The Amenity Value of Natural Views

Timothy L. Hamilton *
Department of Economics
University of Richmond
thamilt2@richmond.edu

Erik B. Johnson
Department of Economics
University of Alabama
ebjohnson5@cba.ua.edu

Abstract

We estimate nonmarket values for natural views in an urban setting. These views contain the aesthetics of natural areas commonly found in public parks and open space, and offer an aspect of property valuation that previous research is unable to disentangle from proximity to parks and open space. We incorporate machine learning techniques on google street view images to identify natural views in an urban setting. We find positive capitalization rates associated with household views of park-like properties. Estimates are robust to a variety of specifications, including models that are identified off of new developments on neighboring properties and falsification tests that help to rule out the effect of a broader neighborhood environment. From a policy perspective, our results inform as to the optimal size, location, and shape of open space. Furthermore, machine learning methods used in the construction of our view variable provide a potentially powerful tool for other nonmarket valuation studies.

Keywords: hedonic valuation; visual amenity; urban open space

JEL: Q51; R14; R31

*Corresponding author

Introduction

Hedonic housing price analysis has emerged as a common tool for nonmarket valuation in urban and environmental economics. Developed by Rosen (1974), the hedonic method estimates the value of location-specific amenities via their capitalization in housing markets and thousands of papers have implemented hedonic methods to measure the value of nonmarket amenities.[1] A major empirical limitation of this tool is quantifying some of the more unexplored amenities which may be important components of home valuation. For example, researchers know very little about hte value of aesthetics both within and outside a home, as well as noise, visual, and olfactory impacts of nearby land uses. These nonmarket amenities are complicated by the presence of other amenities that may confound estimates. Furthermore, only more recently have the presence of geospatial data provided researchers with the ability to observe and measure local amenities in a spatial context.[2] More recent advances in machine learning offer economists another step forward in being able to properly measure local amenities. More specifically, the combination of deep learning and image recognition gives researchers the ability to account for nuanced characteristics of an amenity that may not appear in a discrete set of attributes commonly available with local property records. Similarly, these techniques may be able to categorize observations with more accuracy than standard classification thus reducing measurement error.

In order to advance an understanding of these difficult-to-measure housing amenities, we turn to a novel application in the environment and urban economics space. We ask what is the value of views from a home that encompass greater concentrations of natural space?

[1] A large empirical literature includes applications to air quality (Smith and Huang, 1995), water quality (Walsh et al., 2017), hazardous waste sites (Greenstone and Gallagher, 2008), sunsihne (Fleming et al., 2018), and land use (Muehlenbachs et al., 2015). Hedonic estimation has also been used in the context of non-environmental goods, such as school quality (Black, 1999), crime (Linden and Rockoff, 2008), and drug pricing (Howard et al., 2015).

[2] See Bateman et al. (2002) for an overview of GIS in environmental economics.

Answering this question provides two contributions. First, we provide evidence on the value of residential development patterns that encompass natural views. This contribution will speak to zoning and development regulations that may help increase the value of housing to residents. Second, we provide a new methodology to valuing visual amenities that could have applications to any amenity that is visual in context.

The nonmarket valuation literature has extensively examined open space as an environmental amenity (McConnell and Walls, 2005). The positive capitalization effect of local green space is well-documented. For example, Klaiber and Phaneuf (2010) estimate the effect of various types of open space, including local parks, agricultural areas, and undeveloped space. They find positive effects and significant heterogeneity based on type of open space. Similarly, Turner and Seo (2021) find significant capitalization effects of both private and public green space. Studies often identify physical attributes of the open space, including land-cover designation, size, biodiversity measures, or recreational opportunities. In addition, analyses also measure the magnitude of a household's consumption of such open space through several different channels, including distance to open space, the percent of nearby land that exists as open space, or simply the presence of open space within some radius. Each of these metrics implies a slightly different consumption good. There is also considerable evidence for the value of local amenities and neighborhood designs associated with new urbanism, (Tu and Eppli, 2003; Matthews and Turnbull, 2007). These studies suggest the need for hedonic analysis that can uncover the mechanisms behind heterogeneous values of open space.

An important attribute of open space that may provide considerable benefits to nearby households is that it serves as a visual amenity. Defining the degree of visual amenity is a difficult task. Several studies (Paterson and Boyle, 2002; Walls et al., 2013) use a combination of GIS digital elevation models and land-cover data to characterize a household's view. These analyses provided considerably better measurement of the environmental attribute, relative to previous approaches. In this paper, we continue to improve upon hedonic analysis in this manner by developing a machine learning approach to measuring the visual appeal of properties, which allows us to account for an important location attribute. In particular, we train and utilize an image recognition algorithm to estimate the visual amenity that households consume. We simultaneously estimate the effects of different types of land use in an effort to separately identify the value of the visual component and other proximity-based values. Our primary contribution is a novel approach to quantifying the quality of a household's view. Other work has incorporated advanced methods for identifying environmental goods to be used in hedonic valuation. Franco and McDonald (2018) use remote sensing to measure urban greenness at an aggregate neighborhood level and finds significant capitalization effects. Yang et al. (2021) develops a Green View Index and finds a positive relationship between street-level greenery and commercial office rents. Other work has utilized image recognition methods to explore the price effect of the appearance (Johnson et al., 2020) and architectural (Lindenthal and Johnson, 2021) style of a home. This work thus uses advanced methods to expand the set of structural home attributes. Our work expands the use of such methods in hedonic analysis for valuation of environmental goods, focusing on the view amenity of park-like spaces.

Our analysis rests on the notion that open space conveys value in several ways, including recreational value, ecosystem services, and visual amenity value. Research in nonmarket valuation of open space tends to focus primarily on the first two aspects. We believe that proper characterization of visual amenity value is important for several reasons. First, the omitted visual amenity of open space may generate a bias in attempts to estimate recreation values. For example, proper measurement of the visual characteristics of open space ensures unbiased estimation of recreation demand and correctly informs policies that alter recreational opportunities. Second, estimates of these visual amenity values that are distinct from recreational value play an important role in determining the optimal provision of open space. Given a fixed quantity of land dedicated to open space (rather than residential or commercial development), the spatial allocation of that space could be configured to offer an area and services for recreation, or be allocated in a way to expand the perimeter and thus maximize household exposure to a view amenity through the use of, for example, boulevards, local squares, or small natural buffers. Determining the optimal balance between visual and recreation amenities requires the relative values of such amenities and potential complementarities between public goods play an important role in accurately estimating nonmarket values and in design of efficient policy (Albouy, 2018).

We estimate a hedonic model using housing transactions in the Denver, CO metropolitan area from 2008-2020. We use Google street view to obtain pictures facing away from each of these houses, replicating the view of each household from its front door. For each of these images, we use a convolutional neural network to score its natural view based on the similarity to natural areas of urban parks. Primary results are from a cross-sectional analysis that includes a set of spatial and temporal controls, along with structural housing attributes, to control for the variety of factors that could effect housing prices. We also present a number of alternative specifications that demonstrate the robustness of our empirical results. Finally, we estimate a repeat sales model using a subset of observations. Due to the necessity to match repeat housing sale transactions with multiple images that are close to the respective sales dates, creation of the dataset substantially reduces the sample size. Still, estimates in the repeat sales model show positive capitalization of the visual amenity and provide more evidence for a causal relationship.

This paper is organized as follows. First, we review the literature on hedonic valuation of open space, including analyses that consider visual amenities. We then introduce a method to define a visual amenity and discuss the few places such a technique has been used in the economics literature. After reviewing the current state of the literature, we present our empirical model and describe our data. Then, we present our image recognition algorithm in detail and its role in our hedonic analysis. Finally, we present results of a housing price hedonic model that accounts for visual amenity values and conclude with a discussion of our estimates.

Open Space Hedonic Analyses

While there exists an extensive hedonic literature that attempts to value open space, we limit our discussion here to focus on studies that suggest the value of visual amenities. To begin with, empirical evidence indicates that *type* of open space matters. Neumann et al. (2001) and Liu et al. (2013) find positive housing price premiums associated with proximity to National Wildlife Refuges, a form of open space with a natural aesthetic. These studies also find that effects are highly localized around wildlife refuges, potentially driven by houses with a view of the refuges. These results suggests that the corresponding visual amenity could be an important component. Shultz and King (2001) also report positive amenity values associated with natural areas and wildlife habitat, but find that housing prices decrease with proximity to city/county parks that may be used primarily for recreation. While this could be the result of aversion to congested areas, the visual difference in the type of open space may also be a contributing factor.

Several studies have attempted to directly value a view amenity. Weicher and Zerbst (1973) estimate a hedonic housing price model that includes indicator variables for homes that are adjacent to a municipal park, distinguishing between homes that face a park, back into a park, or face a park with heavy recreational use. They find positive effects associated with facing a park that does not have heavy recreational use and insignificant or negative effects in other cases of close proximity.

