premium as the sample gets larger. These differences suggest a bias in the dataset. Since assessments are based on sales data, among other factors, properties that have not been sold recently (e.g. those properties added as the sample expands from 3.5 to 4A and 4B) may have less reliable assessments. Thus, although it appears that the smaller samples, such as the 2001-2005 sales interval sub-sample, provide a consistent yet large dataset, for completeness's sake both the 2001-2005 sub-sample (sample "A") and the full sample (sample "B") will be examined subsequent models.

4.3 Transit Selection – Distance Rings vs. Continuous Measurement

Although the benchmark model has undergone some important changes up to this point, the focus of the study—transit access and its associated property value premium—remains poorly measured. The benchmark model defines this access as a binary variable which indicates whether a property is inside or outside a given distance ring -

one-half mile; yet, the land value theories upon which this model is based imply that access premiums are manifest as a continuous gradient.

In an attempt to provide an access measure more in line with theory, transit access is now re-defined as the straight-line distance from every property to the nearest MAX light rail stop, and thus is a continuous variable. This GIS-based measurement is not new to the literature, and indeed there are variations (such as the “cost” distance measure discussed above). As noted previously, however, a simple straight-line distance measure seems to provide the best approximation of a Portland-area transit consumer’s time cost of travel to the nearest MAX stop.

Table 15 below presents summary results for this model; when interpreting the coefficient on the MAX access variable, note that this is an elasticity that varies with distance, as opposed to the semi-elasticity reported for Models 1-4. As such, the premium is estimated by the following method, for which the estimated premium is the integral from the initial distance to the final distance of the elasticity of the premium:

(5) $$\text{Premium} = \int [(\beta_{\text{Transit}})(\$Y_{\text{Mean}}/\text{Distance}_{\text{Transit}})(-1)]$$

It is worth noting that, in general, elasticities are not themselves constant. As such, elasticities estimated by multiple regression are calculated at the *mean value* of each of the variables. To generate valid results, then, the value of Y_{Mean} must be the mean assessed property value of a given sample. For sample “A” (Models 5A and 6A) the relevant figure, as displayed in Table 6, is \$272,468; for sample “B” (Models 5B and 6B), the relevant figure of \$272,242 can be found in Table 4.

Impressively, most parameter estimates vary little between Models 4A and 4B and their counterparts for Model 5. Except for the transit access premiums, which are fully discussed in Section 4.6 below, the parameter estimates are substantially the same across the two sets of models; so too are the relevant robustness tests.

4.4 Specification – Adding Community Amenities

Although some previous models have included alternative measures of transportation access, few studies outside the realm of environmental economics have sought to measure the value of access to local services like schools, parks, and other public goods. Seeking to capture the value added of these additional services, Model 6’s specification includes a “Community Amenities” Vector (Vector *A*), which is comprised of the distance from a given property to the nearest school, library, hospital, and park. This model—which uses assessment data, employs a continuous measure of MAX transit access (straight-line distance to the nearest station), and includes this Community Amenities vector—is the same as the target model presented in Section 2.3. The culmination of incremental changes, this model (and its results) is the one ultimately used to estimate the Portland-area MAX transit premium.

The model results, presented in Table 16, fit the pattern of the previous models. It is interesting, however, that proximity to the nearest library has the strongest positive externality, and is substantially larger than the estimated premium associated either school or park proximity. One explanation is that the size of the amenity effect is in large part a function of the proportion of the population that can derive benefit from it; thus positive externalities from one amenity (a library) may appeal to a larger proportion of the population than another amenity (a school). In addition, negative externalities such as crime, noise, and traffic may be larger for hospitals than for the other amenities, explaining the disamenity effect associated with proximity to a hospital.

4.5 Estimated Transit Access Premiums

Having settled on assessment data as the superior proxy of underlying property value, it is easy enough to present and compare the estimated premium from each of the five models. Using the transformations discussed above these parameter estimates for transit access can be converted into nominal premiums; these transformations are split between those models which measure transit access as a binary (Models 1-4) and those that measure access as a continuous variable (Models 5 & 6). Finally, the total estimated increase in estimated property value (and thus property taxes) is presented for the "target" model (Model 6) and as it has been modeled previously (Model 1).

Because Models 5 and 6 assume that property value will be a continuous function that decreases as value increases, it is more appropriate to consider the marginal premium gained by moving from one distance to another. To provide nominal terms, the discount for each property was estimated based on the mean value for Portland properties in the sample, which was a distance of 16,109 feet (3.05 miles) and the relevant mean value for that particular sample. This integration was calculated using STATA. Table 16 presents the marginal premium (e.g. premium of moving one foot) at different distances from transit.

The estimated discounts are notably different from those that would have been estimated by applying a "benchmark" model to the dataset. Although the discount estimates appear to be much greater for the target model than for previous models, this is in large part the because Models 1-4, which use the binary variable transit access selection method, average the access premiums (discounts) for all properties inside and outside of the distance rings.

Table 18, which is based on the marginal premiums presented in Table 17 and normalized at Property E (the mean distance to MAX in Portland), provides a clearer picture of how the premiums compare. The transit access premiums are presented in nominal terms in Table 18 and in terms relative to the sample's mean assessed value in Table 19; in both Tables the estimates increase uniformly from Model 4 to Model 6. Note that, since Property E is the "average" property in Portland, half the properties in our sample will be further away from MAX transit (and have a transit-access-related discount) and half will be closer to transit (and enjoy a premium); as we are most concerned with properties near rail mass transit, four properties are presented that are closer to transit than the mean, while two are presented that are farther than the mean. Figure 4, a density plot of the distance to the nearest MAX station, confirms that the vast majority of Portland-area properties are located within 20,000 feet (approximately 3.75 miles) of a MAX station.

Figure 4 also reveals how inconsistent the distribution of properties from MAX transit stations is. This suggests that the continuous measure of distance used for Models 5 and 6 is appropriate empirically as well as theoretically. Furthermore, the shape suggests that using the traditional binary distance ring approach would miss much of the variation in this distance to transit. Considering how variable the distribution of distance is, the estimates provided by Models 5 and 6 are much more efficient than the estimates provided by Models 1-4,

The elasticities are valid around the mean value of each variable. However, as the distance to the nearest transit facility moves away from its mean value, the presence of a transit access premium implies that, even if all other variables are held constant, the mean property value at any given point will be different than that at the sample mean (for properties closer to transit than the "average" property the mean will be higher; for properties farther from transit the mean will be lower). Although the net effect of this issue, according to how equation (5) is structured, should be to increase the value of the transit access premium (discount) for properties closer to (farther from) an MAX station than the "average" property, it is still worth investigating. For the purposes of this paper, however, the fixed-mean approach presented in equation (5) is adequate and consistent with previous studies.

