Dynamic Analysis Of Open Space Value Using A Repeat Sales/Hedonic Approach.

W. Bowman Cutter[1],Linda Fernandez[2], Ritu Sharma[2],
and Tom Scott[3]

Preliminary Draft: Please Do Note Cite.

Abstract

This research employs a hedonic/repeat-sales method to value proximity to open space for residential values. Riverside County maintained an active program of open space preservations and acquisition along the wild land-urban frontier from 1988 through the end of our sample period in 2004 in order to preserve habitat and species. These new open space reserves allow us to test whether preserving nearby open space adds to the value of a residence. We use a repeat sales approach that measures whether the rate of house price appreciation is greater in a time period where the proximity to open space declines for a house. In addition, we adopt a matching/regression approach from the treatment literature to check the robustness of our results. Our research suggests that there are significant benefits to residential house values from converting open space from temporary, adjustable, uses such as agriculture to permanent preserves.

[1] Corresponding author, Department of Economics, Pomona College, bowman.cutter@pomona.edu
[2] Department of Environmental Sciences, University of California, Riverside, CA, USA
[3] Department of Environmental Policy, Science and Management, University of California, Berkeley

1. Introduction

A significant body of literature examines the benefit of open space on residential property value. However, there are relatively few papers that account for how temporal as well as special changes in open space may influence nearby residential values. This is a key question as open space may be changing over time due to acquisitions made by public resource managers. A dynamic accounting of open space value builds on the literature of how urban sprawl affects open space (e.g. Irwin and Bockstael (2001), Bluffstone et al. 2008).

Economic valuation of open space effects on residential property values in the literature usually relies on a hedonic-pricing approach. McConnell and Walls (2005) review this body of literature. McConnell and Walls (2005) conclude that the literature is mainly uses cross-sectional data and hedonic valuation to value open space amenities. This standard approach precludes the dynamic perspective that would be required to address the change in open space values over time as well as space. Our study is aimed at addressing the dynamic perspective that is missing from the literature.

While proximity to open space generally increases residential property value, open space is not generic and it should matter if the open space is a preserve in perpetuity versus simply temporarily raw land. Smith et al (2002) analyze open space that is fixed in use (golf course, public parks) and adjustable in use (agricultural use now or vacant land). In that study, it is found that the location of adjustable open space is determined by market forces and will be sensitive to buyer expectations and endogeneity of land uses. Alternatively, some have found that open space is exempt from market forces and thus, will be exogenous to housing price (Bockstael (2001) and Walsh (2007). Bockstael (2001) and Walsh (2007) both cite ways in which the government intervention into land management is regulated without a market. In their examples relating to forest land and wetlands, the supply side rather than the demand side is addressed.

Open space value is an example of amenities being capitalized into housing prices. Capitalization occurs when a change in taxes or public goods and services causes a change in house prices (Brasington, 2001). Hoyt (1999) studies cases where all may be equal between neighboring communities, a change in public goods (such as open space) can change the price of housing.Using a Tiebout model, Hoyt (1999) implies residents can costlessly move among all cities and have identical tastes and income in order to focus on the change in public goods' effect.

The finance literature offers approaches to real estate capitalization in the form of options from Merton (1973). The interpretation of the fair price of the real estate option is an equilibrium price each time the residential lot sells. The option value approach is a useful framework in that the open space amenity value is likely to evolve through time. Thus, the buyer has to think about how the value will change over time. An option implies a right to purchase a good at a pre-specified price (the real estate market sale price) and the exercise price. It has value if the market price exceeds the exercise price, as one would expect with appreciating real estate.

Aside from the spatial dimension of the previous studies, a few studies have checked for variation in real estate values over time and space. Geoghegan et al. (1997) validate the classic Von Thunen model showing distance from the commercial center decreases the value of property. The research on the value of proximity to open space typically uses cross-section or repeated cross section data where open-space areas are fixed and constant over the time period of the sample. This type of analysis could result in biased estimates of value because open space proximity could be correlated with unobservables that influence house values.

In this paper we have developed data on the conversion of adjustable to permanent open space over a 16-year period. We use this data to investigate whether designation of open-space preserves in perpetuity is capitalized into housing price values. This conversion value is a somewhat different value than that measured by prior cross section based literature. Open space vales estimated through cross section analysis approximate the value of a marginal change in open-space proximity. In our case, because the designated open space parcels were already open space, our empirical analysis is attempting to measure the difference between the value of open space that may be converted to another use and permanent open space.[4] Depending on the situation, either value could be more policy applicable. The conversion value measure applies in particular to open space preservation on the wildland-urban frontier. This frontier is often where ecosystem service values of preservation are high because contiguous habitats are necessary for biodiversity preservation.

We adopt the hybrid repeated sale/ hedonic price econometric approach of Case et al (2006). They use this methodology to analyze the impact of environmental contamination on condominium prices with dynamic data on the negative externality as well as repeated sales data on condominiums. We use repeated sales data from Western Riverside County (in Southern California). The county maintained an active program of open space acquisition from 1988 through the end of our sample period in 2004. The addition of new open space reserves allows us to test whether preserving nearby open space adds to the rate of appreciation of a residence.[5] The repeat sales methodology allows us to control for all time-invariant house characteristics, whether observed or not. To the best of our knowledge, none of the literature has used a dataset of multiple residential parcels sales with explicit dynamic spatial measures of open space to estimate open space value.

The addition of the open space reserves can be viewed as an experiment with a treatment group where proximity to permanent open space changes and a control group where the distance does not change. As a robustness check, we also use a matching methodology to test whether the control and treatment groups are similar. We employ a doubly-robust variety of propensity-score based matching

[4] The preserved land was typically zoned agricultural. The preservation of this land could generate amenity value either because natural habitat is generally preferred or because homebuyers fear that land will be converted from agricultural to another use.
[5] In our study area, other types of permanent open space such as city parks have not changed in quantity over time, unlike the preserved open space that is the focus of this study.

and regression techniques to compare control and treatment groups. Our results appear robust to a number of different categorization of open-space change.

A factor in our analysis, which may be common to other settings, is that the land that is preserved in perpetuity typically was not zoned residential. Therefore, preservation decisions do not impact housing supply directly. It is a second best framework where zoning of land for residential development is already set and there is no change in zoning from wild or agricultural land to residential development. The land zoned for permanent open space by the resource management agency (Riverside County Board of Supervisors) was open but its legal status changed so that its use cannot be changed. The Board of Supervisors initiated these changes through the Riverside County Integrated Plan (RCIP) initiated in 1998. The RCIP Vision is to "afford the human experience with natural environment and sustain the permanent viability of ecosystems. At the time of the plan launch in 2002, the goal was set a goal of 500,000 acres of open space acquisitions to set aside over time and 43,000 has been set aside since the plan.

Riverside's drastic change in the open space planning gives us an opportunity to examine the role of expectations about open space use in determining its price. The planning first gave broader exposure to the issue of open space disappearance in western Riverside, and then announced a sweeping plan to preserve open space. This information could well be taken as a signal of the likely future availability of open space and affect its marginal value.

The paper begins with an option value model of permanent open space. Then we discuss the policy process that lead to open space preservation. Next, we present the repeat sales/hedonic and matching econometric approaches. Then we discuss the data we use for both the repeat sales and propensity score matching. Finally we discuss our results and conclude.

2. Model

The market value of a residential property is defined as the price prospective residential property buyers are willing to pay for the property under prevailing economic conditions. The rate of appreciation or change in market value per unit of time may be estimated as a combination of observed changes in the sale prices of homes of a similar type over a particular time period. On the demand side of residential real estate, there are N buyers with the following additive utility in residential housing consumption and environmental quality from a public good such as preserved open space:

$$(1)\ u^n = u(g^n, h^n, \theta) = g^n + H(h)^n + \frac{W}{d^n}\theta$$

where g^n is composite good consumption, h^n is residential housing consumption, W is the marginal utility derived by the individual from environmental quality from the public good of preserved open space, d^n is the distance the parcel is from open space, and θ is the volatility coefficient associated with open space preservation value that might change over time because parcel acquisition is not deterministic. H is assumed to be twice differentiable and concave, $H'(\bullet) > 0, H''(\bullet) < 0$. Since the open space is preserved in perpetuity, distance measures can be expected to describe how that open space affects a nearby residential property's value. Since all potential bidders for each site can be expected to bid for a location with the same knowledge, their marginal values reflect proximity to the same open space (Smith et al. 2002).

A buyer of residential real estate faces the following budget constraint:

$$y^n = g^n + p_H \frac{h^n}{d^n} \qquad (1)$$

where y^n is income, and p_H is the price of a residential housing parcel. The price of the composite good is normalized to 1. We can then substitute for g^n by $y^n - p_H \frac{h^n}{d^n}$ in the utility function of the buyer yielding:

$$u^n = u(y^n, h^n, \theta) = y^n - p_H \frac{h^n}{d^n} + H(h)^n + \frac{W}{d^n}\theta \quad (2)$$

On the supply side, the budget constraint of those selling residential real estate is:

$$p_H \frac{h^n}{d^n} + y^n = g^n \quad (3)$$

The sale of the residential property is a source of income for the seller.