While Weicher and Zerbst (1973) define the visual amenity as a discrete variable, other studies use GIS software to define a magnitude of visual amenity. Paterson and Boyle (2002) employ a digital elevation model to determine the scope and distance of visibility from a household's location. This is combined with land cover data to measure the percent of a household's view defined by different land use categories. The study finds that visible developed land generates a negative price impact, but other visual measures are insignificant. However, the inclusion of visual measures does have substantial effects on other model estimates. Walls et al. (2013) take a similar approach and using ArcGIS Viewshed tool to characterize a household's view of different land covers. They find negative price effects of forest views, but positive price effects of farmland and grassy area views. Finally, Tagliafierro et al. (2016) also leverage GIS tools to construct a view measure. These authors use spatial data and the FRAGSTATS software (McGarigal et al. (2002)) to identify discrete objects that characterize a household's view. Results suggest that natural elements of a view have significant value.

Machine Learning for Image Classification

The methods in Paterson and Boyle (2002) and Walls et al. (2013) classify land cover into a specific category, such as residential, forest, or grassland. A view amenity is then defined based on the amount of each category in a household's viewshed. Similarly, Tagliafierro et al. (2016) rely on the identification of well-defined discrete attributes of a view. We take a different approach, relying on a machine learning image classification algorithm. This approach has two advantages. First, we are able to avoid potentially restrictive land-use categories and instead determine the elements of a visual amenity in the process of estimating an image classification model. This allows us to measure considerably more variation in land-use patterns. For example, areas classified as residential may vary substantially in other attributes. Second, by separately measuring view and land-use type we can directly estimate the capitalization effect of the view amenity.

We define the view amenity in our model based on its similarity to an ideal natural view. We discuss the required data and details of our technique later in the paper, but here we briefly outline our approach to measuring the view amenity. First, we use Google Street View to capture images that replicate the view from each household in our dataset: i.e. looking outwards from the property. We then use Street View images of parks that include natural amenities (trees, shrubs, grass) to define a park view. Using these park images, we estimate the parameters of an artificial neural network[3] to build a predictive model that can score any image based on its similarity to the park-like view. The model then allows us to predict the score of each household's view amenity on a $0-1$ scale.

The use of image classification to define environmental attributes is absent in the economics literature. Several studies, however, use machine learning image classification methods in the context of empirical economics. Glaeser et al. (2015) demonstrate the use of support vector regression to address a lack of economic statistics. Using data from Google Street View data, the authors predict median income in a census block group based on images of the neighborhood. The model not only fits the data well, but is able to predict out of sample median income in a separate city with a high degree of accuracy. Additional analysis shows that predicted income is strongly correlated with housing prices, indirectly linking image data to housing prices. In Naik et al. (2014), the authors employ a model that predicts the

[3] This method closely resembles a multivariate discrete choice model in which we define a park category in addition to several other image classifications. The pixel data of the digitized image serves as the input data used to predict a categorical outcome. We use the Inception V3 algorithm and transfer learning to minimize the sample size requirements of the training dataset.

perceived safety of a city from Street View images. The study gathers data on perceived safety based on location images via crowd-sourcing, which can then be used to estimate a support vector regression. Results show a strong positive correlation between a city's median family income and perceived safety. Similarly, Naik et al. (2017) use crowd-sourced data to determine the perceived safety of a location and estimate a predictive model that finds positive correlation between perceived safety and education levels in a neighborhood.

In general, the previously discussed line of research establishes a connection between visual data and economic outcomes but stops short estimating an economic value, such as a capitalization value or willingness to pay. Our study uses an artificial neural network to construct a meaningful economic variable, view amenity, that is difficult to quantify but likely plays a key role in economic transactions. We then use standard non-market valuation techniques to estimate the value of the view. Taking a similar approach in a different context, Guo et al. (2019) train a neural network to score individual faces on how attractive and aggressive they appear. This score is used as a covariate in predicting salaries among college football coaches. Machine learning techniques have been applied extensively and successfully in image classification exercises, but have so far had a limited role in empirical economics.

Empirical Model

We estimate a hedonic housing price function with the objective of identifying the marginal price of the visual amenity associated with open space. To accomplish this, we build a standard hedonic housing price model for the price of house i in subdivision j sold in year t. We use the natural log of nominal transaction prices. The hedonic price function includes a vector of housing attributes X_i, subdivision fixed effects γ_j, and tau_{tm} the interaction of a year

t dummies with Census block group m dummies (Census block groups are geographically larger units than subdivisions) to capture time trends that can vary across space.[4] The primary variables of interest are in the vector of park-related variables, $Park_i$. The hedonic function is

$$\ln P_{ijt} = \beta_0 + \beta_X X_i + \tau_{tm} + \gamma_j + \beta_P Park_i + \varepsilon_{ijt}, \tag{1}$$

where ε_{ijt} is a random i.i.d. error. The vector of housing attributes includes the number of bedrooms, gross living square feet, lot size, and dummies for the year the home was built. Our empirical work includes different specifications of the park amenity vector $Park_i$ to separately identify various park-related amenity values. We use four variables. First, $Prox_i$ measures the inverse distance from the house to the nearest city/county designated park. This captures the most conventional approach in the hedonic literature, treating open-space consumption as distance-related. Next, the variable $Park\_Adj_i$ is a dummy variable equal to 1 if the house is *facing* a designated park. This captures a slightly different type of consumption than $Prox_i$ in the form of a view amenity (though not necessarily a high view score from our image classification). For homes that are adjacent to a park on a different side of the parcel, $Prox_i$ controls for the amenity effect. We also include $Parkway\_Adj_i$, a dummy variable equal to 1 if the house is on a designated parkway, open space that may provide a view amenity but considerably less recreation opportunity than established parks. Finally, the variable $View\_Park_i$ denotes the visual amenity measured from our image classification algorithm. This measure represents the quality of the view that household i has when looking out its front door. As additional controls, we also include image classification scores for residential and commercial property in Equation 1, to more finely define a household's view.

All specifications include spatial fixed effects at the subdivision level. Beyond any location-specific unobservables, one may be concerned about unobservable attributes that are related

[4] Abbott and Klaiber (2011) demonstrate the importance of fixed effects at a fine spatial level.

specifically to a household's view. For example, the $View\_Park_i$ variable could be picking up effects associated with wealthy properties that typically have a high aesthetic quality. Therefore, a set of specifications includes the assessed value of the home across the street. By directly measuring and including the value of that property, we eliminate any potential confounding effect on $View\_Park_i$ coefficient estimate.

Section expands on the details of computing $View\_Park$, but a brief summary of the variable is necessary here. We define a high-quality view as a view that is similar to a forested/landscaped park. Therefore, a high-quality view includes the view of a designated city/county park but may also the view of undeveloped open space, well-manicured private land, or any other land that looks similar to a park. In addition, a household with a view of a designated park may have a lower $View\_Park$ value if that view is of a parking lot, recreational areas, or other elements of a park that do not provide a natural view. Our objective thus is to capture a purely visual amenity. Definitions for $View\_Park_i$ include both continuous and discrete designations. One specification defines $View\_Park_i$ as the predicted park-view amenity score. Since the view score is a somewhat indeterminate measure, we also estimate a dummy variable specification in which $View\_Park\_Ind_i$ is assigned a value 1 if the household's view score is above a particular threshold. Several specifications define this threshold as a percentile in the observed distribution of $View\_Park$, while one specification defines the dummy variable equal to 1 if the park score is the highest score among several categories. This approach fits better with the interpretation of the artificial neural network as a classification model. We discuss construction and interpretation of this variable in more detail following a discussion of our data.

Given a set of park-related variables, including or excluding specific variables in the $Park_i$ matrix allows us to identify the means by which open space generates value. These variables are distinct in important ways. Proximity to an established park, $Prox_i$, indicates value in the form of recreational opportunity. Controlling for proximity, the effect of facing a park suggests value derived from the visual amenity of an established park. Of course, both of these variables will include recreational use value, making it difficult to fully disentangle any visual amenity effects. Similarly, the use of parkways gets closer to picking up solely a view amenity due to the diminished recreational opportunity. While both $Park\_Adj_i$ and $Parkway\_Adj_i$ effects partially signify value related to a view amenity, the quality of this amenity is certainly not uniform. Therefore, we turn to our view score to measure a view amenity that is independent of any other potential values inherent in open space. This view score may distinguish a well-kept boulevard from a nominally designated parkway or the greenspace of a park from formal recreational areas. Moreover, it identifies visual amenities that exist in undesignated open space, such as private residential frontage and undeveloped natural areas.

We also estimate specifications to test for other price impacts. We include the variable Own_Park_i as an additional regressor to capture the visual amenity of a household's *own* property that may be capitalized into its price. This also provides a means of controlling for any potential unobservables at the the neighborhood level that may be related to its average view amenity, rather than the view specific to a household.

Finally, in addition to the cross-section model, we estimate a repeat sales model by taking the difference of Equation 1 across multiple sales. This specification reinforces a causal mechanism by differencing out unobserved property fixed effects.

Data

The first data component of our analysis is housing transaction data. This includes 121,045 observations of housing sale transactions in Denver, Colorado, from 2011-2017. Summary statistics for housing data are shown in the top panel of Table 1. In addition to the previously mentioned structural housing variables and transaction prices, we observe the exact address of each home. This allows us to obtain images of the home and the view across the street, via Google Street View, and to accurately measure distance to the nearest designated park. Next, the exact location of city/county designated parks is obtained from the Denver Department of Parks and Recreation. Figure 1 shows the location of designated parks in the study area. Using GIS software, we overlay housing transaction locations and park locations to measure the distance from the edge of each housing parcel to the edge of the nearest park. Geographic boundary files for parkways are available from the Denver Department of Parks and Recreation so we can similarly determine which houses are located on parkways. To determine which houses are facing an established park, we use a parcel and structure map of all homes in our study area from the Denver Open Data Catalog so we can identify homes that border an established park and are aligned such that the front of its structure faces the park.