As Tables 18 and 19 demonstrate, the target model (Model 6) estimates that a properties' distance to transit affects its value more than prior studies' methods would estimate. Premiums are estimated to be significantly greater for properties close to transit vis-à-vis previous methods' (Models 1-5) estimates, while properties far from transit, such as Property G, are estimated to have lower values under Model 6 than under previous models.

The municipal tax implications of these changes, and in particular the change from a discrete to continuous measure of transit distance, are enormous. As Table 20 indicates, the total increase in property values resulting from transit proximity is estimated to be more than twice as much for properties within a half mile of transit—the area which previous studies have considered as transit-accessible—and is substantially larger when properties further away are taken into account. Note that the total estimated value of transit for the Benchmark Model does not increase from the one-half mile to 5 mile perspective; this is because the Benchmark Model estimates a constant value of $4,929 for each house up to the one-half mile boundary, but $0 for each property past that line.

The result overall suggests "intensive" and "extensive" growth in the size of the estimated premium, as the new estimates provide larger estimates for the same properties (those within one-half mile) while also registering sizable premiums for new properties (those beyond the one-half mile boundary; for Table 20, this extends to 5 miles from transit). Most of the intensive growth in the estimated premium can be attributed to the continuous distance measurement, which allows for a declining marginal premium; the extensive growth is also primarily attributed to the continuous distance measurement, namely in the ability to distinguish between properties 1 mile from transit versus those 10 miles away.

5 Conclusions and Final Remarks

Overall, the model suggests that there are large benefits associated with access to Portland's MAX system. All iterations of the model, from Model 1 to Model 6, find significant transit access premiums. The change to a distance-related access measurement appears to have the most powerful effect on the estimated premium of the three changes that were made, but all changes produced significant improvements in the model. The move from a sales sample to an assessment sample decreased the volatility of the transit access premium estimate, as did the move from single-year to multi-year samples. The change from a binary to continuous measure of transit access transformed the premium estimates from a fixed to variable effect, which more closely approximates the shape of the theoretical land-rent gradient. Finally, the addition of community amenities indirectly impacted the estimated value of the transit access premium, and the estimated dis-amenity associated with proximity to hospitals was unexpected (although not unreasonable).

More broadly, the models' results suggest that transit access premiums—and thus the total value of transit—may be significantly larger than previously estimated. The use of more efficient estimates should increase the estimated property (and thus property tax) value of proposed mass transit projects, likely increasing the number of projects that pass cost-benefit analyses. Furthermore, the powerful effect that mass transit access has on property could be harnessed, perhaps with a transit-based Tax Incremental Financing district, to more efficiently pay for these transportation improvements; such an approach would prove more efficient at taxing those who benefit from such improvements than the more traditional sales tax-based approach.

Although this study attempts to bridge the divide between previous transit studies, and in the process make comparison between them possible, it is only a first attempt at doing so. Left unanswered is the question of whether these estimates and relationships between iterations of the model hold up when applied to other cities, especially those with less restrictive zoning and land-use laws. So too is the question of how the results and estimated premium for cost distances compare to those for straight-line distances; considering that many transit systems are built with significant Park-and-Ride capacity, cost distances may be appropriate for some studies. More fundamentally, the continuous measure of distance should allow future studies to far more accurately predict the actual impact of rail mass transit improvements on surrounding properties. This analysis can and should be extended to predict how such improvements directly affect the tax levies brought in by local governments, as such revenue estimates will ultimately tell whether a system's benefits outweigh the costs of construction and maintenance.

References

1 Alonso, W. (1974). *Location and Land Use; Toward a General Theory of Land Rent*. Harvard University Press, Cambridge, MA.

2 Baum-Snow, N. and M. Kahn. (2000). "The Effects of New Public Projects to Expand Urban Rail Transit." *Journal of Public Economics* 77, 241-63.

3 Bowes, D. and K. Ihlanfeldt. (2001). "Identifying the Impacts of Rail Transit Stations on Residential Property Values." *Journal of Urban Economics* 50, 1-25.

4 Becker, S. (2013). *The New Real Estate Mantra: Location near Public Transportation*. Center for Neighborhood Technology.

5 Cervero, R., C. Ferrell, S. Murphy (2002). *Transit-Oriented Development and Joint Development in the United States: A Literature Review*. Transportation Research Board of the National Academies, Washington DC.

6 Cervero, R. (2004). "Effects of Light and Commuter Rail Transit on Land Prices; Experiences in San Diego County." *Journal of the Transportation Research Forum* 43, 121-38.

7 Dewees, D. (1976). "The Effect of a Subway on Residential Property Values in Toronto." *Journal of Urban Economics* 3, 357-369.

8 Fejarang, R. (1994). "Impact on Property Values: A Study of the Los Angeles Metro Rail." *Transportation Research Board 73rd Annual Meeting*.

9 Gatzlaff, D. and M. Smith. (1993). "The Impact of the Miami Metrorail on the Value of Residences near Station Locations." *Land Economics* 69, 54-66.

10 Koutsopoulis, K. (1977). "The Impact of Mass Transit on Residential Property Values." *Annals of the Association of American Geographers* 67, 564-76.

11 Landis, J., S. Guathakurta, and M. Zhang. (1994). *Capitalization of Transportation Investments into Single-Family Home Prices*. Berkeley, Institute of Urban and Regional Development, University of California, Working Paper 619.

12 Metro (2006). "Regional Land Information Systems (RLIS) Lite." Portland, Oregon.

13 Rosen, S. (1974). "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition." *Journal of Political Economy* 82, 34-55.

14 Tri-Met. (2006). *Facts about Tri-Met*. October 2006. http://www.trimet.org/pdfs/publications/factsheet.pdf.

15 Tri-Met. (2007). *Tri-Met Fixed Route Service and Ridership Information*. http://www.trimet.org/pdfs/ridership/busmaxstat.pdf.

16 U.S. Bureau of the Census. (2000). "Census 2000 Summary File 1 100-Percent Data." Washington, DC.

17 U.S. Bureau of the Census. (2000). "Census 2000 Summary File 3 Sample Data." Washington, DC.

18 Voith, R. (1993). "Changing Capitalization of CBD-Oriented Transportation Systems: Evidence from Philadelphia, 1970-1988." *Journal of Urban Economics* 33, 361-76.

19 Wardrip, K. (2011). *Public Transit's Impact on Housing Costs: A Review of the Literature*. Center for Housing Policy – Insights from Housing Policy Research.

20 Weigher, J. and R. Zerbst (1973). "The Externalities of Neighborhood Parks: An Empirical Investigation." *Land Economics* 49, 99-105.

21 Weinstein, B. and T. Clower (2002). *An Assessment of the DART LRT on Taxable Property Valuations and Transit Oriented Development*. Denton, University of North Texas, Center for Economic Development and Research.