The utility function of the sellers of residential real estate is:

$$u^n = u(y^n, h^n, \theta) = y^n + p_H \frac{h^n}{d^n} + H(h)^n + \frac{W}{d^n}\theta \quad (4)$$

The interaction of the demand and supply will result in an equilibrium price of residential real estate. It will be possible to see the relationship of open space value and distance to open space effects on price. Since supply and demand may transact over time more than once during which distance to preserved open space can change, we expect to see dynamic value changes.

In our paper, econometric analysis helps quantify the specific measure of change dynamically, due to the open space amenity. Open space is designated in perpetuity and this may be empirically estimated as a capitalization effect of locating near permanent amenity value. A perpetuity as a financial earning and not a time designation is defined as an annuity that continues indefinitely. This may imply a deterministic rather than stochastic perception of land use and value. With the potential change in the distance the residential housing is to the open space amenity over time as well as the resale of residential property over time, one would not expect to earn a fixed annuity indefinitely as the value would be changing over time in this case. Our empirical estimation will be able to test for the dynamic change.

In the dynamic setting, the option framework is useful. The option on a residential parcel near open space whose value p_H, the price process, has a boundary condition

$$p_{H0} = p_H . \qquad (5)$$

Let r be the rate of interest. A pseudo price process A according to Black and Scholes (1973) for an option is evaluated according to the equivalent martingale measure, whose existence and uniqueness in a complete real estate market is related to absence of arbitrage by the fundamental theorems of asset pricing. So, the price process A that will help in the analysis is:

$$A_0 = p_H . \qquad (6)$$

Replicating an option guarantees the random value of the option at any time, with probability one. Which price of the repeatedly sold residential real estate internalizes the value of the open space amenity captured in an option? The possible equilibrium price p_H of the option depends on the initial value of the real estate p_{H0}, on the number of years before expiration ∞, and on the exercise price S.

Option pricing of a residential real estate transaction focuses on the price of a unique equilibrium as the fair price. The value of the option in time t is $v(p_{Ht}, S)$ and the discounted value is $e^{-rt}\ v(p_{Ht}, S)$. An upper bound, U, is the largest value of the residential parcel. The strategy for the seller is a put [0, U] and can be indexed by $U \geq 0$. The stopping time is defined as τ_U, of first entry in [0, U] for put options that is tied to the dynamic evolution of real estate.

Assume buyers and sellers in the residential real estate market have common knowledge of the interest rate r and the volatility coefficient associated with the value of open space preservation that might change over time θ through some information such as the County of Riverside announcing the RCIP involving open space preservation. The price of the underlying residential real estate is p_{H0}. At this price both seller and buyer are indifferent between the amount of cash p_{H0} and any martingale J with expected value p_{H0}. The discounted pseudo-price process $J_t := e^{-rt} A_t$ that they are indifferent to p_{H0}.

The solutions of (5) and (6) are given by the formulas

$$A_t = \exp\{\theta\} p_{H0} \tag{7}$$
$$p_{Ht} = \exp\{(r)t\} A_t = \exp\{\theta\} p_{H0}. \tag{8}$$

Equations (7) and (8) have $\exp\theta$ based on information such as the RCIP to formulate an expected value that includes the open space amenity as a function of the RCIP acquisition process variance.

Hence the discounted price process is

$$e^{-rt} p_{Ht} = e^{rt} J_t \tag{9}$$

The buyer and seller could have different opinions about the fluctuation in residential real estate value. Given that the open space is designated in perpetuity by the public resource manager, its designation can be common knowledge as long as a potential buyer has heard about it from a realtor or public announcement. This can translate into reducing the uncertainty in the real estate price process as well as increasing the mean value of the distance to open space that the buyer and seller are interested in.

The seller (I) and buyer (II) preferences can be represented according to the theory of Von Neumann Morgenstern expectations of the utility functions

$$u_I : (P, (\tau_U, \overline{P})) \to P - e^{-r\tau_U} v(A_{\tau_U}, S) \tag{10}$$
$$u_{II} : (P, (\tau_U, \overline{P})) \to e^{-r\tau_U} v(A_{\tau_U}, S) - P \tag{11}$$

if P$\leq \overline{P}$ where $\overline{P}$ is the maximum price the buyer will pay, and $u_I = u_{II}$=0 otherwise. Note, τ_U is the first entrance time of the pseudo price process A in [0, U], for a put.

The fair price of a perpetual American option with strike price S is

$$P^* = \sup_{U \in \Re_+} \mathbf{E}\left[\exp\{-r\tau_U\} v(A_{\tau_U}, S)\right], \tag{12}$$

Where $\mathbf{E}$ is the expectation operator.

The price P^* is an equilibrium price for the transaction described above if there exists an optimal real estate sales policy, namely, if there exists U^* such that

$P^* = \mathbf{E}[\exp\{-r\tau_{U^*}\} v(A_{\tau_{U^*}}, S)]$, then

$$(P^*, (\tau_{U^*}, P^*)) \tag{13}$$

is a Nash equilibrium or price clearing of the transaction between buyer and seller.

The subsequent sections on the empirical estimation will focus on measuring how the price in repeated residential sales capitalizes open space that is preserved dynamically.

3. Background

Several developments have lead to the designation and/or acquisition of additional open-space habitat in western Riverside County. It will be evident from the following description that decisionmaking is largely driven by exogenous biological factors. The initial push for open-space designation was due to the U.S. Fish and Wildlife Service 's (USFWS) decision to list Stephens' kangaroo rat (SKR) as an endangered species under the Endangered Species Act (ESA) in October 1988. To protect the SKR and its habitat from any type of disturbance there was a freeze on new development on more than 22,000 acres throughout western Riverside County. In order to address the perceived severe economic impacts of the SKR listing, the Riverside County Habitat Conservation Agency (RCHCA) was formed in 1990 for the purpose of planning, acquiring, and managing habitat for the SKR and other endangered, threatened, and candidate species. The RCHCA is a Joint Powers Agreement agency comprised of the Cities of Corona, Hemet, Lake Elsinore, Moreno Valley, Murrieta, Perris, Riverside, Temecula, and the County of Riverside. A Short-Term Habitat

Conservation Plan (HCP), approved by the USFWS and CDFG, was prepared by RCHCA in August 1990 as a conservation program designed to afford protection to the SKR while a plan providing for the establishment of permanent preserves could be developed.

Stakeholders and interest groups became concerned that habitat acquisition decisions were insufficiently targeted towards maintaining entire ecosystems and meeting other public needs. In response, on October 20, 1998, the County Board of Supervisors reviewed a set of consensus 'planning principles' submitted by a coalition of interest groups and endorsed their use as initial guidelines in the early stages of developing Riverside County Integrated Project (RCIP). It is a comprehensive, three-part, integrated program, initiated by the Riverside County Board of Supervisors on May 1999 and the draft released for public review in April 2002 (RCIP, 2003). The 3 parts of the RCIP program include: protecting the natural environment by conserving habitat and open space through a Multi-Species Habitat Conservation Plan (MSHCP), reducing traffic congestion by addressing future traffic and circulation issues through the Community & Environmental Transportation Acceptability Process (CETAP) and balancing land-use in the County by determining where our future housing, schools and businesses will be located using the updated County's 'General Plan'. In addition, a Special Area Management Plan (SAMP) planning process addresses watershed management and water-quality issues in the region.

The MSHCP aims to conserve covered species and their habitats in the MSHCP plan area, improve the future economic development in the County by providing a streamlined regulatory process through which development can proceed in an efficient way and provide permanent open space, community edges, and recreational opportunities. This Plan will result in an MSHCP Conservation Area in excess of 500,000 acres and focuses on Conservation of 146 species. The MSHCP Conservation Area includes approximately 347,000 acres on existing Public/Quasi-Public Lands and approximately 153,000 acres of Additional Reserve Land. The public interest in the multi-habitat plans resulted in the development of yearly open-space habitat data from before the start of the KSR preservation plans

One of the important pieces of MSHCP is that it changes the scale both of open-space acquisitions and the public expectation of the scale of acquisitions. In response to the adoption of the plan the public may have increased their expectations of the amount of open-space preservation, and that could lower the marginal value of being close to open space. In addition, the types of areas to be preserved may have changed with the adoption of the MSHCP. This reasoning implies that the value of open-space proximity may differ depending on the period of sale, we test that hypothesis in our empirical specifications.