The final data component is the park-view score for each housing observation. We use Google Street View to obtain pictures of the view from each property in our sample. Pictures are concentrated in years 2014-2016 (71%), with 20% prior to that and 9% in 2017. Across years, pictures are concentrated during May-October (95%). Using the Google Street View API combined with parcel and structure maps, we determine the proper angle so that the Street View image mimics that of looking out from the front of a house. These pictures define the view across the street for each household. We generate a similar dataset of pictures of each property itself. In the next section, we describe in detail the process of using these images to calculate $View\_Park_i$ and Own_Park_i, as well as scores for other categories.

Image Classification

To create the variables $View\_Park_i$ and Own_Park_i, we rely on an image recognition program developed as part of the Google Brain project and re-train it to detect an array of visual property-type attributes. Our objective is to score each property in our dataset based on its visual amenity. We define this visual amenity as similarity to the natural view of a designated park. This technique involves two steps: 1) Train the model on a subset of data to learn what image inputs define a park-like view, and 2) Apply the model to the sample of household view images to classify view types for each housing observation in our dataset. Classification in the second step involves predicting a view score for different view categories.

Training

We use Google's Inception V3 algorithm, a deep convolutional neural network. The Inception V3 classifier, part of the TensorFlow open-source software library, is trained specifically for image recognition on the Imagenet dataset.[5] However, our analysis is focused on identifying views that look like parks, which is outside of the classification domain of the model. Therefore, we re-train (i.e. re-estimate) the final layer of the neural network, effectively teaching the algorithm to distinguish park-like images from other urban land uses.[6]Thus, we continue to use the baseline parameterization of the model to classify images, but estimate the final

[5]Imagenet is an online database of classified images that has been used extensively for training and testing visual object recognition programs.

layer for specific context.

To train the algorithm to properly classify view types, we obtain pictures of parks, residential, commercial and other categories in the Denver area from Google Street View. The park images are manually chosen from among properties that are across from designated city parks. We use images that include naturally green areas (grass, trees, shrubbery) and remove images that include parking lots, fences, maintenance infrastructure, telephone/utility poles, and recreation infrastructure. These images serve as input data to estimate parameters of the algorithm. The online supplementary materials provide examples of images classified as parks for training purposes. We also train the model on other types of views to avoid false positives. Examples are provided in the supplementary materials.[7]

Each Street View image is broken down into a 2048-dimensional vector that captures the pixel representation of each image. Pixel data serves as the input data for the image classification algorithm. The deep convolutional neural network (Inception v3) fits approximately 25 million parameters to classify each image. To save computational cost, we use transfer learning by starting with an Inception-v3 model that is pre-trained on an Imagenet dataset containing 1.4 million images and 1000 image categories. We then re-train only the last bottleneck layer of the network according to our land-use categories. The use of transfer learning means we only have to estimate 8,116 parameters of the approximately 25 million parameters in the Inception v3 algorithm. Table 2 shows the percentage of each type of image (in-sample) that is properly classified by the trained model. Results suggest a high degree of accuracy.

[6] The advantage of using a pre-trained algorithm and estimating the final layer is that 1) less data and less variation in the data is required for identification and 2) the computational burden is reduced considerably. For more on implementation of this algorithm: https://www.tensorflow.org

[7] See Figures S.1 and S.2 of the supplementary materials.

Prediction

We next use the classification algorithm to estimate how similar is each household's view to the park, residential, commercial, and other categories. The image classification algorithm is trained to classify images based on how much the image resembles each of the categories. Classification assigns a score for each category to an image. We use the park category score as a measure of the visual amenity. We obtain images of the view from each house in our dataset and an image of the house itself so that we can then calculate $View\_Park$ and Own_Park. $View\_Park$ measures how similar the view across the street from the property is to a park, while Own_Park measures how similar the property itself is to a park. The supplementary materials show example observations of view images along with the predicted $View\_Park$[8] Figure 2 displays two histograms of predicted park scores. Panel 2a shows the full distribution of predicted park view scores. Note the high concentrations of predictions at very low end. In Panel 2b, we show only the distribution of predicted scores above 0.05 to better illustrate variation in predictions.

An important component of this analysis is controlling for the amenity value of proximity, presumably a recreation-based value, while estimating the amenity of a view. For econometric identification, we need 1) properties that have high park scores but are not in close proximity to parks and parkways and 2) properties that have low park scores and are in close proximity to parks and parkways. The correlation between between parks scores and our proximity measure is 0.004. Similarly, the Spearman rank correlation is -0.019. Spearman rank correlations of $View\_Park$ with $Park\_Adj$ and $Parkway\_Adj$ are 0.11 and 0.12, respectively. These low correlations indicate variation in the data due to view amenities that are due to properties other than designated parks. In addition, proximity to an established park does not ensure a high-quality view amenity. We are therefore confident that our con-

[8] See Figure S.3 of the supplementary materials.

structed view scores have the necessary variation to properly identify the value of a view amenity.

In Table 3 we report the mean and standard deviation of our park view score for deciles of the proximity distribution. Mean and standard deviation of predicted park score are calculated for all observations that are bounded below and above by proximity values. Park scores are relatively consistent throughout the proximity distribution. Similarly, Table 4 shows mean park view score for different values of distance-based park measures. We compare conditional means for observations within close proximity (95^{th} percentile), those adjacent and nonadjacent to parks, and those adjacent and nonadjacent to parkways. Summary measures are slightly different, reflecting differential access to views, but our approach preserves enough random variation for identification. It is evident that park view score is relatively independent of our distance-based measure. Finally, Figure 3 shows histograms of $View\_Park$ across park and parkway adjacency to illustrate a fuller picture of variation in independent variables. These histograms illustrate that observations that are adjacent to parks or parkways tend to have higher concentrations of park view scores at the lower-middle end (i.e. "good" views around 0.2) and at the extreme upper end (i.e. "great" views close to 1.0). Overall, these statistics demonstrate a joint distribution of variables that supports empirical identification.

Image Analysis

Before using the predicted image classification scores as a hedonic covariate, we explore the fundamentals of these scores in more depth. We use a second image recognition algorithm, Google's TensorFlow object detection model, that identifies particular objects of each photo. Since we are primarily interested in separately classifying park, residential, and commercial images, we identify trees and buildings.[9] This algorithm is already trained to recognize such objects. The supplementary materials show examples of two images, one with a high park score relative to its residential score and one with a high park score relative to its commercial score, in which this algorithm isolates trees and buildings. The figure illustrates the use of image segmentation to produce bounding boxes that isolate individual objects and measure their relative size.[10]

We calculate the area of the image that includes trees and buildings, respectively, as a percent of the total area and examine the correlation with park scores. Table 5 shows estimates from regressing predicted scores from our image classification algorithm on individual objects of the image. The dependent variable in these two models is the park score differential, defined as $View\_Park - View\_Res$ or $View\_Park - View\_Com$, where $View\_Res$ and $View\_Com$ are predicted scores for residential and commercial classification, respectively. Regressions also include controls for portions of the image in which trees overlap buildings, or vice versa. For the park-residential differential, results demonstrate a positive and significant effect of tree coverage and negative and significant effect of buildings. The park-commercial differential regression shows nearly identical results. Both cases provided strong corroboration that our park score prediction is driven by park-like views.

We also estimate a spline function to fit a more flexible relationship between parks score and tree coverage, while controlling for buildings and the overlap of objects in the image. Figure 4 shows estimated spline functions for $View\_Park - View\_Res$ or $View\_Park - View\_Com$. The park-residential differential in the first panel appears to be driven by tree concentrations. At lower values of the tree coverage, we do not see a strong positive relationship, likely due to typical residential properties with trees on the property. This result suggests that our image

[9] Wan and Lindenthal (2023) provide a more general approach to examining the workings of similar image classification algorithms.
[10] See Figure S.4 of the supplementary materials.

classification is not overly sensitive to the existence of a tree. In the second panel of Figure 4, we again see a positive and significant effect of tree coverage in distinguishing between park and commercial views. This relationship appears to hold across the entire distribution.