22 Wu, J., R. Adams, and A. Plantinga (2004). "Amenities in an Urban Equilibrium Model: Residential Development in Portland, Oregon." *Land Economics* 80, 19-32.

Figure 1: Transit Distance Rings and UGB Boundary Overlaid onto Portland Map

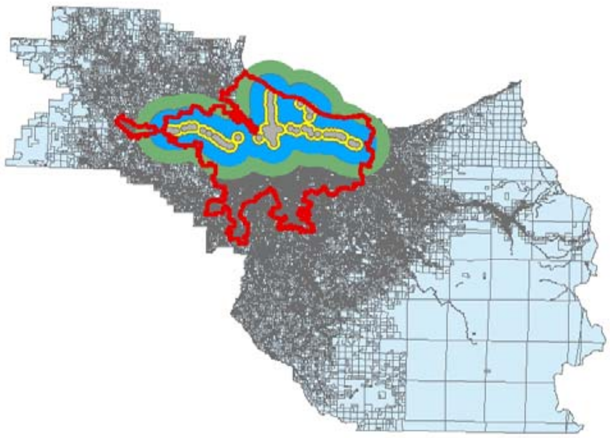

Figure 2: Density Plot of Total Assessed Value if Less than $100,000 (No Filter)

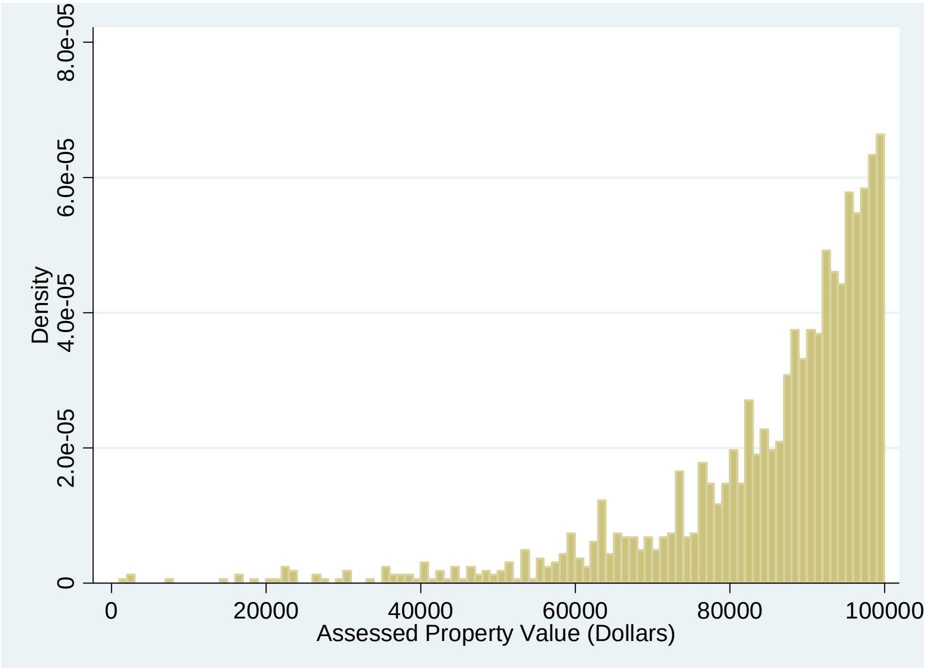

Figure 3: Density Plot of Assessed Value ("Full Sample" after Modified Filter)

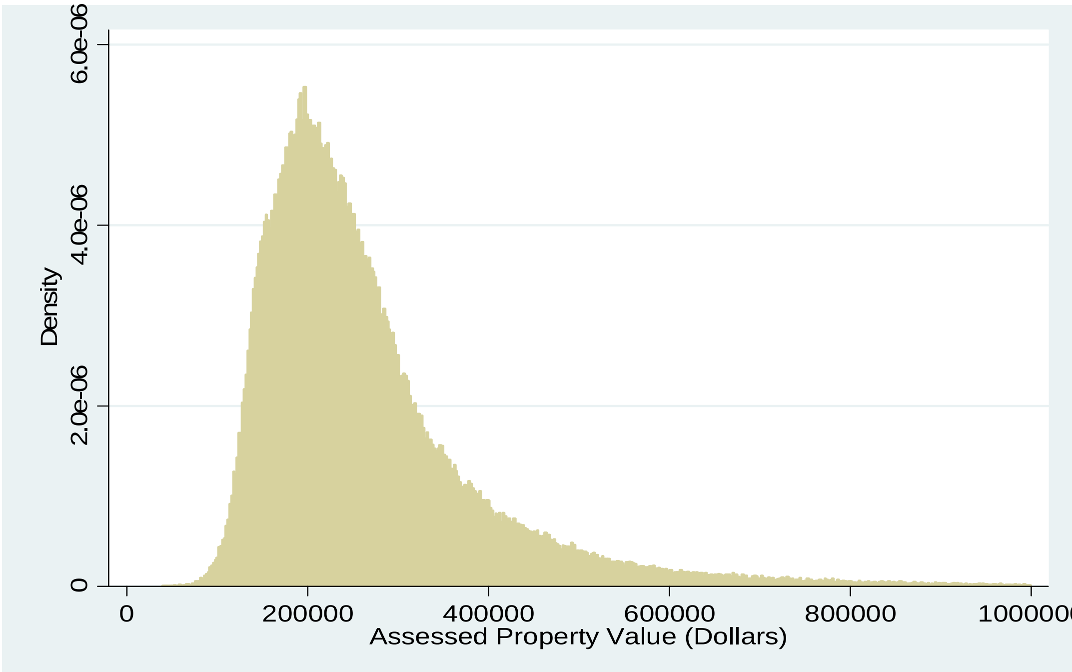

Figure 4: Density Plot of Distance to Nearest MAX Transit Station (Full Sample)