4. Empirical Model

We derive our repeat sales price-ratio equation using a hybrid of hedonics and the repeat sales model similar to that used by Case et al (2006). Typically, repeat-sale analysis is based on the assumption that the attributes and the parameters are constant through time. Suppose that houses prices follow the equation:

$$P_i = \gamma e^{\beta_1 Y_i + \tau_1 T_{i1} + \tau_2 T_{i2} + \tau_3 T_{i3} + \ldots + \tau_n T_{in}} \qquad (14)$$

where in equation 1, Pi is the price of property i, Y_i is a vector of property attributes that may change through time and $T_{i\Phi}$ is a dummy time variable such that

$$T_{i\Phi} = \begin{cases} 1, & if\ \Phi = t_i \\ 0, & if\ \Phi \neq t_i \end{cases}$$

and ti is the year of sale of the ith property. The year zero time dummy variable is omitted from the equation.

If a property sells twice, once at year t and once at an earlier year $\tilde{t}$ where the ~ denotes earlier year. The ratio of the two predicted prices would then be:

$$\frac{P_i}{\tilde{P}_i} = \frac{\gamma e^{\beta_1 Y_i + \tau_1 T_{i1} + \tau_2 T_{i2} + \tau_3 T_{i3} + \ldots + \tau_n T_{in}}}{\gamma e^{\beta_1 \tilde{Y}_i + \tau_1 \tilde{T}_{i1} + \tau_2 \tilde{T}_{i2} + \tau_3 \tilde{T}_{i3} + \ldots + \tau_n \tilde{T}_{in}}}$$

$$= e^{\tau_1 (T_{i1} - \tilde{T}_{i1}) + \tau_2 (T_{i2} - \tilde{T}_{i2}) + \tau_3 (T_{i3} - \tilde{T}_{i3}) + \ldots\ldots + \tau_n (T_{in} - \tilde{T}_{in})}$$

If we assume the property attributes and coefficients are constant, then the logarithmic transformation of this equation is:

$$\ln \frac{P_i}{\tilde{P}_i} = \tau_1 (T_{i1} - \tilde{T}_{i1}) + \tau_2 (T_{i2} - \tilde{T}_{i2}) + \tau_3 (T_{i3} - \tilde{T}_{i3}) + \ldots\ldots + \tau_n (T_{in} - \tilde{T}_{in}) \qquad \text{(Model 1)}$$

Model one is the well-known repeat-sale analysis equation in which the dependent variable is the ratio of prices, the attributes and the implicit prices of the attributes do not change over time, and the time variables in parentheses take on the value -1 if the first sale occurs during that period, 1 if the second sale occurs during that period and 0 if no sale occurs during that period. The equation is estimated by taking the natural logarithm of both sides and using ordinary least squares regression. In addition, in all the specifications in this paper we use a robust variance estimate with clustering at the property level. This controls for the heteroskedasticity at the property level that is noted by Case et al. (2006).

The price effect on the distance from the parcel to the preserve is analyzed by incorporating additional distance change variables to the above model. If we let X_i be the distance to open space for parcel i then we have the formulation:

$$\ln \frac{P_i}{\tilde{P}_i} = \tau_1 (T_{i1} - \tilde{T}_{i1}) + \tau_2 (T_{i2} - \tilde{T}_{i2}) + \tau_3 (T_{i3} - \tilde{T}_{i3}) + \ldots\ldots + \tau_n (T_{in} - \tilde{T}_{in}) - \beta_1 \mathit{distchange} \qquad \text{(Model 2a)}$$

where *distchange* is the decrease in distance from the first to the second sale.

As outlined in the theory section, it is possible that open space values are only slowly incorporated into housing value. In order to account for this time lag we include a variable measuring the years between the sale date and the establishment of the preserve (*yrsbfsale*) and an interaction term between *distchange* and *yrsbfsale, AlEdgChYrsVa*:

$$\ln\frac{P_i}{\tilde{P}_i} = \tau_1(T_{i1} - \tilde{T}_{i1}) + \tau_2(T_{i2} - \tilde{T}_{i2}) + \tau_3(T_{i3} - \tilde{T}_{i3}) + ...+ \tau_n(T_{in} - \tilde{T}_{in})$$
$$- \beta_1 distchange + \beta_2 AlEdgChYrsVa \qquad \text{(Model 2b)}$$

This specification allows us to test whether the distance change is gradually, rather than immediately capitalized into house prices. Our final specification examines whether the stages in the RCIP process affected the marginal value of open space. The specification starts with the hypothesis that the open space distance coefficient changes over the three periods (pre-planning is before 1999, the planning announcement period is 1999-2001, and open space draft plan release period is post 2001.) In the repeat sales model this can be formulated as:

$$\frac{P_i}{\tilde{P}_i} = \frac{\gamma e^{\alpha_1 dist_i^1 + \alpha_2 dist_i^2 + \alpha_3 dist_i^3 + \tau_1 T_{i1} + \tau_2 T_{i2} + ... + \tau_n T_{in}}}{\gamma e^{\alpha_1 \tilde{d}ist_i^1 + \alpha_2 \tilde{d}ist_i^2 + \alpha_3 \tilde{d}ist_i^3 + \tau_1 \tilde{T}_{i1} + \tau_2 \tilde{T}_{i2} + ... + \tau_n \tilde{T}_{in}}}$$

Each $dist_i^j$ is the distance to preserved open space multiplied by a time period dummy, so each α_j is the coefficient on distance for the different time periods. We take the logs of both sides and collect terms that share a common coefficient to obtain Model 2c.[6] This specification roughly captures whether open space value differs after new information became available.

$$\ln\frac{P_i}{\tilde{P}_i} = \tau_1(T_{i1} - \tilde{T}_{i1}) + \tau_2(T_{i2} - \tilde{T}_{i2}) + \tau_3(T_{i3} - \tilde{T}_{i3}) + ...+ \tau_n(T_{in} - \tilde{T}_{in})$$
$$+ \alpha_1 RCIP0\_edge + \alpha_2 RCIP1\_edge + \alpha_3 RCIP2\_edge \qquad \text{(Model 2c)}$$

[6] Because the value of open space proximity is presumed to change, open space proximity must be included in Model 2c for pairs of sales that do not take place in the same period even if the distance does not change.

Our matching approach (described below) is predicated on a binary treatment indicator.[7] Therefore, we use several different binary indicators as approximations to the distance change variable. We examine quarter mile increments of the absolute distance change (distchange25 equals one for all changes over one quarter mile, distchange50 equals one for all changes over one half mile etc.) We also use an indicator of whether the property moved from greater than one mile to less than one mile distance from open space (EdgChDum). This is a commonly used distance cutoff for real estate appraisal comparisons. These lead to specification of the form:

$$\ln\frac{P_i}{\tilde{P}_i} = \tau_1(T_{i1} - \tilde{T}_{i1}) + \tau_2(T_{i2} - \tilde{T}_{i2}) + \tau_3(T_{i3} - \tilde{T}_{i3}) + ...+ \tau_n(T_{in} - \tilde{T}_{in}) - \beta_1 W_i \quad \text{(Model 3)}$$

where W_i is one of the dichotomous variables discussed in the previous paragraph. The W_i are the treatment indicators we discuss in the next section.

5. Empirical Strategy

Our analysis relies on a comparison of the price appreciation between properties that become closer to open space as opposed to other properties. We are attempting to discern a treatment effect T in the following equation:

$$Y_i = \beta_0 + \beta_1 W_i + Yr_i{}'\tau \tag{15}$$

Where Y_i is the log of the price ratio, W_i is a continuous or binary indicator of open space distance change, and $Yr_i{}'$ is the vector of sale-year indicators explained in the methodology section. Applications of the repeat-sales methodology usually rely on the assumption that time-invariant factors not included in the regression will not bias coefficient estimates. The hedonic/repeat sales methodology allows attributes of a property to have different coefficients in different time periods. However, because of the sheer number of different property attributes one quickly encounters a curse

[7] Ho et al. 2007 recommend using several binary approximations when the underlying treatment is multi-valued or continuous. In many binary treatment papers the underlying treatment is continuous but only treatment participation is observed.

of dimensionality if all attributes are allowed to have different coefficients in each time period. The researcher has to make the assumption that at least some attributes have constant coefficients over time.

However, this adds some doubt to the analysis of open space value. It seems possible that open space habitat designation could be correlated with confounding property and neighborhood socioeconomic attributes. Figure 1 supports this concern. The open space habitat designations are clearly not randomly spread throughout western Riverside county but tend to be in specific areas, such as near previously designated open space. This is consistent with the ESA's and RCIPs attempt to maintain unfragmented habitat that preserves ecosystems.

