Landscape and Urban Planning 104 (2012) 373–387

Contents lists available at SciVerse ScienceDirect

Landscape and Urban Planning

journal homepage: www.elsevier.com/locate/landurbplan

Estimating the value of urban green areas: A hedonic pricing analysis of the single family housing market in Los Angeles, CA

Jean-Daniel Saphores[a,b,c,*], Wei Li[c,1]

[a] Department of Civil and Environmental Engineering, University of California, Irvine, CA 92697, USA
[b] Department of Economics, University of California, Irvine, CA 92697, USA
[c] Department of Planning, Policy and Design, University of California, Irvine, CA 92697, USA

ARTICLE INFO

Article history:
Received 21 September 2010
Received in revised form
11 November 2011
Accepted 15 November 2011
Available online 15 December 2011

Keywords:
Hedonic pricing
Cliff–Ord
Geographically weighted regression
Single family housing
Urban forests
Urban land cover

ABSTRACT

We analyze 20,660 transactions of single family detached houses sold in 2003 and 2004 in the city of Los Angeles, CA, to estimate the value of urban trees, irrigated grass, and non-irrigated grass areas. To deal with spatial autocorrelation and unobserved neighborhood characteristics, we contrast two models: a geographically weighted regression model, and a Cliff–Ord model with spatial lags in the dependent variable, the exogenous variables, and the disturbances as well as submarket fixed effects and an extensive set of covariates. We find that Angelenos like lawns: over 88% of the properties examined would gain value with additional irrigated grass on their parcel, and even more (89%) in their neighborhood. Although more non-irrigated grass/bare soil on parcels typically hurts property values, it often has the opposite effect at the neighborhood level. Moreover, additional parcel trees would decrease the value of almost 40% of the properties examined and they would have only a small positive impact on most of the others. By contrast, additional neighborhood trees would slightly increase the value of over 97% of the properties analyzed. This suggests that while Los Angeles residents may want additional trees, they are unwilling to pay for them. These results have implications for urban tree planting programs that rely primarily on private property owners.

1. Introduction

The importance of urban green spaces, and especially urban forests, is increasing worldwide because of the expansion of urban land fueled by urbanization. According to the Population Reference Bureau (2010), half of the world's population is now living in urban areas. In the United States, which is 79% urbanized, the percentage of urban land may soar from 2.6% in 2002 (Lubowski, Vesterby, Bucholtz, Baez, & Roberts, 2006) to over 8% by 2050 (Nowak & Dwyer, 2007, chap. 2). At the same time, there is growing evidence of links between urban green spaces, health, and social safety (Groenewegen, van den Berg, de Vries, & Verheij, 2006; Tzoulas et al., 2007). As highlighted in Conway, Li, Wolch, Kahle, and Jerrett (2010), empirical research on the amenity value of neighborhood green spaces is still limited although a dynamic literature has been exploring for some time the value of open space and urban parks (e.g., see Brander & Koetse, 2011).

There has also been a keen interest for urban trees, which may be explained by their many potential benefits. These benefits include providing habitat to various species (insects, birds, small rodents); controlling erosion and limiting water runoff; improving air quality by intercepting particulate matter, ozone, and nitrogen dioxide; removing carbon dioxide from the atmosphere; providing shading, which decreases energy use and mitigates the urban heat island effect; and beautifying neighborhoods as well as enhancing some people's sense of spiritual well-being. Some of these benefits (habitat provision, CO_2 removal from the atmosphere, and runoff reduction) are public goods, while others (esthetic qualities, air quality improvements, erosion reduction, and shading) are more like private goods so they are likely to be capitalized in the housing market. However, still little appears to be known about the value of urban trees in a Mediterranean climate that characterizes Los Angeles and many developing megacities.

Our paper starts bridging these gaps and suggests some explanations for the slow rate of tree planting experienced by the large tree planting program started in September 2006 in Los Angeles, CA. Our dataset includes records of 20,660 single family detached homes sold in 2003–2004 in Los Angeles, CA, and high resolution (2 ft) land use data (McPherson, Simpson, Xiao, & Wu, 2007) for the parcel of these properties as well as a 200 m area surrounding each of them. To analyze the value of urban land cover, we estimate a Cliff–Ord

* Corresponding author at: Department of Civil and Environmental Engineering, University of California, Irvine, CA 92697, USA. Tel.: +1 949 824 7334.
E-mail addresses: saphores@uci.edu (J.-D. Saphores), wli3@uci.edu (W. Li).
[1] Tel.: +1 949 233 3437.

0169-2046/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.landurbplan.2011.11.012

hedonic model (Anselin, 1988) with submarket fixed effects and an extensive set of covariates to deal with spatial autocorrelation and unobserved neighborhood characteristics. Moreover, we estimate a geographically weighted regression model (Cleveland & Devlin, 1988; Fotheringham, Brunsdon, & Charlton, 2002) to confirm the robustness of our findings.

As suggested by Table 1, the HPM has become the approach of choice to study the value of urban land cover thanks to advances in remote sensing, geographic information systems (GIS), and econometrics (including fixed effects and spatial econometrics models). It has been widely applied to study different environmental externalities (Sirmans, MacDonald, Macpherson, & Zietz, 2006) and it is particularly well suited here since local land cover is readily observable.