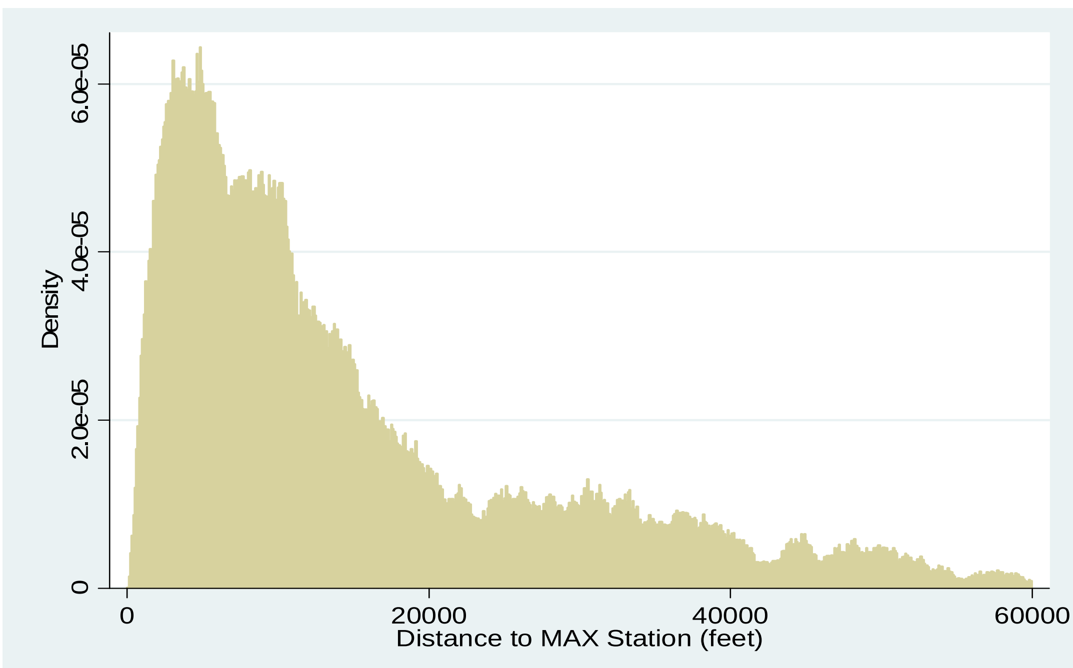

Table 1: Proportion of Sale Price to Indexed Assessed Value				
Year Property Sold	Mean Proportion	Minimum Proportion	Maximum Proportion	Number of Observations
2001	1.02	0.073	10.39	14,561
2002	1.00	0.056	8.36	16,625
2003	1.01	0.075	17.81	21,022
2004	0.99	0.069	18.40	25,411
2005	1.12	0.064	74.15	31,123
Year Property Built	Mean Proportion	Minimum Proportion	Maximum Proportion	Number of Observations
2001	1.11	0.087	11.40	4,384
2002	1.14	0.068	30.01	4,616
2003	1.05	0.053	25.94	4,722
2004	0.99	0.067	16.10	5,201
2005	1.27	0.044	74.15	4,252
2005=1.00				

Table 2: Breakdown of Properties Omitted from 2001-2005 Sales Samples, Using Different Selection Techniques						
		2001	2002	2003	2004	2005
Initial Number of Observations		14,834	16,897	21,318	27,023	33,540
Cervero (2004)	Percentage of All Properties Omitted Because Indexed Sale Price Less than 90% of Assessed Value	22.75%	28.54%	33.54%	36.51%	26.24%
	Percentage of All Properties Omitted Because Indexed Sale Price Greater than 110% of Assessed Value	33.75%	26.30%	22.89%	20.05%	38.26%
	Total Percentage of Initial Observations Omitted	56.50%	54.84%	56.43%	56.56%	64.51%
Modified Filter/ "Full Sample"	Total Percentage of Initial Observations Omitted	4.93%	4.38%	3.57%	9.92%	14.86%

Table 3: Single-Family Property Characteristics		
Variable	Mean	Std. Dev.
Land Value	$114,722	$67,635
Total Value	$272,242	$155,130
Building Size (Square Feet)	1,834	820
Building Age (Years)	44.3	29.0
Property Lot Size (Square Feet)	10,642	28,798
Distance (Feet) to Nearest:		
School	1,942	1,569
Library	7,488	4,937
Hospital	13,677	8,557
Park	1,567	1,972
MAX rail stop	16,109	14,902
Freeway	9,511	8,445
Surrounding Census Tract:		
Median Age	35.6	4.2
Median Household Income	$51,666	$15,594
Median Household Size (People Per Unit)	2.6	0.3
Percent of Census Tract who are Not White	17.5%	10.8%
Percent of Tract Housing Units which are Detached	69.0%	18.7%
Percent of Households in Census Tract which are Vacant	13.5%	4.5%
Percent of Households in Census Tract which are Rented	33.8%	16.0%
Percent of Households in Census Tract with Income Exceeding $100,000	14.8%	11.4%
Average Commute Time (Minutes)	23.9	2.8
Number of Observations	340,364	

Table 4: Single-Family Property Characteristics, Properties Sold in 2001-2005, Cervero's Filter												
	2001		2002		2003		2004		2005		Pooled 2001-05	
Variable	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Sale Price	$205,034	$104,732	$209,985	$121,661	$224,949	$125,835	$246,722	$136,023	$273,391	$143,349	$236,568	$131,878
Land Value	$114,359	$54,960	$103,575	$55,772	$97,843	$47,024	$99,554	$50,677	$114,010	$59,650	$105,440	$54,193
Total Value	$281,874	$143,988	$266,691	$155,742	$263,001	$148,402	$266,779	$147,069	$273,981	$144,677	$269,858	$147,961
Building Size (Sq. Feet)	1,872	768	1,832	780	1,839	813	1,860	829	1,843	799	1,849	802
Building Age (Years)	39.0	30.7	38.2	30.5	38.2	30.5	39.8	30.5	35.2	27.6	38.0	29.9
Property Lot Size (Sq. Feet)	8,436	8,740	8,235	8,862	8,507	10,072	8,682	16,729	8,548	10,258	8,504	11,788
Distance (Feet) to Nearest:												
School	1,945	1,338	1,961	1,333	1,944	1,300	1,974	1,417	1,985	1,378	1,964	1,359
Library	7,872	5,050	7,667	4,792	7,588	4,808	7,480	4,668	7,969	4,901	7,706	4,833
Hospital	14,190	8,409	13,660	8,352	13,501	8,762	13,263	8,451	14,959	9,023	13,916	8,657
Park	1,512	1,827	1,516	1,723	1,563	1,875	1,562	1,803	1,494	1,602	1,531	1,763
MAX rail stop	17,473	15,756	17,072	15,885	17,182	16,093	16,502	15,714	16,298	14,971	16,826	15,659
Freeway	9,860	8,278	9,743	8,493	9,799	8,992	9,339	8,568	10,535	8,922	9,861	8,699
Surrounding Census Tract:												
Median Age (years)	35.3	4.3	35.4	4.2	35.6	4.1	35.5	4.1	34.9	4.2	35.3	4.2
Median Household Income	$52,607	$15,342	$51,505	$15,390	$51,075	$15,996	$51,002	$16,126	$52,170	$15,549	$51,608	$15,741
Median Household Size (People Per Unit)	2.6	0.3	2.6	0.3	2.6	0.3	2.6	0.3	2.6	0.3	2.6	0.3
Percentage who are Not White	17.8%	11.2%	17.8%	11.0%	17.4%	10.9%	17.9%	11.0%	17.6%	9.6%	17.7%	10.7%
Percentage of Housing Units which are Detached	69.0%	18.8%	69.4%	18.6%	69.5%	18.9%	69.4%	18.9%	67.9%	18.5%	69.0%	18.8%
Percentage of Housing Units which are Vacant	13.4%	4.5%	13.4%	4.4%	13.6%	4.3%	13.6%	4.3%	13.3%	4.4%	13.5%	4.4%
Percentage of Households which are Rented	33.5%	15.9%	33.1%	15.6%	33.1%	16.1%	33.4%	16.2%	33.9%	15.6%	33.4%	15.9%
Percentage of Households with Income Exceeding $100,000	15.4%	11.2%	14.5%	11.2%	14.2%	11.5%	14.2%	11.7%	14.8%	11.2%	14.6%	11.4%
Average Commute Time (minutes)	23.9	2.7	24.1	2.8	24.3	2.9	24.3	2.9	24.2	2.7	24.2	2.8
Number of Observations	6,452		7,628		9,284		11,208		10,871		45,443	
As Proportion of Observations	43%		45%		44%		41%		32%		40%	