The treatment W_i is therefore likely to be correlated with property attributes. If property price appreciation differs across these property attributes (equivalent to the marginal effect of attributes differing over time), then our estimate of the treatment effect is likely to be biased. We address this potential bias by using a doubly robust matching approach as in Imbens (2008).[8] This estimation uses propensity score matching (Rosenbaum and Rubin 1983) as a first step. In doubly-robust estimation matching is used as a preprocessing step to choose a treatment sample that is as similar as possible to the non-treatment sample. Then a normal parametric estimator (linear regression, survival analysis, etc.) is used as a second stage. The advantage of this two stage estimator is that estimates of the treatment affect are less dependent on the parametric assumptions in the second stage. Ho et al. 20007 explain:

> When the data are of sufficiently high quality so that proper matches are available … causal effect estimates do not vary much even when changing parametric modeling assumptions. Finally, since most of the adjustment for potentially confounding control variables is done nonparametrically, the potential for bias is greatly reduced compared to parametric analyses based on raw data. Furthermore, in many situations, the same preprocessing also leads to a reduction in the variance of the estimated causal effects, and so the mean squared error will normally be lower

[8] Our discussion closely follows the Ho et al. (2007) paper.

too.

In our application we use propensity score matching to define groups of treatment and control observations with similar covariate distributions and then use the hedonic/repeat sales methodology as the parametric step to generate estimates of the treatment effect.

Our objective in this paper is to estimate the average treatment effect. Each unit can be assumed to have a random causal effect that is the outcome of a random variable such that:

$$\text{Random causal effect for unit i} = Y_i(1) - Y_i(0) \tag{16}$$

where $Y_i(1)$ is the price ratio with the open space distance change and $Y_i(0)$ is the price ratio if the open space distance change does not occur. Of course, only one of these outcomes is observed in the actual data, the counterfactual must be estimated. We are interested in the mean causal effect, defined in Ho et al. (2007) as:

$$\begin{aligned} \text{Mean causal effect} &= \mathrm{E}[Y_i(1) - Y_i(0)] \\ &= \mu_1 - \mu_0 \\ \text{where } \mu_1 = \mathrm{E}[Y_i(1)] &\text{ and } \mu_0 = \mathrm{E}[Y_i(0)] \end{aligned} \tag{17}$$

There are several different choices for the average treatment effect of interest, (See Imbens and Wooldridge 2008 for a full explanation.) In this paper we focus on the Conditional Average Treatment Effect. This is the effect of an open-space distance change conditional on the pretreatment covariates:

$$CATE = \frac{1}{\sum_{i=1}^{n} T_i} \mathrm{E}[Y_i(1) - Y_i(0) \mid Z_i] \tag{18}$$

where Z_i is a pretreatment vector of covariates that affect treatment status and/or the treatment effect. Pretreatment implies that the Z_i should be determined prior to treatment status. In our case the

Z_i include a vector of property characteristics and neighborhood characteristics drawn from 1990 Census data.

The key assumptions on Z_i are unconfoundedness and overlap. Uncounfoundedness can be expressed as:

$$Y_i \perp T_i \mid Z_i$$

The outcome is independent of the treatment status given Z_i. This is equivalent to the assumption that Z_i should include all pretreatment variables that are correlated with T_i and affect Y_i conditional on T_i.

Another key assumption for propensity score matching is overlap, “that the conditional distribution of Z_i given T_i=0 overlaps completely with the conditional distribution of Z_i given T_i=1” (Imbens and Wooldridge 2008). Formally, overlap is defines as:

$$0 < \Pr(T_i = 1 \mid Z_i = z) < 1, \text{ for all z.} \qquad (19)$$

The goal of propensity score matching step is to select a subsample of data such that T_i and Z_i are unrelated , or:

$$\begin{aligned}&\tilde{p}(Z \mid T = 1) = \tilde{p}(Z \mid T = 0)\\&\text{where } \tilde{p}(.) \text{ is the observed empirical density.}\end{aligned} \qquad (20)$$

The easiest way to select such a subsample of data would be to use exact one-to-one matching where each treated unit is matched to a control unit with the same characteristics. However, exact matching quickly becomes impossible with a large number of covariates, which this application has since we must consider both house and neighborhood characteristics as covariates that could be correlated with T_i and influence Y_i.

In this application, one-to-one matching is not possible, so instead we use propensity score matching for the first stage. Propensity score matching predicts the probability that unit *i* will receive treatment. Rosenbaum and Rubin (1983) show that under the unconfoundedness assumption the outcome and treatment are independent, conditional on the propensity score $e(x)$:

$$T_i \perp Y_i \mid X_i \Rightarrow T_i \perp Y_i \mid e(X_i)$$

As Imbens and Wooldridge (2008) explain, "within subpopulations with the same value for the propensity score, covariates are independent of the treatment indicator and thus cannot lead to biases." In practice, $e(x)$ is not known and is estimated with either logit or probit estimators to obtain $\hat{e}(x)$.

For most matching applications, researchers calculate a simple difference of means after matching to estimate the treatment effect. This would be highly misleading in our case because the timing of the sales is critical to the amount of price appreciation. However, the pretreatment requirement for the covariates implies that the sale year dummies cannot be included in the matching step, because the second sale in any pair is post-treatment. Therefore, this application requires that matching be combined with the parametric repeat sales model.

A number of approaches have been developed for doubly-robust estimators that combine propensity-score matching (Robins 1999 and Ho et al. 2008 are two examples.) Imbens and Wooldridge recommend a subclassification approach for combining matching and parametric estimation. After estimating the propensity score the scores are divided into *J* strata with boundary values $0 = c_0 < c_1 < c_2 < ... < c_J = 1$. Following their notation define B_{ij} for i=1,…N and j=1,…,J-1 as:

$$B_{ij} = \{ \begin{array}{ll} 1 \text{ if } c_{j-1} \leq e(X_i) \leq c_j \\ 0 & \text{otherwise} \end{array} \quad \text{and } B_{iJ} = 1 - \sum_{j=1}^{J-1} B_{ij}$$

The B_{ij} binary variables define each strata. Within each strata the propensity scores are very similar and we can analyze within the strata as if the propensity scores were constant. The general practice is to use the five quantiles as the strata. Because of the large size of the data set we use 10 equal sized strata. We combine this with regression by regressing Model 3 for each of the strata.

Because within each strata the propensity scores and therefore covariate distributions are similar, there individual regressions are not extrapolating far out of sample, as often happens with regression.[9] Imbens and Wooldridge (2008) provide formulae for aggregating the treatment effect and variance over the strata (p. 37).

6. Data

Dataquick provided the information from their multiple sales file on the price and other characteristics of single family residential parcels in Riverside County in Southern California 10 years before and 4 years after RCIP was established (data span 1988 – 2004.) Dataquick is a company that compiles all transaction data from county assessors' offices and supplies it to the real estate industry. The usable transactions are summarized in Table 1.[10]

For our analysis, following the Appendix of Case et al (2006), we considered all those pairs of sales within a particular parcel which occurred in different years. We can have N-1 independent sale pairs, where N is the number of times a property sold. If all the transactions occured in different years then we simply take the price ratios of consecutive transactions. However, when there are multiple sales within a year using consecutive transactions for price ratio does not work. For example, if a property sold 4 times, first two in year one, then you can have the consecutive pairs for 2nd-3rd transactions and 3rd-4th transactions but not the 1st-2nd transactions. Since this property sold 4 times there are 3 independent price ratios that can be formulated. Two of these have already been mentioned above. For the third price ratio we need to choose among either the 1st-3rd, 1st-4th or 2nd-4th transaction pairs. We choose the transaction pair with the closest sequence order. As indicated in

---

[9] Robbins(1999) develops an inverse probability weight (IWP) approach to regression after matching. This approach has difficulty when there are units with very high or low probabilities of receiving treatment (Imbens and Wooldridge 2008). Since we have many control observations with little chance of being treated, we judge the subclassification approach superior in our data set.

[10] From this data, we drop several thousand parcels that transacted more than 10 times, were very large, or had implausibly high or low prices.

Table 1, out of a total of 651,749 possible transaction pairs, only 125,424 could be used for the analysis as 526,325 transactions took place either in the same year or were not independent in nature.

The Dataquick data also contain a number of property characteristics we incorporate into our vector of pre-treatment variables Z_i. These include the number of bedrooms (bedrooms), distance from Corona and Temecula (Corona and Temecula), the distance to open space in 1988 (mfirstdis*),* the lot square feet (lot_sqft), the number of bathrooms (bathrooms), and the square footage of the main structure of the property (sqft_stru). See Table 2 for summary statistics on these variables.

Our second major data source is the information on open space designation. We constructed GIS maps of each open-space habitat preserve and its date of preservation from 1988 through 2004. Then the distance from each house to the nearest preserve was calculated for each sale date, providing the basic data for measuring the distance to open space. Figure 1 gives the map of the preserves for the Riverside County. The preserves which were already in place before 1990 are denoted by the yellow area in the map. Green denotes the preserves which were established during the years 1991-1995, while blue areas are for those during 1996-2000 and red is for the preserves which were established during 2001-2006. The distance from the properties to the preserve ranges from 0.004-7.54 miles with the average distance being 1.24 miles. There are 11,135 observations which showed a change in distance of the property from the preserve over time. Figure 2 gives the frequency of the properties where distances from the preserves changed between sales. The y-axis shows the frequency while the x-axis shows the amount by which the distance changed.