Apart from the HPM, different approaches have been used for estimating the value of urban land cover. Early studies (Morales, Boyce, & Favretti, 1976; Payne, 1973) focused on urban trees and analyzed hand-picked datasets. During the 1980s and 1990s (see Kestens, Thériault, & Des Rosiers, 2004 or Sander, Polasky, & Haight, 2010), researchers broadened their inquiries to urban green spaces and explored a variety of techniques. A number of papers (see Brander & Koetse, 2011, for references) relied on the contingent valuation method (CVM), which asks people for their willingness to pay for changes in environmental quality under various hypothetical scenarios (Carson, Flores, & Meade, 2001). However, the CVM relies on stated preferences that may not translate into actual behavior. An alternative is the travel cost method, but with the exception of Dwyer, Peterson, and Darragh (1983), it has not been used for valuing urban green spaces because people typically do not travel specifically to enjoy them.

Understanding how the value of land cover is capitalized in the real estate market is important not only to real estate developers who could profit from building more desirable residential communities, but also to planners and local officials, so they can foster the adequate provision of the local public goods provided by urban green spaces by designing better zoning and land-use regulations. This is especially salient since a number of US cities have recently committed to large tree planting programs, including Baltimore, Denver, Houston, and New York, to name a few. One of the most ambitious tree planting programs, however, started in September 2006 in Los Angeles with the stated goal of planting one million trees by the end of 2010 (McPherson et al., 2007); it provides one of the motivations of this study.

2. Methods

2.1. Theoretical considerations

Following the standard hedonic framework (Rosen, 1974), we explain the market price P_h of a single family detached house based on its structural ($\mathbf{S_h}$) and neighborhood ($\mathbf{N_h}$) characteristics, as well as environmental variables ($\mathbf{G_h}$):

$$P_h = f(\mathbf{S_h}, \mathbf{N_h}, \mathrm{G_h}, e_h). \quad (1)$$

In Eq. (1), variables in bold are vectors, and the error term, e_h, reflects uncertainty in the measurement of variables and in the preferences of individual homebuyers. The partial derivative of f with respect to one of its arguments is an implicit price; it represents a consumer's marginal willingness to pay for the corresponding characteristic.

Rosen's (1974) framework requires strong assumptions to be valid, however: first, the market considered should be in equilibrium; second, it should be perfectly competitive; third, buyers and sellers need to have perfect information about product characteristics; and fourth, there should be a continuum of products.

MacLennan (1977) argued that equilibrium may be assumed if the housing market does not suffer severe shocks and if the study period is reasonably short, which is the case here. In addition, Meese and Wallace (1997) found that housing markets typically adjust quickly to small shocks so the equilibrium assumption is reasonable. Assuming a continuity of products is sensible in Los Angeles' large housing market but perfect competition and perfect information are more difficult to justify. Fortunately, Bajari and Benkard (2005) showed that the demand side of the market guarantees the existence of a function relating price and product characteristics even under imperfect competition and even if the number of products considered is small. For a more in-depth discussion of theoretical issues, see Taylor (2008).

2.2. Empirical considerations

We hypothesize that the value of different types of land cover may depend on lot size, population density, income levels, school quality, crime rates, neighborhood age composition, as well as ethnic make-up. For example, the presence of neighborhood trees or irrigated grass is probably more valuable in a dense area, while having a larger lot may increase the effectiveness of parcel trees against various urban externalities, including noise. Our model also accounts for the distance to components of the urban "green infrastructure" (parks, golf courses, lakes, rivers, and cemeteries) as previous research found that they impact housing values (e.g., Anderson & West, 2006). Since Los Angeles is a polycentric city, we do not include measures of distance to a central business district.

The best data collection efforts cannot escape the threat of omitted variable bias, however. Omitted (or unobservable) variables may be spatially correlated and create spatial autocorrelation in the error terms of hedonic models; examples include local climate or neighborhood quality (Case, 1991). It is also well known that omitting variables leads to biased and inconsistent estimators if the omitted variables are correlated with the included explanatory variables (Kennedy, 2003). A second difficulty is that economic theory tells us little about the form of the hedonic function represented by Eq. (1). A third difficulty is to correctly account for spatial dependence in the data, because ignoring it may lead to biased estimators and misleading inferences (Anselin & Arribas-Bel, 2011). A battery of tests on the residuals of the ordinary least squares model obtained by setting λ, ρ, and $\mathbf{b_L}$ to zero in Eq. (2) (no spatial dependence of any kind) for a block group contiguity weight matrix clearly indicates the presence of spatial dependence: Moran's I is 0.284 (p-value < 0.001), the Lagrange Multiplier lag and error test statistics are 10,662 and 50,884 respectively, both with tiny p-values; moreover, their robust versions are 6686 and 1165 respectively, with p-values < 1E−6.

To overcome these difficulties, we contrast two models: first, we estimate a general Cliff–Ord spatial hedonic model (Anselin, 1988; Cliff & Ord, 1981) with lags in the dependent variable, the exogenous variables, and the disturbances as well as submarket fixed effects and an extensive set of explanatory variables; and second, a geographically weighted regression model. As shown by Anselin and Arribas-Bel (2011), spatial hedonic models are superior to spatial fixed effects models because the latter correctly remove spatial correlation only in special cases. In addition, we pay special attention to heteroskedasticity because it can cause maximum likelihood estimators to be inconsistent in spatial models (Arraiz, Drukker, Kelejian, & Prucha, 2010).

To explore an adequate functional form for our models, we first inspected graphically the relationship between price and key explanatory variables in our dataset. This exploration suggested that a log–log functional form is appropriate for continuous structural and location variables. Moreover, a Box–Cox transformation of our dependent variable yielded an estimated power coefficient