Table 5: Single-Family Property Characteristics, Properties Sold in 2001-2005, With Modified Filter/Full Sample												
	2001		2002		2003		2004		2005		Pooled (2001-05)	
Variable	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Sale Price	$207,010	$125,416	$217,532	$130,394	$234,872	$156,801	$252,983	$149,377	$285,397	$202,491	$246,544	$164,263
Land Value	$112,229	$61,198	$111,475	$60,401	$110,650	$56,598	$112,367	$62,490	$109,110	$63,921	$110,972	$61,283
Total Value	$274,414	$151,995	$274,813	$156,524	$274,741	$149,205	$279,417	$154,646	$262,620	$146,057	$272,468	$151,316
Building Size (Sq Feet)	1,833	779	1,833	792	1,843	811	1,861	821	1,808	802	1,835	804
Building Age (Years)	40.4	30.4	39.4	30.6	38.3	30.7	38.5	30.8	41.3	30.2	39.6	30.5
Property Lot Size (Sq Feet)	9,134	18,436	8,772	16,052	8,588	11,131	8,786	15,659	8,824	13,011	8,802	14,666
Distance (Feet) to Nearest:												
School	1,988	1,478	1,989	1,556	1,959	1,319	1,981	1,488	1,938	1,420	1,967	1,447
Library	7,677	4,978	7,701	4,869	7,846	4,986	7,783	4,937	7,597	4,862	7,717	4,922
Hospital	14,001	8,535	14,045	8,566	14,269	8,704	14,287	8,667	14,095	8,672	14,154	8,643
Park	1,534	1,844	1,532	1,791	1,531	1,767	1,531	1,793	1,502	1,665	1,524	1,760
MAX rail stop	16,761	15,325	16,495	15,227	16,429	15,184	16,379	15,254	15,809	14,729	16,302	15,106
Freeway	9,619	8,357	9,709	8,345	10,022	8,575	9,961	8,479	9,864	8,513	9,861	8,471
Surrounding Census Tract:												
Median Age (years)	35.5	4.3	35.4	4.2	35.3	4.2	35.2	4.2	35.1	4.1	35.3	4.2
Median Income	$52,052	$15,621	$52,140	$15,910	$52,148	$15,875	$52,313	$15,958	$51,072	$15,605	$51,876	$15,799
Median Household Size (People Per Unit)	2.6	0.3	2.6	0.3	2.6	0.3	2.6	0.3	2.6	0.3	2.6	0.3
Percentage Not White	17.5%	10.8%	17.7%	10.9%	17.7%	10.7%	17.9%	10.8%	18.2%	10.8%	17.9%	10.8%
Percentage of Housing Units which are Detached	68.6%	18.9%	68.8%	18.7%	68.9%	18.8%	68.9%	18.8%	68.5%	18.7%	68.7%	18.7%
Percentage of Housing Units which are Vacant	13.5%	4.4%	13.4%	4.4%	13.5%	4.4%	13.4%	4.4%	13.6%	4.4%	13.5%	4.4%
Percentage of Households which are Rented	33.6%	16.0%	33.5%	15.8%	33.3%	15.9%	33.5%	16.0%	34.2%	15.8%	33.7%	15.9%
Percentage of Households with Income Exceeding $100,000	15.1%	11.4%	15.1%	11.5%	15.0%	11.5%	15.1%	11.5%	14.2%	11.3%	14.8%	11.4%
Average Commute Time (minutes)	24.0	2.8	24.0	2.8	24.1	2.7	24.0	2.8	24.1	2.8	24.1	2.8
Number of Observations	14,103		16,157		20,557		24,342		28,556		103,715	
As Proportion of Initial Observations	95%		96%		96%		90%		85%		91%	

Table 6: How the Models Compare

Model	Measurement of Underlying Property Value	Sample Filter Used	Measurement of Access to MAX Transit	Specification
1	Sale Prices	Cervero	Binary	"Benchmark"
2		Modified		
3	Assessment Values	Cervero		
3.5		Modified		
4A		Modified without Threshold Restriction		
4B				
5A			Continuous	
5B				
6A				Expanded (includes "Community Amenties")
6B				

Table 7: Model Using Sales Price as Endogenous Factor, Cervero's Filter (Model 1)						
	2001	2002	2003	2004	2005	Pooled (2001-05)
Variable	Coefficient	Coefficient	Coefficient	Coefficient	Coefficient	Coefficient
Building Characteristics [C]						
Building Size (Square Feet) (Log)	0.5983	0.5776	0.5960	0.6020	0.5852	0.5908
Building Age (Years) (Log)	-0.0417	-0.0434	-0.0366	-0.0382	-0.0593	-0.0418
Lot Size (Square Feet) (Log)	0.0976	0.1088	0.0999	0.1149	0.1340	0.1117
Neighborhood Characteristics [N]						
Median Census Tract Age (Years) (Log)	-0.3314	-0.3256	-0.3562	-0.4553	-0.3489	-0.3753
Median Tract Income (Dollars) (Log)	0.0755	0.1045	0.1217	0.0590	0.1575	0.1108
Median Tract Household Size (Log)	-0.9474	-1.0496	-1.0832	-1.1757	-1.0059	-0.9831
Percent of Tract which is not White	0.0017	0.0024	0.0023	0.0026	0.0029	0.0021
Percent of Tract Housing Units which are Detached	0.0009	0.0022	0.0024	0.0028	0.0020	0.0019
Percent of Tract Housing Units which are Rented	0.0003†	0.0015	0.0017	0.0007**	0.0012	0.0013
Percent of Tract Housing Units which are Vacant	-0.0011**	-0.0008*	-0.0009**	-0.0003†	-0.0006†	-0.0004†
Percent of Tract with Income above $100,000	0.0108	0.0111	0.0106	0.0116	0.0092	0.0106
Transportation Access [T]						
Access to MAX (within 1/2 mile of MAX stop (1=yes, 0=no)	0.0165*	0.0221	0.0127*	0.0261	0.0064†	0.0181
Distance to Nearest Freeway (Feet) (Log)	0.0209	0.0228	0.0139	0.0176	0.0192	0.0175
Average Tract Commute Time (Minutes) (Log)	-0.2216	-0.2081	-0.2196	-0.2141	-0.3029	-0.2404
Constant	8.4697	8.1384	8.1846	9.1373	8.0099	8.3484
Summary Statistics						
Number of Observations (N)	6,452	7,628	9,284	11,208	10,871	45,443
F-statistic	1974.46	2499.21	3018.57	3905.39	3620.74	10430.07
Adjusted R-Squared	0.8107	0.821	0.8198	0.8299	0.8234	0.7626
Root Mean Squared Error	0.1673	0.1622	0.1701	0.1711	0.1695	0.2011
Unless noted, estimates are significant at the 1% level.						
*p<.10, **p<.05 (two-tailed test); † not statistically significant						