To control for neighborhood characteristics we matched the properties to zip code characteristics based on the 1990 census. The characteristics included median income, education-level variables, and racial makeup. In order to summarize the neighborhood characteristics we include median income and also use factor analysis to estimate three summary variables based on a vector of education and racial characteristics. Future versions of this paper will include more geographically specific tract or block group level data. This may improve the propensity score estimation. We also have data on air pollution levels but there is little variation over Western Riverside county so we do not use it in either estimation step. Table 2 contains summary statistics. [11]

7. Results

Our empirical strategy leads to three sets of results. First, we examine the factors that are associated with treatment and use a logit model to predict the propensity scores. Then we examine repeat sales and doubly-robust estimates of the dichotomous treatment indicators. Finally, we present the repeat sales approach to estimating open-space proximity values with a continuous treatment variable.

7.1 Propensity Score Results

Our first objective is to estimate the propensity score and examine whether propensity score adjustment improves the balance of the covariates. The propensity score estimation assumes the probability of receiving treatment W_i follows a logit model with covariates Z_i. We attempted to include the set of variables in Z_i that is generally included as controls in hedonic regressions such as property and neighborhood characteristics. In addition, we control for the time-period of the first sale with three time period dummies (yyper1 1 is for the first sales before 1995 and is the omitted dummy, yyper2 is for sales between 1995 and 2000, and yyper3 is for sales after 2000.) Our reasoning is that the time period of the first sale in any pair of sales could influence the measured price appreciation because of the movements in the general real estate market.[12] Also, since the first sale is prior to any treatment it is a pretreatment variable. We also use the quadratic of all continuous variables, dropping the squared term where it is insignificant. This corresponds to the advice in Imbens and Wooldridge

[11] Because of the common support requirement and data trimming the alternative dichotomous treatment measures are measured on different data subsamples.
[12] The results are essentially the same if we use dummies for the year of first sale. We use the time period dummies for brevity.

2008 to use flexible forms with higher order terms with large data sets. The results for each of the dichotomous treatment variables are in Table 3.

The results for EdgChDum, the binary variable indicating whether a property went from more to less than a mile between sales, are representative. Nearly all variables are significant, indicating large differences in the distribution of the treatment and control groups. The mfirstdis variable has the largest marginal effect with properties that are farther from reserves in 1988 having much higher probability of going from greater to less than one mile away. Another notable result is that sales transaction pairs are much more likely to be in the treatment group if they occur later in the sample time period. This is just a mechanical consequence of open-space preservation activity occurring more in the latter years of the sample.

The logit results show that the treatment and control groups have quite different covariate distributions in variables that are commonly presumed to influence house values. This implies that standard regression results are questionable since house appreciation rates may differ across the values of these covariates, and that a standard regression measurement would be extrapolating beyond the range of the data in estimating the treatment effect. We use the logit results to generate the estimated propensity scores $\hat{e}(x)$, the estimated probability of treatment given the covariates.

However, after controlling for the propensity score the bias diminishes substantially. Table 4 shows the average difference and the difference in bias between the raw and matched sample when the treatment variable is EdgChDum. In the full sample the treated group properties are larger with more bedrooms, lot area, and interior square footage. They are also located in higher income zip codes. Such significant differences amplify the concern that the treated group houses may appreciate at a different rate than the control group. The matched samples compares the 30 nearest neighbor control units and reduces bias substantially.[13] The average bias falls from approximately 43 percent in the unmatched sample to 1.8 percent in the matched sample. The results are similar for the other treatment measures. When comparing across observations with similar propensity scores, this table shows that there is little difference in covariate means. This finding justifies our choice to estimate treatment effects within each propensity score block.

The estimated propensity scores reveal that there are a large numbers of control observations with near-zero estimated propensity scores (see figure 3 for a histogram comparison.) Regression after subclassification would result in several strata with no or only a few treatment observations. The current practice in these situations is to drop observations with propensity scores close to one or zero, we "trim" the data by dropping observations **with $\hat{e}(x)$ <.025 or $\hat{e}(x)$ >.975** following Crump, Hotz, Imbens and Mitnik (2009). The number of observations dropped due to this trimming varies by the treatment indicator but is always a significant portion of the data set. The trimming improves the efficiency of the estimator but limits the external validity of the estimates. The data does not allow us to draw inferences for properties have very low treatment propensities.

7.2 Results: Dichotomous treatment measures.

The propensity score analysis allows us to proceed with the doubly robust approach. For each treatment variable we divide the propensity sores into deciles and then run Model 3 within each decile. Finally, we compile them into the Imbens and Wooldridge (2008) subclassification matching/regression estimator. For comparison purposes we also present the standard repeat sales treatment effect estimate for the entire sample (see Table 5). The aim of this comparison is to assess whether the doubly-robust approach gives significantly different results than standard regression.

We first look at any residence that moves from more than one mile to less than one mile is distance to open space (EdgChDum). Then, because many residences in this category have fairly

[13] Bias is the mean difference between the averages of the control and treatment unit, divided by the square root of the mean of their respective variances.

small distance changes, for robustness we examine houses that move from more to less than one mile with distance changes of at least 1/20th and 1/10th of a mile (EdgChDum2 and EdgChDum3, respectively.) In the full data sample regression, the coefficients on all the treatment indicators are positive and significant at the 1% level. In the matching estimator, the coefficients on these treatment indicators have coefficients of similar magnitude but somewhat less significant. The coefficient on EdgChDum is significant at the 5% level while the other two coefficients are significant only at the 10% level.

Our second set of indicators looks at the distance change in quarter mile increments. The coefficients on the quarter and half mile distance changes (AlEdgChDum25 and AlEdgChDum50) are insignificant and small in both the regression and matching model. The coefficient on AlEdgChDum75 is positive and significant in the matching model but not in the regression model, and the coefficient is much larger in the matching model. For distance changes of at least one mile (AlEdgChDum100) both coefficients are positive and significant at least the 5% level. The coefficient on AlEdgChDum100 in the matching estimator is almost twice as large as the coefficient in the regression approach, however.

Overall, the matching results are quite similar to the standard regression results. This is especially noteworthy because the samples for the matching estimator are quite a bit smaller than the full sample. If there is a significant difference in the results, it is that the standard regression approach may somewhat under-estimate the open space coefficient relative to the more rigorous matching approach. In evaluating the hedonic/repeat-sales analysis we should keep in mind that it may underestimate the marginal value of open space.

7.3 Results: Continuous specifications

The goal of our continuous specifications is three fold: 1) test whether changes in open space distance significantly affect price appreciation; 2) test whether open space value, if any, is gradually capitalized into house prices; 3) test whether the announcements of the open space plans have any effect on open space values. See Table 6 for all results.

We first estimate Model 2A (Table 6, column 1) to test whether changes in open space distance are significant. We find that the coefficient on *distchange* is positive and significant at the one percent level. The coefficients on the year dummies (unreported) are consistent with the general trend of real estate prices in Riverside County. In order to judge the economic significance of the coefficient, we estimate the average gain in value for those houses that experienced a change in preserved open-space distance proximity relative to the counterfactual where they had not change in proximity. We estimate the houses that underwent a decrease in open-space distance increased their value on average by $2,918 , or slightly less than 1%, relative to the counterfactual where no distance change occurred. The total increase amounts to slightly over $30 million (2004 dollars.) This seems like a significant value given that we are only measuring the change in value from temporary to permanent open space.

As a specification check on these results we include a false treatment dummy. The matching results show that houses where open space distance changed have quite different characteristics than houses where open space distance did not change. Another method to test whether price appreciation rates differ because of these confounding factors is a false treatment dummy. Our data contains many houses that sold several times over the period and where the proximity to open space changed in one sale pair, but not in others. In these cases, we create a false treatment dummy that is positive for an observation if : 1) the house had a change in proximity to open space at some transaction pair; and, 2) this particular observation (transaction pair) did not have an open space proximity change. If the coefficient on the false treatment dummy is positive and significant, it suggests that open space preserves are placed in areas where house values are appreciating in any case. In our case, the false treatment dummy coefficient is insignificant and near zero. Table 6, column 2, shows the result, neither the coefficient magnitude for *distchange* nor significance changes significantly and the coefficient on the false treatment dummy (AlFlseTrtDum) is insignificant. It appears that houses with a change in open space in one transaction pair do not experience greater than normal appreciation in periods where there is not a change in open space proximity.

Our regressions do not support the hypothesis of a lagged capitalization effect. Table 6, column three shows the results of the specification in Model 2b. The coefficient on AlEdgChYrsVa, which is the interaction between *distchange* and the years between the designation of the open space and the second sale in the transaction, is insignificant. This result should be interpreted with care, since most houses do not sell frequently, it may be that there are too few sales to accurately estimate the capitalization effect of open space. Also, in this data a plurality of second sales occurred less than two years after preserve designation (43%) so we may lack the time span necessary to observe a capitalization effect.