Results

We first present estimates from regressions that exclude park-view image scores. These baseline results replicate the conventional approach to measuring amenity values associated with natural areas. All regressions include the full set of housing attributes in Table 1, subdivision fixed effects and block-group time trends, with standard errors clustered at the subdivision level. Table 6 shows estimates and standard errors for specifications that include proximity to the nearest park ($Prox$), adjacency to a park ($Park\_Adj$), and adjacency to a parkway ($Parkway\_Adj$). Columns 1 and 2 indicate that park proximity and adjacency are both statistically insignificant, while column 3 implies adjacency to a parkway leads to a 9% decrease. The full specification in column 4 shows similar results. The negative effect of parkway adjacency is likely due to traffic externalities such as noise, air pollution, and safety concerns. The lack of a price premium on the other park variables is consistent with the existence of some positive benefits of parks, coupled with potentially negative effect of activity near parks. There exists a large literature that finds negative housing price impacts of parks.[11]

In Table 7, the predicted park view score $View\_Park$ is included in each specification to directly capture the view amenity. To ensure the relevance of views, we restrict our sample to include only homes for which the corresponding Street View image was captured within 1 year of the sale date. The estimated coefficients on proximity, park adjacency, and parkway

[11] See, for example, Smith et al. (2002) and Anderson and West (2006).

adjacency are consistent with earlier results. The coefficient on $View\_Park$ is positive and significant. The score variable is measured on a 0 to 1 scale, so this coefficient can be interpreted as a price premium for the highest-quality park-like view relative to no park-like view. The impact of an increase in score from 0 to 1 is approximately an 4.6% increase in price. This corresponds to a change in view from a typical commercial or residential property to that of a park-like natural area. Given a median home value in our sample of $330,000$, this effect translates to a price increase of $\$15,180$. Alternatively, a one standard deviation increase in score corresponds to a roughly 1.1% housing price increase. In addition to a positive and significant capitalization effect, inclusion of the view score has no significant impact on the estimated effect of other park-related variables, suggesting that the view score is indeed capturing a pure view amenity. We also note that there is a slight increase in the magnitude of the negative coefficient on $Park\_Adj$ relative to the baseline model. Without explicitly controlling for $View\_Park$, its positive capitalization effect of $View\_Park$ is partially captured in $Park\_Adj$, confounding the effect of the view amenity. Table A.1 presents coefficient estimates for structural housing attributes for our main specifications.

We utilize several other spatial controls to ensure that the effect of $View\_Park$ is capturing the park-view amenity. The image classification algorithm used in this study focuses on identifying park-like views, but also includes *commercial*, *residential*, and *other* as additional categories. The *other* category primarily captures views that are dominated by road intersections, but also includes the view of parking lots or capital infrastructure. Of course, every household has a view so that the predicted $View\_Park$ score measures the degree of park-like view, rather than other potential categories. We therefore estimate our model with the inclusion of view scores for commercial and residential properties covariates, as well as $View\_Park$. In Table 8, we re-estimate the four previous specifications and include the predicted score for commercial and residential views, $View\_Res$ and $View\_Com$, additional categories from our image classification algorithm. Furthermore, to address the potential confounding effect associated with proximity to wealthy households, we also include the natural log of the total assessed value of the property across the street from a house, $Across\_Value$. In these specifications, the coefficient on $View\_Park$ is larger than previous estimates, suggesting a housing price effect of roughly 6%. Residential and commercial views are marginally statistically significant. The positive coefficients should be interpreted in the context of $View\_Other$ (primarily road intersections) as an omitted measure. Results are nearly identical when assessed land value is used in place of total assessed value.

In Table 9, we show estimates in which our park view variable is defined as a dummy variable equal to 1 if the view score is above a particular threshold. The threshold is determined by quantiles of the observed distribution of $View\_Park$. We estimate several different specifications where the dummy is equal to 1 if the park view score is greater than the q^{th} quantile. In the final specification, the dummy is equal to 1 if the park view score is the largest among all view categories. The rows of Table 9 correspond to specifications with different q thresholds. All regressions include controls for proximity, park adjacency, and parkway adjacency, but we report only coefficients on the park score. A general pattern emerges in these estimates, where the estimated effect of a park view increases when the definition of such a park view becomes more stringent. We see a positive and significant price impacts of a park-like view only when it is the dominant view type, indicating a price increase ranging from 3.0% to 3.5% for having a park-like view.

Our results provide strong evidence for the capitalization of a view amenity in housing prices.[12] While the continuous magnitude of the view score variable is somewhat difficult to interpret, the monetary value of having a park-like view, relative to a purely residential or commercial view, is several percentage points of a home's value. Moreover, location-

[12]We also estimated regressions that include interaction effects among proximity and view variables. Estimates are consistent with earlier results but interaction coefficients add little to the overall analysis. Results are available upon request.

based measures that may coarsely include access to a view fail to capture the amenity value of a natural view. We conduct a back-of-the-envelope calculation to determine the total capitalized value of the view amenity in the study area. We take the average observed view amenity and housing value in our sample and extrapolate to the total number of homes in entire City and County of Denver. We are therefore assuming that our observed transactions are a random sample of the locality.

The average home has a $View\_Park$ score of 0.0884 and the 2010 census reports 622,900 owner-occupied homes in the city. We use a coefficient estimate of .0463, from the specification in which park view score is treated as a continuous variable. With an average housing price of $\$444,657$, an average capitalization value of \$1,819 leads to an aggregate value of \$1.13 billion. Using 95% confidence intervals on the coefficient estimate, the total capitalization value is between \$1.84 - \$3.02 billion.

The average home has a $View\_Park$ score of 0.1073 and the 2010 census reports 622,900 owner-occupied homes in the city. We use a coefficient estimate of .0818, from the specification in which park view score is treated as a continuous variable. With an average housing price of $\$444,657$, an average capitalization value of \$3,902 leads to an aggregate value of \$2.43 billion. Using 95% confidence intervals on the coefficient estimate, the total capitalization value is between \$0.688 - \$1.58 billion.

Repeat Sales Model

To better control for location unobservables and demonstrate a causal relationship between the view amenity and housing prices, we estimate a repeat sales model using homes that were sold twice during the study period. This empirical specification estimates a first-differenced specification of Equation 1 for a home sold in period t and in period k,

$$\ln P_{ijt} - \ln P_{ijk} = \beta_X (X_{it} - X_{ik}) + (\tau_t - \tau_k) + \beta_P (Park_{it} - Park_{ik}) + \varepsilon_i. \tag{2}$$

The cross-section analysis uses housing transaction during the 2008-2018 time period. To create the dataset, additional data from 2019-2020 is required to match the available Google Street View images.[13] The reported variables in this additional data also require us to reduce to set of X variables to only include changes in the number of Bedrooms, Bathrooms (full and half), and Stories. Thus we lose total square-footage and lot area. While changes in square-footage will be mostly captured with the available variables, lot area is unlikely to change over time.

We estimate two different specifications that address time effects differently. The first is a direct application of Equation 2 that includes a tk fixed effect to capture the combination of sale years. We also estimate a version of Equation 2 that deflates P_{ijt} using Federal Housing Finance Agency annual housing price indices for Denver and does not include any time fixed effects. Given the relatively small size of the repeat sales dataset, this approach helps to avoid the considerable loss of estimation power.

Table 10 displays estimates for the two approaches to time effects, as well as with and without *commercial* and *residential* view scores. Results continue to indicate a positive effect of the visual amenity. The coefficient on $View\_Park$ remains statistically significant and similar in magnitude to cross-sectional estimates. Column 4 in Table 10 shows a much larger estimated effect of 7.5%. Across specifications, standard errors are considerably larger but statistical significance remains. Generally, the results of the repeat-sales analysis provide

[13] Google Street View images in the study area are heavily concentrated during 2013-2016 and 2019-2020.

strong evidence for a positive capitalization effect caused by the view amenity.

Robustness

We estimate a number of additional specifications to address the sensitivity of our empirical results. While the repeat sales model provides evidence of a causal effect, the required reduction in observations due to finding multiple sales is substantial. In most cases, this makes meaningful robustness checks difficult. We therefore continue to emphasize the results of the cross-sectional model as well-grounded estimates. We briefly discuss the impact of the relative timing of pictures and sales in the repeat sales model, but focus our robustness checks on the cross-sectional model.

Timing of Picture

We consider variations in the sample due to the relative timing of pictures and sales. Recall that our main results restrict the sample to homes whose sale took place within 52 weeks of the date its corresponding Street View image was taken. Table A.2 shows summary statistics for the number of weeks between the sale and image capture for each year. There does not appear to be any pattern over time. Still we estimate regressions with variation in this restriction as a sensitivity analysis. For cross-sectional models, we implement three variations: images taken within 26 weeks of sale, images taken within 78 weeks of sale, and images take 78 weeks before sale. For the repeat sales model, we estimate the regression with observations for which corresponding Street View images are taken within 78 weeks prior to the sale.[14]

Results of variation in timing restrictions for the cross-sectional model are shown in Table A.3. Results are relatively similar to our main empirical results. When the time frame is reduced to 26 weeks, the estimated coefficient decreases and has a p-value slightly over 0.1. Given the size of standard errors on this estimate and those from earlier models, we do not attempt to attribute the small coefficient to anything other than sampling variation. The other specifications suggest a significant coefficient that is close in magnitude to earlier estimates. Table A.4 presents results from the repeat sales models with a image timing restriction of 78 weeks prior to sale. Models include the same two time fixed-effects specifications reported in the main results. Coefficient estimates double in magnitude, but large standard errors (the sample includes on 268 observations) suggest that we should not interpret these estimates as statistically different from previous specifications.

Additional Robustness Checks

The following robustness checks are applied only to the cross-sectional model. First, we design a falsification test to rule out the possibility that our hedonic results are driven by neighborhood effects. For each observation, we choose a random home in the same neighborhood and use that home's park score to re-assign the score of the base observation. We then re-estimate the hedonic regression. By choosing a home that is out of view but in the same neighborhood, this regression will pick up potential neighborhood effects rather than a home's specific view. We define neighborhoods by looking at areas around a home bounded

[14] In both models, considering only 52 weeks prior to sale reduces the sample size considerably due to the timing for Google Street View images in this area. Similarly, in the repeat sales model, restricting the dataset to observations within 26 weeks of the sale reduces the sample to just over 100 properties. Picture dates in our sample are concentrated in June through October. Home sales in the spring, which are potentially important for observing consumers' home preferences, are lost in our analysis when the time between pictures and sales is overly restrictive.

by rings of distance $d1$ and $d2$ meters from the home. Specifications include neighborhoods defined by homes within $50-100$ meters, $50-150$ meters, $100-150$ meters, and $150-200$ meters.