Table 8: Model Using Sales Price as Endogenous Factor, Modified Filter (Model 2)						
	2001	2002	2003	2004	2005	Pooled (2001-05)
Variable	Coefficient	Coefficient	Coefficient	Coefficient	Coefficient	Coefficient
Building Characteristics [C]						
Building Size (Square Feet) (Log)	0.5885	0.6077	0.6080	0.6138	0.5848	0.6030
Building Age (Years) (Log)	-0.0450	-0.0305	-0.0357	-0.0321	-0.0258	-0.0291
Lot Size (Square Feet) (Log)	0.0937	0.0914	0.0718	0.1141	0.1336	0.1014
Neighborhood Characteristics [N]						
Median Census Tract Age (Years) (Log)	-0.2375	-0.2959	-0.2208	-0.3589	-0.4320	-0.3363
Median Tract Income (Dollars) (Log)	0.0306†	0.014†	0.0968	-0.0274*	0.0226†	0.0099†
Median Tract Household Size (Log)	-0.9540	-1.0632	-1.0325	-1.2604	-1.3364	-1.1175
Percent of Tract which is not White	0.0012	0.0018	0.0019	0.0028	0.0032	0.0021
Percent of Tract Housing Units which are Detached	0.0016	0.0022	0.0032	0.0035	0.0032	0.0030
Percent of Tract Housing Units which are Rented	0.0007*	0.0006*	0.0023	0.0006**	0.0002†	0.0010
Percent of Tract Housing Units which are Vacant	0.0005†	0.0007*	0.0016	0.0001†	0.0006*	0.0008
Percent of Tract with Income above $100,000	0.0113	0.0122	0.0094	0.0125	0.0119	0.0116
Transportation Access [T]						
Access to MAX (within 1/2 mile of MAX stop (1=yes, 0=no)	0.0113†	0.0208	0.0044†	0.0213	0.0038†	0.0155
Distance to Nearest Freeway (Feet) (Log)	0.0161	0.0193	0.0208	0.0139	0.0178	0.0175
Average Tract Commute Time (Minutes) (Log)	-0.1654	-0.1684	-0.2680	-0.1644	-0.1545	-0.1857
Constant	8.5486	8.8328	8.1014	9.5192	9.4448	9.1395
Summary Statistics						
Number of Observations (N)	14,103	16,157	20,557	24,342	28,556	103,715
F-statistic	2830.44	3652.18	3229.24	5559.72	5156.24	14992.45
Adjusted R-Squared	0.7375	0.7598	0.6874	0.7617	0.7165	0.6693
Root Mean Squared Error	0.2112	0.2019	0.2379	0.2084	0.2288	0.2515
Unless noted, estimates are significant at the 1% level.						
*p<.10, **p<.05 (two-tailed test); † not statistically significant						

Table 9: Estimated Premium With Cervero's Filter and With Modified Filter, Using Sales Data

Sample Year	Cervero's Filter (Model 1)	Modified Filter (Model 2)
2001	$4,689	$2,758
2002	$5,959	$5,776
2003	$3,361	$1,212
2004	$7,055	$6,015
2005	$1,759	$1,000
Pooled (2001-05)	$4,929	$4,256

Table 10: Model Using Assessed Value as Endogenous Factor, Cervero Filter (Model 3)

	2001	2002	2003	2004	2005	Pooled (2001-05)
Variable	Coefficient	Coefficient	Coefficient	Coefficient	Coefficient	Coefficient
Building Characteristics [C]						
Building Size (Square Feet) (Log)	0.59849	0.57966	0.60124	0.60198	0.58163	0.59323
Building Age (Years) (Log)	-0.04310	-0.04538	-0.04085	-0.04358	-0.06424	-0.04636
Lot Size (Square Feet) (Log)	0.09859	0.10754	0.10090	0.11310	0.13555	0.11231
Neighborhood Characteristics [N]						
Median Census Tract Age (Years) (Log)	-0.34476	-0.38156	-0.38106	-0.47458	-0.36814	-0.41178
Median Tract Income (Dollars) (Log)	0.17272	0.16437	0.17223	0.10291	0.20441	0.17140
Median Tract Household Size (Log)	-0.93289	-1.02559	-1.04035	-1.11863	-0.98546	-1.02876
Percent of Tract which is not White	0.00240	0.00266	0.00254	0.00257	0.00307	0.00268
Percent of Tract Housing Units which are Detached	0.00050*	0.00196	0.00236	0.00268	0.00180	0.00195
Percent of Tract Housing Units which are Rented	0.00079*	0.00196	0.00219	0.00121	0.00158	0.00160
Percent of Tract Housing Units which are Vacant	-0.00106**	-0.00139	-0.00143	-0.00111	-0.00095**	-0.00125
Percent of Tract with Income above $100,000	0.00957	0.01092	0.01020	0.01111	0.00891	0.01010
Transportation Access [T]						
Access to MAX (within 1/2 mile of MAX stop (1=yes, 0=no)	0.02015**	0.03430	0.01820	0.02938	0.00789†	0.02169
Distance to Nearest Freeway (Feet) (Log)	0.02842	0.02984	0.01832	0.02075	0.02345	0.02436
Average Tract Commute Time (Minutes) (Log)	-0.28414	-0.27241	-0.27449	-0.27103	-0.34948	-0.30226
Constant	7.91098	8.05419	7.93969	8.95325	7.69616	8.11836
Summary Statistics						
Number of Observations (N)	6,452	7,628	9,284	11,208	10,871	45,443
F-statistic	1996.59	2559.33	3202.64	4058.81	3745.82	15394.57
Adjusted R-Squared	0.8124	0.8244	0.8284	0.8352	0.8283	0.8259
Root Mean Squared Error	0.1686	0.1634	0.1679	0.1699	0.1690	0.1691

Unless noted, estimates are significant at the 1% level.