Our final specification is Model 2c and is presented in Table 6, column four. This specification allows the open space value to differ by the key periods in the RCIP process. For the pre-1998 period, the coefficient on RCIP0_edge is positive and significant and the coefficient is similar in magnitude to the coefficient in our base specification in column one. However, in the 1999-2001 period the coefficient on RCIP1_edge is approximately 25% smaller than the coefficient on RCIP0_edge and only significant at the 25% level. In the post-2001 period, the coefficient on RCIP2_edge is close to zero and insignificant.

The results are consistent with the marginal value of open space declining as the planning process generates information that open space supply is likely to be large. Pre-1998, it would be reasonable for house buyers to assume that little open space would be preserved because of the rapid pace of development in western Riverside, so proximity to open space would be valuable. However, the announcement of the MSHCP may have added dramatically to the expectation for future open-space preservation and thus lowered the value of proximity. However, we should be cautious drawing conclusions because it is difficult to differentiate between the hypothesis that open space values fell due to the RCIP process and the hypothesis that preferences for open space proximity fell over the period.

8.Conclusion.

This paper presents a new hedonic/repeat sales approach to estimating open-space value. Repeat sales approaches, where there is data on changing open-space designation over time, presents the possibility of removing possible confounding variables through a fixed-effects approach. Our repeat sales/hedonic approach shows statistically and economically significant open-space proximity values.

We also employ a matching approach from the treatment literature to check the robustness of the results. In this case, our treatment is any pair of sales where there is a change in one of our categorization of open space proximity, and controls are the transactions where there is not a change in proximity. Our concern is that the repeat-sales/hedonic approach could be biased if the treatment and control groups are dissimilar. We first estimate a propensity score for the likelihood of receiving a dichotomous treatment and then estimate the repeat-sales/hedonic regression with strata of the propensity score. Our results for dichotomous treatment variables are generally similar for the matching/regression and standard regression approaches.

The propensity score approach may be applicable to the broader open space literature, including cross-section approaches. Our results show that properties that increased their proximity to open space were “better” than control properties- larger lots, houses, and higher income neighborhoods. If these systematic differences hold for other cases where open-space value is measured then a similar doubly-robust approach would be useful in reducing model dependence.

We also find that the open-space values seem to decline coincident with the announcement of open-space planning and acquisition. This is consistent with the real-estate market pricing in new information on the increased availability of open-space amenities. However, the findings are by no means definitive as in this data it is difficult to distinguish between a general time-trend in open space value, a lagged-capitalization effect, and information effects.

Our paper opens up several questions about what the value added is when open space is converted from some other open-space use (usually agriculture) to preserved habitat. Our empirical estimates suggest that homeowners do value this conversion, but do not shed light on the reason for this value. Since this conversion of adjustable to non-adjustable open space is a key decision on the critical urban/wildlands frontier, there is a need for additional empirical and theoretical research on the reasons why this conversion generates value.

Bibliography

[1] S.T. Anderson, and S.E. West, Open space, residential property values, and spatial context. Regional Science and Urban Economics 36 (2006) 773-789.

[2] H. Bang, Doubly robust estimation in missing data and causal inference models. Biometrics 61 (2005) 962-972.

[3] F. Black, and M. Scholes, Pricing Of Options And Corporate Liabilities. Journal of Political Economy 81 (1973) 637-654.

[4] R. Bluffstone, M. Braman, L. Fernandez, T. Scott, and P.Y. Lee, Housing, sprawl, and the use of development impact fees: The case of the Inland Empire. Contemporary Economic Policy 26 (2008) 433-447.

[5] D.M. Brasington, A model of urban growth with endogenous suburban production centers. Annals of Regional Science 35 (2001) 411-430.

[6] D.M. Brasington, Capitalization and community size. Journal of Urban Economics 50 (2001) 385-395.

[7] B. Case, P.F. Colwell, C. Leishman, and C. Watkins, The impact of environmental contamination on condo prices: A hybrid repeat-sale/hedonic approach. Real Estate Economics 34 (2006) 77-107.

[8] R.K. Crump, V.J. Hotz, G.W. Imbens, and O.A. Mitnik, Dealing with limited overlap in estimation of average treatment effects. Biometrika 96 (2009) 187-199.

[9] J. Geoghegan, L.A. Wainger, and N.E. Bockstael, Spatial landscape indices in a hedonic framework: an ecological economics analysis using GIS. Ecological Economics 23 (1997) 251-264.

[10] D.E. Ho, K. Imai, G. King, and E.A. Stuart, Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference. Political Analysis 15 (2007) 199-236.

[11] W.H. Hoyt, Leviathan, local government expenditures, and capitalization. Regional Science and Urban Economics 29 (1999) 155-171.

[12] G.W. Imbens, and J.M. Wooldridge, Recent Developments in the Econometrics of Program Evaluation. Journal of Economic Literature 47 (2009) 5-86.

[13] E.G. Irwin, and N.E. Bockstael, The problem of identifying land use spillovers: Measuring the effects of open space on residential property values, ASSA Winter Meeting, New Orleans, Louisiana, 2001, pp. 698-704.

[14] R.C. Merton, Theory Of Rational Option Pricing. Bell Journal of Economics 4 (1973) 141-183.

[15] P.R. Rosenbaum, and D.B. Rubin, The Central Role Of The Propensity Score In Observational Studies For Causal Effects. Biometrika 70 (1983) 41-55.

[16] V.K. Smith, C. Poulos, and H. Kim, Treating open space as an urban amenity. Resource and Energy Economics 24 (2002) 107-129.

[17] R. Walsh, Endogenous open space amenities in a locational equilibrium. Journal of Urban Economics 61 (2007) 319-344.