Coefficient estimates on the $View\_Park$ variable are reported in Table A.5. When the falsified observation is defined by close proximity (50-100 meters), results are insignificant with small magnitudes. If previous results were driven by the overall impression of a local area (finer than subdivisions, which are controlled for with fixed effects), these regressions should return similar coefficients to baseline models. Thus our falsification text supports our initial findings.

Next, we recognize that the presence of view amenities may not be exogenously determined. Neighborhoods may have zoning restrictions or historical architectural trends that lend themselves to high $View\_Park$ scores. Much of this is controlled for through the inclusion of fixed effects at a fine geographic scale in all specifications. We strengthen this by including a variable that measures a home's own view score: i.e. the view of the home. We use the same image classification algorithm discussed earlier, matching Street View images to the home in the image. Thus we control for the appearance of the home itself, in addition to its view. This will capture any unobserved attributes that may be correlated with general views in an area, ensuring that our model is estimating the impact of an individual's home view. Table A.6 reports estimates when $View\_Own$ is included in the model. Estimates on $View\_Own$ are statistically insignificant. Importantly, the coefficient on $View\_Park$ remains positive and statistically significant. Estimates on this view score are close to those of our main specification and indicate considerable capitalization of the views.

In Table A.7, we provide coefficient estimates for several additional robustness checks for variations on the model in Equation 1. In Column 1 we use only transactions after 2012, removing potential effects of the 2008 housing market crisis and its short-term aftermath. Estimates are qualitatively similar. The $View\_Park$ coefficient retains statistical significance but is smaller in magnitude. Results in Column 2 use month-of-year fixed effects to control for potential seasonality.[15] Only fixed effects for January, February, and December (unreported) are statistical significant, but the $View\_Park$ coefficient is similar to earlier results. We also estimate a specification that interacts the $View\_Park$ variable with an indicator variable for whether the transaction took place during the spring-summer months (March-September). Again, results associated with view variables are essentially unchanged. The coefficient on the interaction term is 0.0023, relative to a baseline coefficient of 0.0451, and is statistically insignificant. Next, we limit the sample to only single-story structures to focus on homes that only have a view from the ground level. The coefficient on $View\_Park$ is slightly smaller but not statistically different from previous estimates. Columns 5 and 6 include additional spatial amenities that could be driving variation in housing prices.[16] Controlling for proximity to golf courses, historical sites, and points of interest, we find no change in the $View\_Park$ coefficient. We also do not find statistically significant effects on any of the spatial variables, likely due to the inclusion of subdivision fixed effects. Column 6 includes proximity to open space as measured by undeveloped parcels. This final specification addresses potential concerns about the correlation between $View\_Park$ and proximity to local open space. We do not find a statistically significant coefficient on the open space variable, and the $View\_Park$ coefficient is not affected.[17]

Finally, we consider the predictive strength of our image classification. Recall that the algorithm assigns a score to each potential image category: *parks*, *residential*, *commercial*, and *other*. While the park score is meant to indicate a quality of view, it may also partially reflect the level of uncertainty in the prediction. Our earlier results address this with a set of

[15] Fixed effects refer to the date of the housing sale transaction, not the date of the picture.

[16] Denver Open Data Catalog: https://denvergov.org/opendata

[17] We were only able to access usable data for 2018 so the open space measure is imperfectly measured.

specifications in which we use an indicator variable equal to 1 when $View\_Park$ is greater than some threshold. We examine this further by focusing on predictions with a relatively high degree of certainty. For each observation we calculate the Herfindahl-Hirschman Index (HHI)[18] as the sum of the squared category scores. Since the four category scores are restricted to sum to 1, the HHI is bounded between 0.25 and 1. High index values indicate the existence of a single category with a high score, while low index values are driven by scores that are roughly equal across categories. Therefore, we estimate our model using only those observations with a relatively large index value. The columns of Table A.8 show results for subsets of observations in which the HHI is greater than the first, second, and third quartiles of the HHI distribution, respectively. Estimates are similar to our main results, suggesting a slightly larger capitalization of view, though statistical significance decreases with the reduction in sample size.

Conclusion

While a valuable view amenity creates potential for welfare gains through spatial allocation of open space that maximizes total view to the area, there are tradeoffs to such allocation. For example, while the perimeter of an open space parcel is important for the view amenity, total size or depth of open space may generate larger values linked to recreational opportunities or ecosystem services. Functional aspects, such as parking lots and facilities may increase recreational value but diminish quality of the view. Our results suggest, however, an important additional component of open space, namely the quality and quantity of households' view of such space, that should be considered in valuation. Finally, the potential complementarity of view amenities may be an important factor in nonmarket valuation when

[18]The Herfindahl-Hirschman Index is typically used to measure market concentration, but more generally measures the dispersion of a set of numbers whose sum is fixed.

considering other environmental goods.

In this analysis, we provide evidence for the amenity capitalization of a household's view. Using machine learning techniques, we develop a new approach to quantitatively measuring view. This allows us to identify the view amenity separately from any amenity value that works through proximity. This second channel, proximity to open space, has dominated the literature on valuation of open space. Instead, our results indicate significant effect of view and suggest that value is created through properties other than publicly designated parks. These results are robust to a number of different specifications. A causal effect is further supported by a repeat sales model. From a policy perspective, such identification could indicate the optimal size and location of open space for welfare maximization. In addition to illustrating an important component of the nonmarket valuation of open space, our approach to defining a new variable is widely applicable in hedonic analysis.

References

Abbott, J. K. and Klaiber, H. A. (2011), 'An embarrassment of riches: Confronting omitted variable bias and multiscale capitalization in hedonic price models', *The Review of Economics and Statistics* **93**(4), 1331–1342.

Albouy, D. (2018), 'Unlocking amenities: Estimating public-good complementarity', *Working Paper* .

Anderson, S. T. and West, S. E. (2006), 'Open space, residential property values, and spatial context', *Regional Science and Urban Economics* **36**, 773–789.

Bateman, I. J., Jones, A. P., Lovett, A. A., Lake, I. R. and Day, B. H. (2002), 'Applying geographic information systems (gis) to environmental and resource economics', *Environmental and Resource Economics* **22**(2202), 219–269.

Black, S. E. (1999), 'Do better schools matter? parental valuation of elementary education', *The Quarterly Journal of Economics* **114**(2), 577–599.

Fleming, D., Grimes, A., Lebreton, L., Maré, D. and Nunns, P. (2018), 'Valuing sunshine', *Regional Science and Urban Economics* **68**, 268–276.

Franco, S. F. and McDonald, J. L. (2018), 'Measurement and valuation of urban greenness: Remote sensing and hedonic applications to lisbon, portugal', *Regional Science and Urban Economics* **72**, 156–180.

Glaeser, E. L., Kominers, S. D., Luca, M. and Naik, N. (2015), 'Big data and big cities: The primises and limitations of improved measures of urban life', *NBER Working Paper* **21778**. NBER Working Paper.

Greenstone, M. and Gallagher, J. (2008), 'Does hazardous waste matter? evidence from the housing market and the superfund program', *The Quarterly Journal of Economics* **123**(3), 951–1003.

Guo, G., Humphreys, B. R., Nouyet, M. I. and Zhou, Y. (2019), 'Attractive or aggressive? a face recognition and machine learning approach for estimating returns to visual appearance', *Economics Faculty Working Papers Series* **36**.

Howard, D. H., Bach, P. B., Berndt, E. R. and Conti, R. M. (2015), 'Pricing in the market for anticancer drugs', *Journal of Economic Perspectives* **29**(1), 139–162.

Johnson, E. B., Tidwell, A. and Villupuram, S. V. (2020), 'Valuing curb appeal', *The Journal of Real Estate Finance and Economics* **60**, 111–133.

Klaiber, H. A. and Phaneuf, D. J. (2010), 'Valuing open space in a residential sorting model of the twin cities', *Journal of Environmental Economics and Management* **60**(2), 57–77.

Linden, L. and Rockoff, J. E. (2008), 'Estimates of the impact of crime risk on property vlues from megan's laws', *The American Economic Review* **98**(3), 1103–1127.

Lindenthal, T. and Johnson, E. B. (2021), 'Machine learning, architectural styles and property values', *The Journal of Real Estate Finance and Economics* .

Liu, X., Taylor, L., Hamilton, T. and Grigelis, P. (2013), 'Amenity values of proximity to national wildlife refuges: An analysis of urban residential property values', *Ecological Economics* **94**, 37–43.

Matthews, J. and Turnbull, G. (2007), 'Neighborhood street layout and property value: The interaction of accessibility and land use mix', *The Journal of Real Estate Finance and Economics* **35**(2), 111–141.

McConnell, V. and Walls, M. (2005), 'The value of open space: Evidence from studies of nonmarket benefits', *Resources for the Future Working Paper* .