*p<.10, **p<.05 (two-tailed test); † not statistically significant

Table 11: Model Using Assessed Value as Endogenous Factor, Modified Filter (Model 3.5)

	2001	2002	2003	2004	2005	Pooled (2001-05)
Variable	Coefficient	Coefficient	Coefficient	Coefficient	Coefficient	Coefficient
Building Characteristics [C]						
Building Size (Square Feet) (Log)	0.56644	0.58659	0.58680	0.59670	0.54460	0.57384
Building Age (Years) (Log)	-0.05219	-0.05039	-0.04992	-0.05761	-0.06962	-0.05801
Lot Size (Square Feet) (Log)	0.10368	0.10452	0.09959	0.11187	0.13081	0.11301
Neighborhood Characteristics [N]						
Median Census Tract Age (Years) (Log)	-0.37372	-0.39355	-0.41607	-0.45672	-0.40977	-0.41457
Median Tract Income (Dollars) (Log)	0.27391	0.24806	0.24674	0.18202	0.27773	0.24515
Median Tract Household Size (Log)	-1.01904	-1.00939	-0.97953	-1.01161	-0.98235	-1.00204
Percent of Tract which is not White	0.00319	0.00306	0.00269	0.00267	0.00294	0.00290
Percent of Tract Housing Units which are Detached	0.00142	0.00167	0.00159	0.00187	0.00119	0.00153
Percent of Tract Housing Units which are Rented	0.00211	0.00232	0.00195	0.00143	0.00202	0.00191
Percent of Tract Housing Units which are Vacant	-0.00060†	-0.00060*	-0.00079**	-0.00075**	-0.00092	-0.00078
Percent of Tract with Income above $100,000	0.00896	0.00988	0.00927	0.00969	0.00853	0.00923
Transportation Access [T]						
Access to MAX (within 1/2 mile of MAX stop (1=yes, 0=no)	0.01649	0.02262	0.01652	0.02402	0.01119	0.01768
Distance to Nearest Freeway (Feet) (Log)	0.02618	0.02802	0.02537	0.02469	0.02788	0.02656
Average Tract Commute Time (Minutes) (Log)	-0.33282	-0.29944	-0.37260	-0.34693	-0.36299	-0.34776
Constant	7.26647	7.28710	7.68685	8.32361	7.39525	7.63111
Summary Statistics						
Number of Observations (N)	14,103	16,157	20,557	24,342	28,556	103,715
F-statistic	3997.93	4980.13	6396.41	7396.54	7447.43	29962.05
Adjusted R-Squared	0.7987	0.8118	0.8133	0.8097	0.785	0.8018
Root Mean Squared Error	0.1824	0.1788	0.1789	0.1848	0.1943	0.1857

Unless noted, estimates are significant at the 1% level.

*p<.10, **p<.05 (two-tailed test); † not statistically significant

Table 12: Model Using Assessment Data, Modified Filter‡ (Models 4A & 4B)					
	Model 4B	Model 4A			
	Full Sample	2001-2005	1996-2005	1991-2005	1981-2005
Variable	Coefficient	Coefficient	Coefficient	Coefficient	Coefficient
Building Characteristics [C]					
Building Size (Square Feet) (Log)	0.56277	0.57523	0.57762	0.57537	0.57432
Building Age (Years) (Log)	-0.06235	-0.06126	-0.06166	-0.06240	-0.06205
Lot Size (Square Feet) (Log)	0.10930	0.11403	0.11159	0.11084	0.11012
Neighborhood Characteristics [N]					
Median Census Tract Age (Years) (Log)	-0.37627	-0.42388	-0.41194	-0.40624	-0.39621
Median Tract Income (Dollars) (Log)	0.28161	0.24040	0.23624	0.25484	0.26486
Median Tract Household Size (People) (Log)	-0.98032	-1.02024	-1.00162	-1.00490	-1.00281
Percent of Tract which is not White	0.00301	0.00303	0.00287	0.00298	0.00304
Percent of Tract Housing Units which are Detached	0.00124	0.00161	0.00149	0.00147	0.00143
Percent of Tract Housing Units which are Rented	0.00204	0.00188	0.00185	0.00196	0.00198
Percent of Tract Housing Units which are Vacant	-0.00093	-0.00075	-0.00086	-0.00092	-0.00095
Percent of Tract with Income above $100,000	0.00864	0.00929	0.00935	0.00907	0.00882
Transportation Access [T]					
Access to MAX (within 1/2 mile of MAX stop (1=yes, 0=no)	0.00842	0.01894	0.01392	0.01211	0.01120
Distance to Nearest Freeway (Feet) (Log)	0.02369	0.02652	0.02437	0.02416	0.02398
Average Tract Commute Time (Minutes) (Log)	-0.34790	-0.34954	-0.34506	-0.34643	-0.35003
Constant	7.25719	7.72002	7.72773	7.54242	7.42780
Summary Statistics					
Number of Observations (N)	340,364	110,398	171,175	210,298	235,158
F-statistic	94444.38	31761.39	50209.41	61409.45	67557.44
Adjusted R-Squared	0.7953	0.8027	0.8042	0.8057	0.8009
Root Mean Squared Error	0.1894	0.1850	0.1874	0.1820	0.1885

Note: All coefficients signficant at the 1% level

‡ For Assessment Data, the 50% threshold was not imposed, as this threshold filters out dubious sale price data but is not necessary when applied to assessments. Thus for this model and subsequent "A" and "B" samples, this threshold will not be used; the filter will still include the year built and minimum nominal assessment ($40,000)

Table 13: Estimated Premium Using Sales and Assessment Data, Cervero Filter		
Sample Year	Using Sale Price Data (Model 2)	Using Assessment Data (Model 3)
2001	$4,689	$5,737
2002	$5,959	$9,306
2003	$3,361	$4,830
2004	$7,055	$7,954
2005	$1,759	$2,170
"Pooled" (2001-05)	$4,929	$5,917

Table 14: Estimated Premium for Properties Sold in Selected Intervals, Using Assessment Data		
Sales Interval	Premium	Observations
2001-2005 (Model 3.5)*	$4,814	103,715
2001-2005 (Model 4A)	$5,211	110,398
1996-2005	$3,880	171,175
1991-2005	$3,375	210,298
1981-2005	$3,067	235,158
Full Sample (Model 4B)	$2,347	340,364
*The sample for Model 3.5 includes the 50% restriction, while the other samples do not		