Figure 1: Location of the Preserves in Riveside County

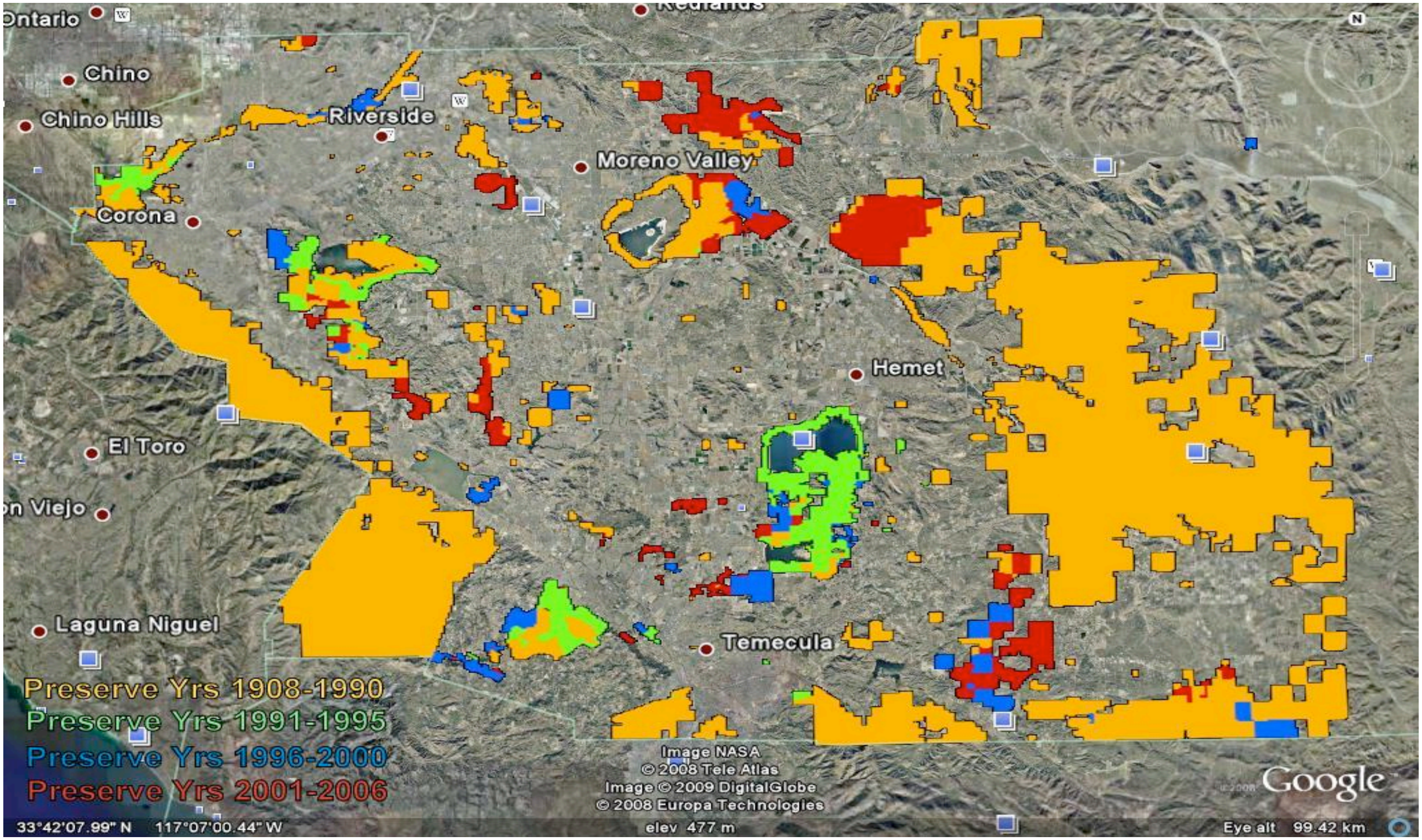

Figure 2: Frequency of Distance Changed Over Multiple Sales

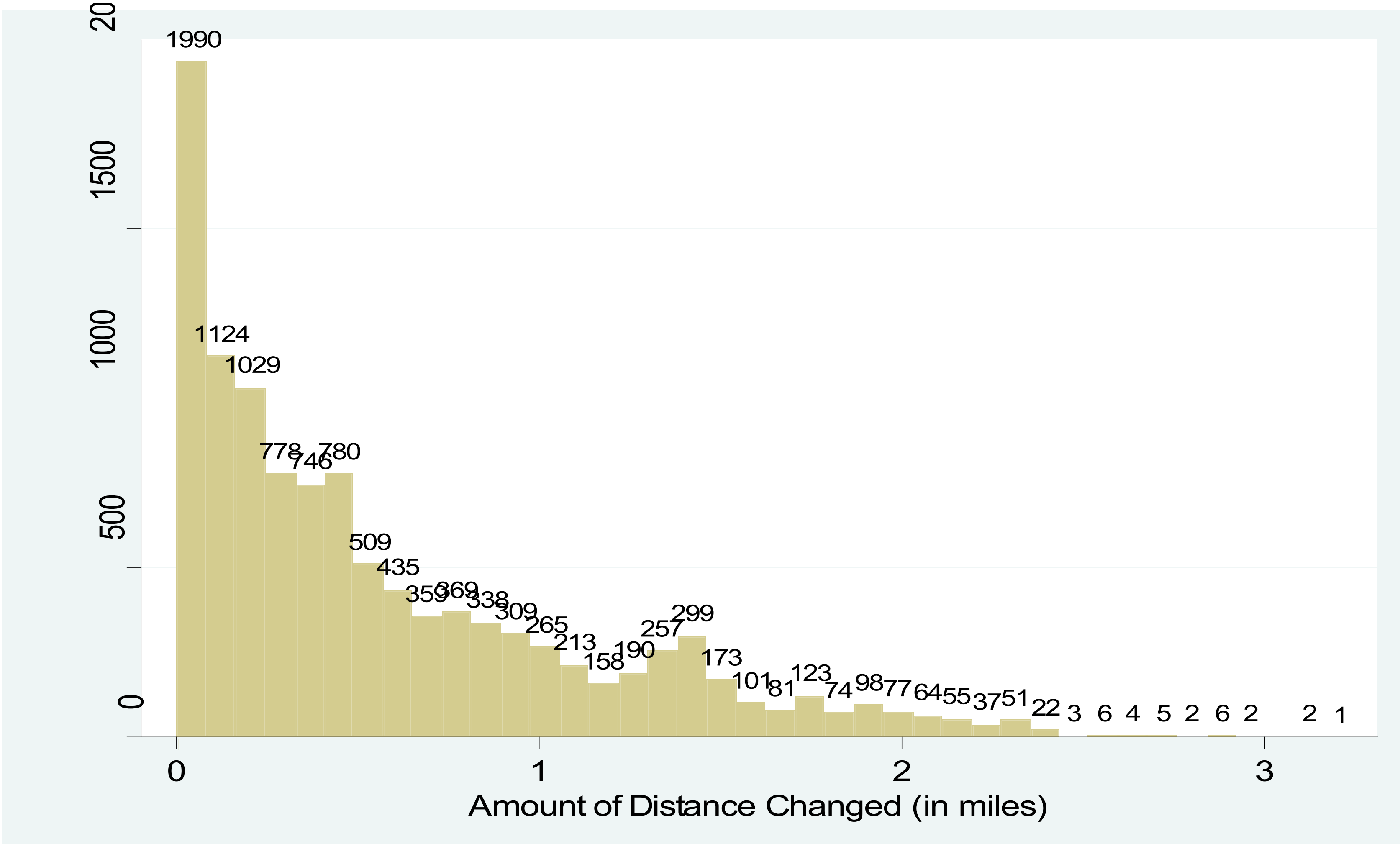

Figure 3: The Propensity Score Distribution for Control Units is Concentrated Near Zero

(Density of Propensity Score by Treatment Status.)

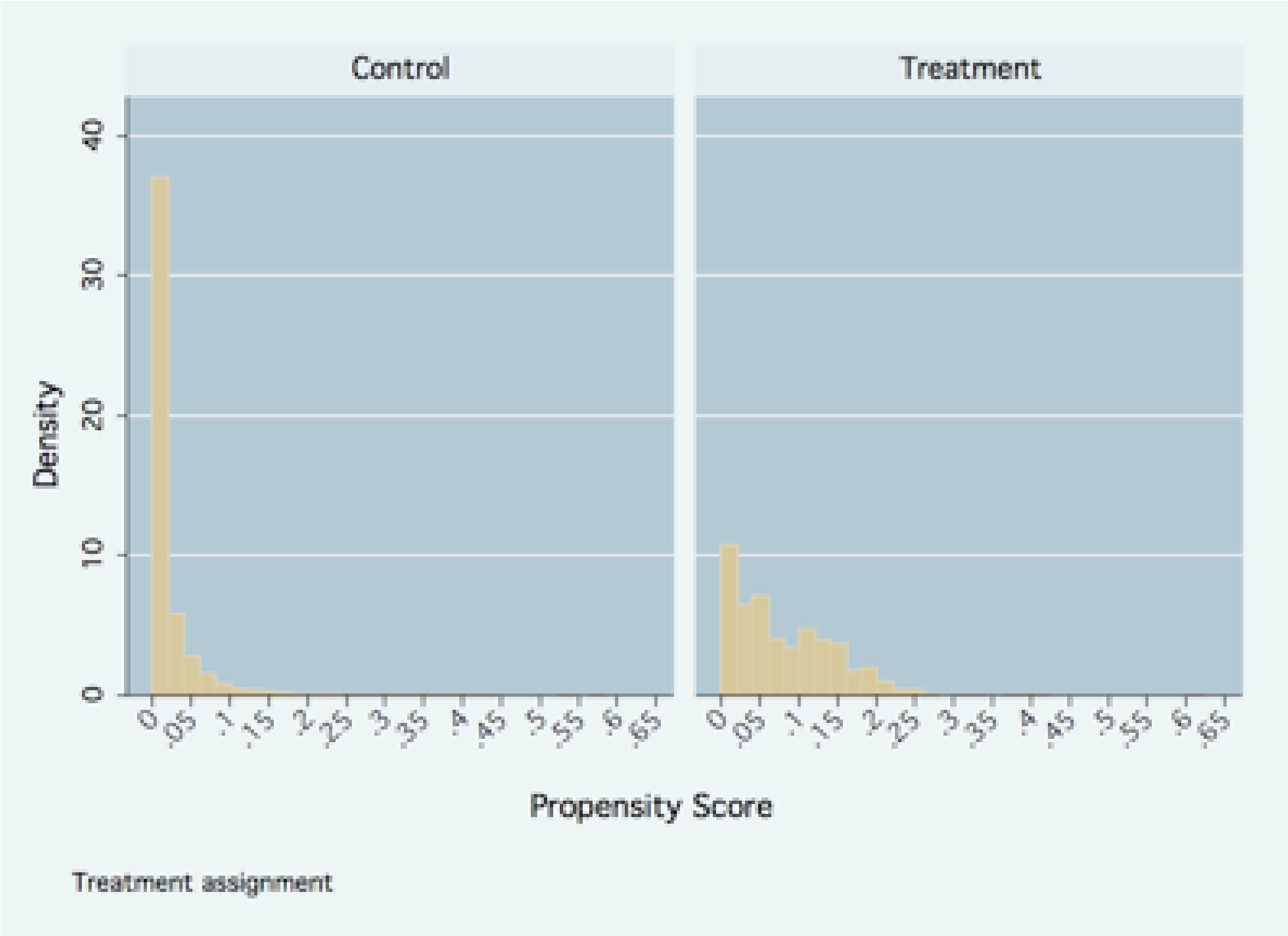

Table 1: Summary of Transactions of the Parcels.