McGarigal, K., Cushman, S. A., Neel, M. C. and Ene, E. (2002), 'FRAGSTATS: Spatial pattern analysis program for categorical maps', *University of Massachusetts, Amherst* .

Muehlenbachs, L., Spiller, E. and Timmins, C. (2015), 'The housing market impacts of shale gas development', *The American Economic Review* **105**(12), 3633–3659.

Naik, N., Kominers, S. D., Raskar, R., Glaeser, E. L. and Hidalgo, C. A. (2017), 'Computer vision uncovers predictors of physical urban change', *Proceedings of the National Academy of Sciences* **114**(29), 7571–7576.

Naik, N., Philipoom, J., Raskar, R. and Hidalgo, C. A. (2014), 'Streetscore – predicting the perceived safety of one million streetscapes', Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition Workshops, Washington, DC.

Neumann, B. C., Boyle, K. J. and Bell, K. P. (2001), 'Property price effects of a national wildlife refuge: Great meadows national wildlife refuge in massachusetts', *Land Use Policy* (26), 1011–1019.

Paterson, R. W. and Boyle, K. J. (2002), 'Out of sight, out of mind? using gis to incorporate visibility in hedonic property value models', *Land Economics* **78**(3), 417–425.

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

Shultz, S. D. and King, D. A. (2001), 'The use of census data for hedonic price estimates of open-space amenities and land use', *Journal of Real Estate Finance and Economics* **22**(2), 239–252.

Smith, V. and Huang, J.-C. (1995), 'Can markets value air quality? a meta-analysis of hedonic property value models', *Journal of Political Economy* **103**(1), 209–227.

Smith, V., Poulos, C. and Kim, H. (2002), 'Treating open space as an urban amenity', *Resource and Energy Economics* **24**, 107–129.

Tagliafierro, C., Boeri, M., Longo, A. and Hutchinson, W. G. (2016), 'Stated preference methods and landscape ecology indicators: An example of transdisciplinarity in landscape economic valuation', *Ecological Economics* **127**, 11–22.

Tu, C. C. and Eppli, M. J. (2003), 'Valuing new urbanism: The case of kentlands', *Real Estate Economics* **27**(3), 425–451.

Turner, T. and Seo, Y. (2021), 'House prices, open space, and household characteristics', *Journal of Real Estate Research* **43**(2), 204–225.

Walls, M., Kousky, C. and Chu, Z. (2013), 'Is what you see what you get? the value of natural landscape views', *Resources for the Future Discussion Paper* **RFF DP 13-25**.

Walsh, P., Griffiths, C., Guignet, D. and Klemick, H. (2017), 'Modeling the property price impact of water quality in 14 chesapeake bay counties', *Ecological Economics* **135**, 103–113.

Wan, W. X. and Lindenthal, T. (2023), 'Testing machine learning systems in real estate', *Real Estate Economics* **51**(3), 754–778.

Weicher, J. C. and Zerbst, R. H. (1973), 'The externalities of neighborhood parks: An empirical investigation', *Land Economics* **49**(1), 99–105.

Yang, J., Rong, H., Kang, Y., Zhang, F. and Chegut, A. (2021), 'The financial impact of street-level greenery on new york commercial buildings', *Landscape and Urban Planning* **214**.

Table 1: Summary Statistics

		Mean	Std. Deviation	0.05 Perc.	Median	0.95 Perc.
Housing	**Sale Price**	574530	1937142	119000	330000	905000
	# of Bedrooms	2.50	0.88	1	2	4
	# of Bathrooms	2.16	0.94	1	2	4
	Square Feet (1000s)	1.45	0.68	0.68	1.25	2.87
	Lot Area (1000s)	4.63	3.03	0.423	4.69	9.52
	# of Stories	1.44	0.60	1	1	3
	Year Built	1965	37.71	1898	1964	2015
	Year of Sale	2015	1.63	2011	2015	2017
Park Measures	**Park Proximity**	0.0283	0.1457	0.0009	0.0033	0.0386
	Park Adjacency	0.0201	0.1405	0	0	0
	Parkway Adjacency	0.0485	0.2149	0	0	0
View Scores	**park**	0.0884	0.1808	0	0.0101	0.5075
	residential	0.5533	0.3290	0.0138	0.5972	0.9787
	commercial	0.1288	0.2556	0	0.0047	0.8500

Table 2: In-Sample Prediction Accuracy (proportion of true class)

		Ground Truth			
		parks	residential	commercial	other
Prediction	parks	0.8200	0.0370	0.0000	0.0455
	residential	0.1667	0.9352	0.0000	0.0455
	commercial	0.0133	0.0093	0.9714	0.0000
	other	0.0000	0.0185	0.0286	0.9091

Table 3: Park Summary Conditional on Proximity

Proximity Quantile (upper bound)	Park Score Mean	Park Score S.D.
0.1	0.0994	0.2031
0.2	0.1041	0.2021
0.3	0.0934	0.1934
0.4	0.0909	0.1889
0.5	0.0788	0.1693
0.6	0.0775	0.1700
0.7	0.0957	0.1561
0.8	0.0733	0.1618
0.9	0.0991	0.1997
1.0	0.0657	0.1427

Table 4: Park Score Conditional Means

Distance-Based Measure	=1	=0
Proximity ($> 95^{th}$ percentile)	0.075	0.089
Park Adjacency	0.084	0.089
Parkway Adjacency	0.159	0.085

Table 5: Elements of Park Score Prediction

	$View\_Park - View\_Res$	$View\_Park - View\_Com$
Tree Area	0.0494***	0.2732***
	(8.10)	(65.05)
Building Area	-0.0366***	-0.7151***
	(3.30)	(-92.87)
N	179,667	179,667

Dependent variable is the difference in image classification scores. t-statistics, using standard errors clustered at the subdivision level, are shown in parentheses.
*** indicates statistical significance at the 1% level.

Table 6: Baseline Estimates

Dependent Variable = $ln(price_i)$

	(1)	(2)	(3)	(4)
$Prox_i$	0.0178			-0.0076
	(0.07)			(0.03)
$Prox_i^2$	0.0249			0.0499
	(0.11)			(0.19)
$Park\_Adj$		0.0073		0.0048
		(0.49)		(0.29)
$Parkway\_Adj$			-0.0893***	-0.0894***
			(-6.24)	(-6.25)
N	29,247	29,247	29,247	29,247
R^2	0.85	0.85	0.85	0.85

All specifications include fixed effects for subdivision, year built, and year of sale, structural housing attributes, and block group time trends. t-statistics, using standard errors clustered at the subdivision level, are shown in parentheses.
*** indicates statistical significance at the 1% level.

Table 7: Estimates of Park View Score

Dependent Variable = $ln(price_i)$

	(1)	(2)	(3)	(4)
$Prox_i$	-0.117			-0.0271
	(-0.48)			(-0.10)
$Prox_i^2$	0.1574			0.0717
	(0.66)			(0.27)
$Park\_Adj$		-0.0116		-0.0174
		(-0.75)		(-1.01)
$Parkway\_Adj$			-0.0999	-0.1009
			(-6.94)	(-7.01)
$View\_Park$	0.0376	0.0387	0.0434	0.0463
	(4.22)	(4.25)	(4.88)	(4.98)
N	29,247	29,247	29,247	29,247
R^2	0.85	0.85	0.85	0.85

All specifications include fixed effects for subdivision, year built, and year of sale, structural housing attributes, and block group time trends. t-statistics, using standard errors clustered at the subdivision level, are shown in parentheses.
*** indicates statistical significance at the 1% level.

Table 8: Estimates of Park View Score with Additional Spatial Controls

Dependent Variable = $ln(price_i)$

	(1)	(2)	(3)	(4)
$Prox_i$	-0.2004			0.0315
	(0.44)			(0.06)
$Prox_i^2$	0.2381			0.0157
	(0.54)			(0.03)
$Park\_Adj$		-0.0346		-0.0434
		(1.16)		(1.21)
$Parkway\_Adj$			-0.1018***	-0.104***
			(-3.56)	(-3.66)
$View\_Park$	0.0477***	0.0516***	0.0531***	0.0591***
	(2.87)	(2.97)	(3.16)	(3.40)
$View\_Res$	0.017**	0.0171**	0.0163**	0.0161*
	(1.73)	(1.74)	(1.68)	(1.64)
$View\_Com$	0.0331*	0.0332*	0.0348*	0.0346*
	(1.55)	(1.56)	(1.63)	(1.63)
Acr_Val	-0.0026*	-0.0026*	-0.0025*	-0.0025*
	(-1.53)	(-1.63)	(-1.56)	(-1.56)
N	28,989	28,989	28,989	28,989
R^2	0.85	0.85	0.85	0.85

All specifications include fixed effects for subdivision, year built, and year of sale, structural housing attributes, and block group time trends. t-statistics, using standard errors clustered at the subdivision level, are shown in parentheses.
***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Table 9: View as Discrete Variable: Park View Coefficient

Dependent Variable = $I(ln(price_i) > q)$		
quantile threshold	Coefficient	t-statistic
0.4	0.0101*	(1.63)
0.5	0.0139**	(1.88)
0.6	0.0136**	(1.72)
0.7	0.0176***	(2.26)
0.8	0.0105	(1.22)
0.9	0.0348***	(3.08)
max	0.0316***	(2.84)