Table 15: Model Using Assessment Data, Continuous Distance Function, and All Properties (Model 5)		
	2001-05 Sales Sample (Model 5A)	Full Sample (Model 5B)
Variable	Coefficient	Coefficient
Building Characteristics [C]		
Building Size (Square Feet) (Log)	0.57470	0.56201
Building Age (Years) (Log)	-0.06293	-0.06497
Lot Size (Square Feet) (Log)	0.11671	0.11263
Neighborhood Characteristics [N]		
Median Census Tract Age (Years) (Log)	-0.42407	-0.37571
Median Tract Income (Dollars) (Log)	0.22801	0.26895
Median Tract Household Size (People) (Log)	-1.01675	-0.98118
Percent of Tract which is not White	0.00263	0.00256
Percent of Tract Housing Units which are Detached	0.00157	0.00119
Percent of Tract Housing Units which are Rented	0.00170	0.00178
Percent of Tract Housing Units which are Vacant	-0.00066	-0.00081
Percent of Tract with Income above $100,000	0.00956	0.00890
Transportation Access [T]		
Distance to Nearest MAX Stop (Feet) (Log)	-0.01418	-0.01683
Distance to Nearest Freeway (Feet) (Log)	0.02730	0.02559
Average Tract Commute Time (Minutes) (Log)	-0.33401	-0.32931
Constant	7.92514	7.47334
Summary Statistics		
Number of Observations (N)	110,398	340,364
F-statistic	31885.36	95035.17
Adjusted R-Squared	0.8017	0.7963
Root Mean Squared Error	0.1879	0.1889
Note: All coefficients are significant at the 1% level		

Table 16: Model Using Assessment Data, Continuous Distance Function, Community Amenities, and All Properties (Model 6)		
	2001-05 Sales Sample (Model 6A)	Full Sample (Model 6B)
Variable	Coefficient	Coefficient
Building Characteristics [C]		
Building Size (Square Feet) (Log)	0.57513	0.56146
Building Age (Years) (Log)	-0.06644	-0.06939
Lot Size (Square Feet) (Log)	0.12121	0.11786
Neighborhood Characteristics [N]		
Median Census Tract Age (Years) (Log)	-0.39887	-0.35202
Median Tract Income (Dollars) (Log)	0.24915	0.28722
Median Tract Household Size (People) (Log)	-0.99187	-0.96321
Percent of Tract which is not White	0.00269	0.00263
Percent of Tract Housing Units which are Detached	0.00143	0.00103
Percent of Tract Housing Units which are Rented	0.00170	0.00172
Percent of Tract Housing Units which are Vacant	-0.00052	-0.00061
Percent of Tract with Income above $100,000	0.00927	0.00860
Spatial Vector [S]		
Distance to Nearest School (Feet) (Log)	-0.00432	-0.00477
Distance to Nearest Library (Feet) (Log)	-0.02956	-0.02814
Distance to Nearest Hospital (Feet) (Log)	0.01671	0.01429
Distance to Nearest Park (Feet) (Log)	0.00029	-0.00174
Transportation Access [T]		
Distance to Nearest MAX Stop (Feet) (Log)	-0.01705	-0.01910
Distance to Nearest Freeway (Feet) (Log)	0.02698	0.02543
Average Tract Commute Time (Minutes) (Log)	-0.33244	-0.32357
Constant	7.71920	7.32425
Summary Statistics		
Number of Observations (N)	110,398	340,364
F-statistic	25098.14	74721.8
Adjusted R-Squared	0.8036	0.798
Root Mean Squared Error	0.1870	0.1881
Note: All coefficients are significant at the 1% level		

Table 17: Marginal Premiums* at Selected Points				
Distance to Transit (feet)	Model 4		Model 5	
	4A (β=-.01418)	4B (β=-.01683)	5A (β=-.01705)	5B (β=-.01910)
500	$7.78	$9.16	$9.36	$10.40
1,320	$2.95	$3.47	$3.54	$3.94
2,640	$1.47	$1.74	$1.77	$1.97
5,280	$0.74	$0.87	$0.89	$0.98
16,109	$0.24	$0.28	$0.29	$0.32
26,400	$0.15	$0.17	$0.18	$0.20
52,800	$0.07	$0.09	$0.09	$0.10
*Per 1 foot change				

Table 18: Nominal Transit Access Premium for Selected Properties, vs. Property E*							
Property	A	B	C	D	E*	F	G
Distance to Transit							
(Miles)	0.09	0.25	0.50	1.00	3.05	5.00	10.00
(Feet)	500	1,320	2,640	5,280	16,109	26,400	52,800
Estimated Premium							
Model 1†	$4,929	$4,929	$4,929	$0	$0	$0	$0
Model 2†	$4,256	$4,256	$4,256	$0	$0	$0	$0
Model 3A	$5,917	$5,917	$5,917	$0	$0	$0	$0
Model 3B	$2,347	$2,347	$2,347	$0	$0	$0	$0
Model 4A	$13,513	$9,736	$7,038	$4,341	$0	-$1,922	-$4,620
Model 4B	$15,680	$11,197	$7,995	$4,794	$0	-$2,640	-$5,841
Model 5A	$16,248	$11,706	$8,463	$5,219	$0	-$2,311	-$5,555
Model 5B	$17,795	$12,707	$9,073	$5,440	$0	-$2,996	-$6,629
*Reflects Mean Transit Distance and Mean Total Value as reported in Table 4							
Note: Premiums reflect increase in value resulting from moving "average" property from E to that location (e.g to A)							

Table 19: Relative Transit Access Premium for Selected Properties, vs. Property E*							
Property	A	B	C	D	E*	F	G
Distance to Transit							
(Miles)	0.09	0.25	0.50	1.00	3.05	5.00	10.00
(Feet)	500	1,320	2,640	5,280	16,109	26,400	52,800
Estimated Premium							
Model 1†	1.83%	1.83%	1.83%	0.00%	0.00%	0.00%	0.00%
Model 2†	1.56%	1.56%	1.56%	0.00%	0.00%	0.00%	0.00%
Model 3A	1.80%	1.80%	1.80%	0.00%	0.00%	0.00%	0.00%
Model 3B	0.86%	0.86%	0.86%	0.00%	0.00%	0.00%	0.00%
Model 4A	4.96%	3.57%	2.58%	1.59%	0.00%	-0.71%	-1.70%
Model 4B	5.76%	4.11%	2.94%	1.76%	0.00%	-0.97%	-2.15%
Model 5A	5.96%	4.30%	3.11%	1.92%	0.00%	-0.85%	-2.04%
Model 5B	6.54%	4.67%	3.33%	2.00%	0.00%	-1.10%	-2.44%
*Reflects Mean Transit Distance and Mean Total Value as reported in Table 4							
Note: Premiums reflect increase in value resulting from moving "average" property from E to that location (e.g to A)							

Table 20: Estimated Transit Access Premium with Old (Benchmark) and New ("Target") Measurements		
Premium:	Benchmark Model	"Target" Model
Up to 0.5 miles from MAX	$140,738,000	$392,000,000
Up to 5.0 miles from MAX	$140,738,000	$5,640,000,000