Each regression defines a park view as having $View\_Park$ value above the specified quantile. All specifications include fixed effects for subdivision, year built, and year of sale, structural housing attributes, and block group time trends. Standard errors are clustered at the subdivision level.
***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Table 10: Repeat Sales Model

	Year Fixed Effects		Index-Deflated Prices	
	(1)	(2)	(3)	(4)
$View\_Park$	0.0485**	0.0429*	0.0483**	0.0749***
	(1.78)	(1.34)	(1.83)	(2.39)
$View\_Res$		-0.0669***		-0.0192
		(-2.40)		(-1.10)
$View\_Com$		0.0227		0.0683*
		(0.47)		(1.52)
N	660	660	660	660
R^2	0.78	0.79	0.13	0.13

All specifications include structural housing attributes. t-statistics, using standard errors clustered at the subdivision level, are shown in parentheses.
***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Appendix A Appendix Tables

Table A.1: Hedonic Function Estimates

Dependent Variable = $ln(price_i)$

	(1)	(2)	(3)	(4)	(5)
Intercept	11.71***	11.70***	11.70***	11.65***	11.65***
	(86.19)	(85.67)	(85.93)	(84.45)	(84.01)
$Prox$		-0.117			-0.0271
		(-0.25)			(-0.05)
$Prox^2$		0.1574			0.0717
		(0.35)			(0.14)
$Park\_Adj$			-0.0116		-0.0174
			(-0.38)		(-0.50)
$Parkway\_Adj$				-0.0999***	-0.1009***
				(-3.57)	(-3.63)
$View\_Park$		0.0376***	0.0387***	0.0434***	0.0463***
		(2.58)	(2.52)	(2.95)	(3.04)
$\#of Bedrooms$	0.0321***	0.0324***	0.0323***	0.0323***	0.0323***
	(5.79)	(5.82)	(5.84)	(5.85)	(5.83)
$\#of Bathrooms$	0.0902***	0.0901***	0.0902***	0.0899***	0.0899***
	(14.26)	(14.21)	(14.20)	(14.15)	(14.14)
$SquareFeet$	0.224***	0.220***	0.220***	0.220***	0.00022***
	(11.20)	(11.15)	(11.15)	(11.20)	(11.20)
$LotArea$	0.031***	0.031***	0.031***	0.031***	0.031***
	(7.75)	(7.75)	(7.75)	(7.75)	(7.75)
$StoriesNbr$	-0.0264**	-0.0265**	-0.0265**	-0.0266**	-0.0267**
	(-1.83)	(-1.84)	(-1.84)	(-1.85)	(-1.86)
N	29,247	29,247	29,247	29,247	29,247
R^2	0.85	0.85	0.85	0.85	0.85

All specifications include fixed effects for subdivision, year built, and year of sale, in addition to block group time trends. t-statistics, using standard errors clustered at the subdivision level, are shown in parentheses.
***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Table A.2: Weeks Between Sale and Image Capture

SaleYear	Min	Mean	SD	Max
2008	1.0	21.9	16.8	51.6
2009	3.1	35.8	13.4	51.7
2010	26.4	43.9	4.8	51.9
2011	0	15.8	11.6	51.6
2012	0	30.1	13.4	51.9
2013	13.1	38.6	9.0	51.9
2014	0	19.4	13.9	51.9
2015	0	28.9	15.9	51.9
2016	0	20.8	14.1	51.9
2017	9.0	34.9	10.1	51.9

Summary statistics, by year-of-sale, for the number of weeks between date of image and housing sale.

Table A.3: Picture-to-Sale Time Variation: Cross-Sectional Model

	26-week window	78-week window	78 weeks before
$Prox_i$	0.0929	-0.1379	-0.1026
	(0.16)	(-0.31)	(0.15)
$Prox_i^2$	-0.0327	0.1794	0.1659
	(-0.06)	(0.42)	(0.25)
$Park\_Adj$	-0.0759**	0.0094	-0.0008
	(-1.75)	(0.32)	(-0.02)
$Parkway\_Adj$	-0.0773***	-0.0988***	-0.0963***
	(-2.12))	(-3.71)	(-3.26)
$View\_Park$	0.0243*	0.042***	0.032**
	(1.51)	(2.88)	(1.79)
N	14,530	43,454	20,183
R^2	0.90	0.81	0.88

This table replicates baseline estimates with a stricter timing restriction between sale and time of picture. All specifications include fixed effects for subdivision, year built, and year of sale, structural housing attributes, and block group time trends. t-statistics, using standard errors clustered at the subdivision level, are shown in parentheses.
***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Table A.4: Repeat Sales Model

	Year Fixed Effects		Index-Deflated Prices	
	(1)	(2)	(3)	(4)
$View\_Park$	0.0959**	0.0525	0.1069**	0.1121**
	(1.97)	(0.87)	(2.26)	(2.02)
$View\_Res$		-0.081*		0.0221
		(-1.45)		(0.76)
$View\_Com$		-0.0759		0.015
		(-0.79)		(0.23)
N	268	268	268	268
R^2	0.78	0.78	0.06	0.06

This table replicates estimates from the repeated cross-section with a stricter timing restriction between sale and time of picture. t-statistics, using standard errors clustered at the subdivision level, are shown in parentheses.
***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Table A.5: Falsification Test: Park Score Coefficient

Dependent Variable = $ln(price_i)$

Radius	Coefficient	t-statistics
50-100	0.0133	(0.67)
50-150	-0.0167	(-0.68)
100-150	-0.0186	(-0.79)
100-200	-0.0125	(-0.66)

The falsification tests assigns to a home the $Park\_View$ score of a random home within the specified radius.
All specifications include fixed effects for subdivision, year built, and year of sale, structural housing attributes, and block group time trends. Standard errors are clustered at the subdivision level.

Table A.6: Inclusion of Own Park View Score

Dependent Variable = $ln(price_i)$

	(1)	(2)	(3)	(4)
$Prox_i$	-0.0848			-0.0026
	(-0.18)			(-0.01)
$Prox_i^2$	0.1263			0.0479
	(0.28)			(0.10)
$Park\_Adj$		-0.0098		-0.0159
		(-0.32)		(-0.45)
$Parkway\_Adj$			-0.098	-0.0988
			(-3.54)	(-3.58)
$View\_Park$	0.0357***	0.0367***	0.0416***	0.0441***
	(2.43)	(2.37)	(2.81)	(2.86)
$View\_Own$	0.0071	0.0071	0.0102	0.0098
	(0.40)	(0.40)	(0.58)	(0.56)
N	29,247	29,247	29,247	29,247
R^2	0.85	0.85	0.85	0.85

All specifications include fixed effects for subdivision, year built, and year of sale, structural housing attributes, and block group time trends. t-statistics, using standard errors clustered at the subdivision level, are shown in parentheses.
***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Table A.7: Robustness Checks

Dependent Variable = $ln(price_i)$

	Post-2012	Month FEs	Month Interaction	1-Story Homes	Other Amenities	Open Space
$Prox$	0.0870	-0.0003	-0.0269	0.3722	-0.0020	-0.0116
	(0.15)	(0.01)	(-0.05)	(0.48)	(0.01)	(-0.02)
$Prox^2$	-0.0095	0.0523	0.0715	-0.2723	0.0465	0.0561
	(-0.02)	(0.11)	(0.14)	(-0.36)	(0.09)	(0.11)
$Park\_Adj$	-0.0162	-0.0200	-0.0174	-0.0195	-0.0179	-0.0157
	(-.45)	(-0.57)	(-0.50)	(-0.37)	(-0.51)	(-0.45)
$Parkway\_Adj$	-0.0982***	-0.1042***	-0.1009***	-0.1161***	-0.1018***	-0.0985***
	(-3.61)	(-3.79)	(-3.63)	(-4.36)	(-3.66)	(-3.58)
$View\_Park$	0.0266**	0.0444***	0.0451***	0.0331**	0.0455***	0.0444***
	(2.25)	(2.92)	(2.13)	(1.88)	(2.99)	(2.88)
N	27,290	29,247	29,247	17,872	29,247	29,247
R^2	0.86	0.85	0.85	0.88	0.85	0.85

Robust checks for transactions after 2012 (Column 1); inclusion of month fixed effects (Column 2); inclusion of a month-of-year interaction (Column 3); sample restricted to 1-story homes (Column 4); inclusion of other spatial amenities (Column 5); inclusion of undeveloped parcels (Column 6).
All specifications include fixed effects for subdivision, year built, and year of sale, structural housing attributes, and block group time trends. t-statistics, using standard errors clustered at the subdivision level, are shown in parentheses.
***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Table A.8: Sample Restrictions using Herfindahl-Hirschman index

Dependent Variable = $ln(price_i)$

	(HH > Q_1)	(HH > Q_2)	(HH > Q_3)
-2.95 -2.80 -2.30 2.55 2.22 1.54 $Prox$	0.3238	0.2947	0.8351
	(0.73)	(0.60)	(1.05)
$Prox^2$	-0.2871	-0.2375	-0.8551
	(-0.67)	(-0.49)	(-1.08)
$Park\_Adj$	-0.0353	-0.0438	-0.0245
	(-0.83)	(-0.92)	(-0.46)
$Parkway\_Adj$	-0.0958***	-0.0959***	-0.1053**
	(-2.95)	(-2.80)	(-2.30)
$View\_Park$	0.0441***	0.0516**	0.0511*
	(2.55)	(2.22)	(1.54)
N	21,935	14,624	7,313
R^2	0.85	0.85	0.88

Each column restricts the sample to homes with a HHI above the specified quartile. All specifications include fixed effects for subdivision, year built, and year of sale, structural housing attributes, and block group time trends. Standard errors, clustered at the subdivision level, are shown in parentheses.
***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

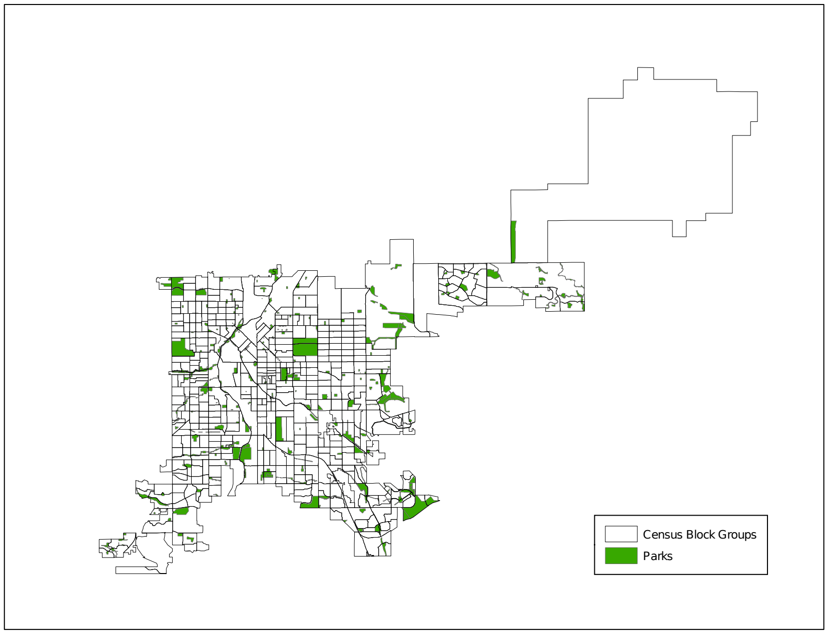
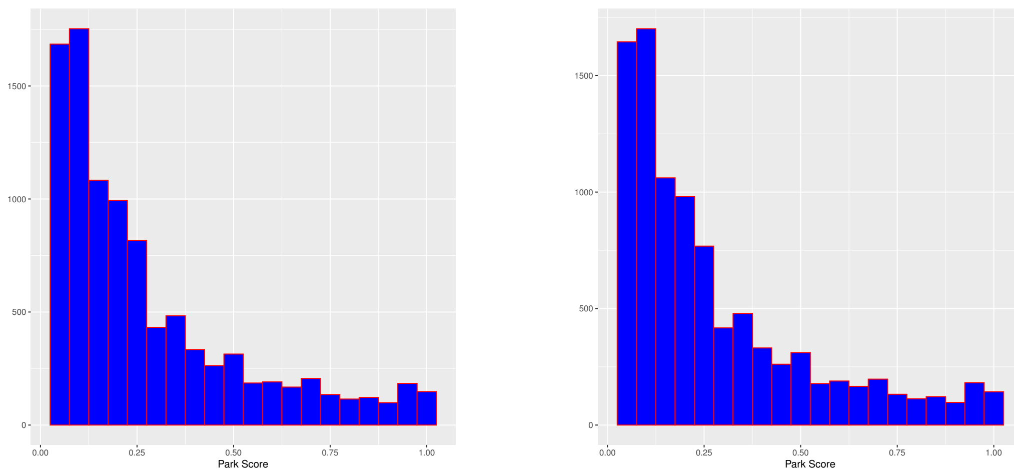


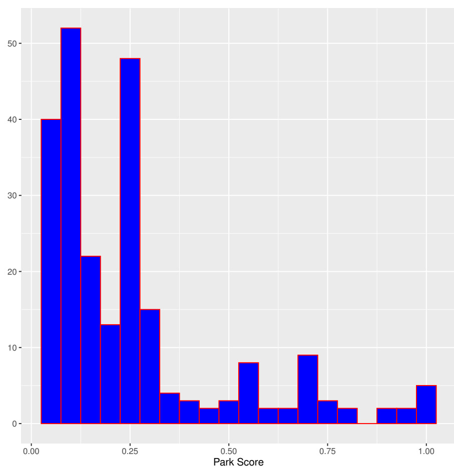
Figure 1: Study Area



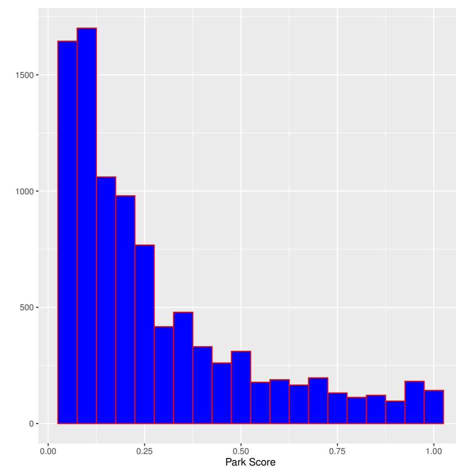
(a) Full Distribution

(b) park score > 0.05

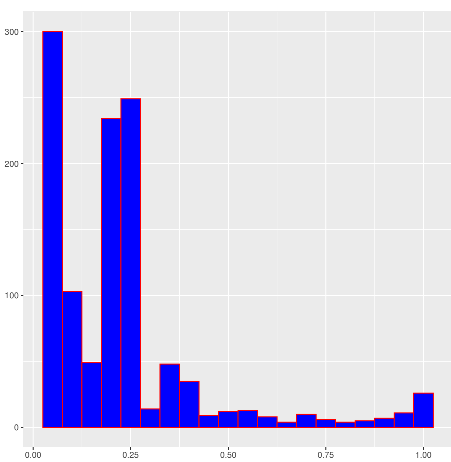
Figure 2: Distribution of $View\_Park$



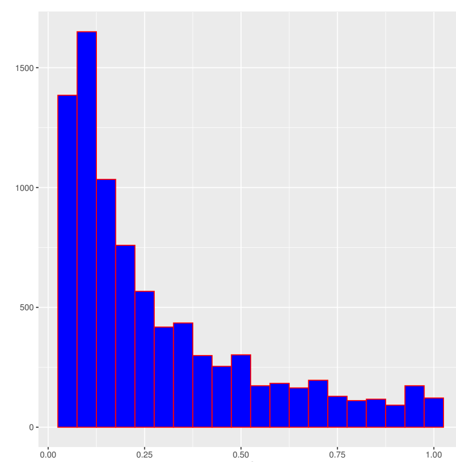
(a) $Park\_Adj = 1$



(b) $Parkway\_Adj = 1$

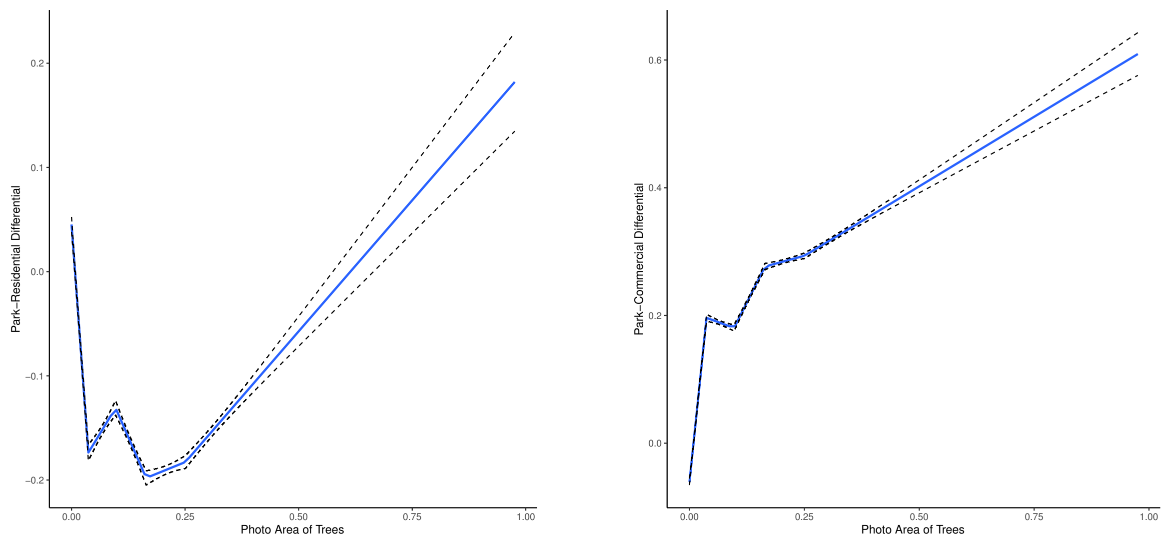


(c) $Park\_Adj = 0$



(d) $Parkway\_Adj = 0$

Figure 3: Distribution of Park Scores conditional on Park/Parkway Adjacency



(a) Park-Residential Differential

(b) Park-Residential Differential

Figure 4: Linear spline for predicted park score differential