No. of Times Transacted	Properties	Transactions	All Transaction Pairs	Usable Transaction Pairs	Unusable Transaction Pairs
2	54,120	108,240	216,480	54,120	162,360
3	21,647	64,941	194,823	43,294	151,529
4	6,184	24,736	98,944	18,552	80,392
5	1,464	7,320	36,600	5,856	30,744
6	284	1,704	10,224	1,420	8,804
7	65	455	3,185	390	2,795
8+	82	1,874	65,340	1,792	63,548
Totals	83,846	209,270	625,596	125,424	500,172

Table 2: Summary Statistics.

variable	Units	mean	min	max	sd	N
Log Price Ratio	-	0.2143777	-7.698683	1.609438	0.4709472	121,853
A. Propensity Score Controls						
sqft_stru	square feet	1634.149	121.000	18369.000	596.947	121,327
bedrooms	number	3.232	1.000	31.000	0.828	121,298
lot_acres	Acres	0.280	0.010	5.000	0.480	121,853
mfirstdis	miles	1.314	0.004	7.544	0.844	121,853
corona	km	46.314	0.274	99.738	21.061	121,853
temecula	Km	50.104	0.000	113.326	24.598	121,853
income_med	000's	47.857	22.095	70.992	12.438	121,853
factor1	-	-0.202	-1.913	1.326	0.874	121,853
factor2	-	-0.060	-1.304	2.059	0.556	121,853
factor3	-	0.116	-1.480	1.403	0.357	121,853
B. Binary Treatment Indicators						
EdgChDum	-	0.020	0	1	0.139	121,853
EdgChDum2	-	0.020	0	1	0.139	121,853
EdgChDum3	-	0.019	0	1	0.138	121,853
AlEdgChDum	-	0.086	0	1	0.280	121,853
AlEdgChDum25	-	0.056	0	1	0.229	121,853
AlEdgChDum50	-	0.037	0	1	0.189	121,853
AlEdgChDum75	-	0.026	0	1	0.160	121,853
AlEdgChDum100	-	0.018	0	1	0.134	121,853

Table 2 continued, panel C:

C. Continuous Treatment Variables

AllEdgChVa	Miles	0.042	0.000	5.166	0.219	121,853
AlFlseTrtDum	-	0.042	0.000	1.000	0.200	121,853
AlEdgChYrsVa	-	0.177	0.000	26.568	0.948	121,853
RCIP0_edge	Miles	0.601	0.000	7.410	0.860	121,853
RCIP1_edge	Miles	-0.163	-5.781	5.790	0.904	121,853
RCIP2_edge	Mile	-0.386	-5.790	1.150	0.715	121,853

Table 3: Properties In The Treatment Lots Are Larger And In Higher Income Areas. (Logit Estimates Of Treatment Determinants.)

	(1)	(2)	(3)	(4)	(6)	(7)	(8)
	> 1 mile to < 1 mile			All distance changes > than:			
		&change>.05 miles	&change>.1 miles	.25 miles	.5 miles	.75 miles	1 mile
Model: Logit							
Dependent Variable:	EdgChDum	EdgChDum2	EdgChDum3	AlEdgChDum25	AlEdgChDum50	AlEdgChDum75	AlEdgChDum100
Pseudo R2	0.192	0.194	0.196	0.269	0.292	0.359	0.398
bedrooms	0.0871***	0.0926***	0.0910***	0.0306	0.0909***	0.1213***	0.1187***
	[0.0040]	[0.0020]	[0.0026]	[0.1726]	[0.0003]	[0.0000]	[0.0004]
bathrooms	1.7383***	1.8329***	1.8758***	1.5120***	1.4752***	1.4811***	1.5674***
	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]
bathrooms2	-0.1894***	-0.2051***	-0.2121***	-0.1894***	-0.1817***	-0.1901***	-0.1910***
	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0013]
sqft_stru	-0.0006***	-0.0007***	-0.0007***	-0.0001	-0.0004***	-0.0001*	0.0004
	[0.0000]	[0.0000]	[0.0000]	[0.3598]	[0.0000]	[0.0999]	[0.1024]
sqft_stru2	0.0000***	0.0000***	0.0000***	-0.0000	0.0000***	0.0000*	-0.0000*
	[0.0000]	[0.0000]	[0.0000]	[0.7616]	[0.0001]	[0.0888]	[0.0548]
lot_acres	-0.2143	-0.1854	-0.2055	0.3062***	-0.2715***	-0.5733***	-0.4257***
	[0.1399]	[0.2014]	[0.1608]	[0.0001]	[0.0093]	[0.0000]	[0.0048]
lot_acres2	0.1078***	0.1026***	0.1071***	-0.0084	0.1042***	0.1467***	0.1210***
	[0.0005]	[0.0010]	[0.0006]	[0.6387]	[0.0000]	[0.0000]	[0.0003]
mfirstdis	5.2404***	5.2959***	5.3404***	2.4735***	2.8794***	3.8465***	4.6567***
	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]
mfirstdis2	-1.2361***	-1.2458***	-1.2531***	-0.1864***	-0.2394***	-0.3735***	-0.5421***
	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]
corona	0.0741***	0.0722***	0.0710***	0.0719***	0.0789***	0.0930***	0.1334***
	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]
corona2	-0.0008***	-0.0007***	-0.0007***	-0.0012***	-0.0012***	-0.0013***	-0.0016***
	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]
temecula	0.0546***	0.0552***	0.0565***	-0.0363***	-0.0045	0.0097**	0.0460***
	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.2617]	[0.0405]	[0.0000]
temecula2	-0.0005***	-0.0005***	-0.0005***	0.0004***	0.0000	-0.0002***	-0.0004***
	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.8260]	[0.0006]	[0.0000]
income_med_hse	0.2601***	0.2568***	0.2601***	0.1334***	0.1148***	0.1350***	0.2076***
	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]
income_med_hse2	-0.0026***	-0.0025***	-0.0026***	-0.0015***	-0.0013***	-0.0014***	-0.0018***
	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]
factor1	0.8558***	0.8515***	0.8517***	0.9081***	1.0249***	1.2365***	1.3808***
	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]
factor2	0.5135***	0.4639***	0.4792***	0.2496***	0.6155***	1.0448***	0.8752***
	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]
factor3	-0.8107***	-0.8075***	-0.8224***	-1.3478***	-1.3069***	-1.2561***	-1.3375***
	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]
yyper2	0.9057***	0.8857***	0.8792***	0.9448***	0.5879***	0.3816***	0.3020**
	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0001]	[0.0301]
yyper3	2.2562***	2.2454***	2.2388***	1.6201***	1.5164***	1.6435***	2.1747***
	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]
Constant	-21.2626***	-21.3938***	-21.5488***	-12.5871***	-13.4723***	-17.0400***	-24.4485***
	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]
Observations	121255	121255	121255	121255	121255	121255	121255

Robust p values with clustering at the property le

*** p<0.01, ** p<0.05, * p<0.1

Table 4: Controlling for Propensity Score Significantly Reduces Bias.
(EdgChDum is Treatment Variable.)

		Mean			%reduction
Variable	Sample	Treated	Control	%bias*	in bias
bedrooms	Unmatched#	3.4365	3.2282	26.7	
	Matched##	3.4367	3.4416	-0.6	97.7
bathrooms	Unmatched	2.2961	2.0406	43.9	
	Matched	2.296	2.3011	-0.9	98
sqft_stru	Unmatched	1801.9	1630.8	27.7	
	Matched	1801.9	1812.7	-1.8	93.7
lot_acres	Unmatched	0.30296	0.27179	5.7	
	Matched	0.30298	0.29328	1.8	68.9
mfirstdis	Unmatched	1.781	1.3066	65.9	
	Matched	1.7811	1.774	1	98.5
corona	Unmatched	50.788	46.201	22.5	
	Matched	50.789	51.322	-2.6	88.4
temecula	Unmatched	38.014	50.337	-49.5	
	Matched	38.009	36.84	4.7	90.5
income_med	Unmatched	54.479	47.705	56.8	
	Matched	54.474	54.395	0.7	98.8
factor1	Unmatched	0.18475	-0.20851	54.8	
	Matched	0.18468	0.21707	-4.5	91.8
factor2	Unmatched	0.14746	-0.0683	40.8	
	Matched	0.14657	0.13484	2.2	94.6
factor3	Unmatched	-0.03997	0.12013	-52.3	
	Matched	-0.03937	-0.0353	-1.3	97.5
yyper2	Unmatched	0.134	0.3269	-47.1	
	Matched	0.13405	0.13553	-0.4	99.2
yyper3	Unmatched	0.84561	0.55714	66.4	
	Matched	0.84555	0.84142	0.9	98.6

*Bias is the mean difference between the averages of the control and treatment units, divided by the square root of the mean of their respective variances.
Unmatched sample is full set of suitable transactions.
Matched sample compares the treatment observations with the 30 controls with the closest propensity scores.

Table 5: Standard Regression and Matching Regression Estimates are Close.

(Coefficient Estimates for Binary Treatment Indicators.)

	Treatment Indicator	(1) Regression	(2) Matching Estimator
		Model: General Linear model.	
		Dependent Variable:	Ln(PriceRatio)
> 1 mile to < 1 mile			
	EdgChDum	0.0299***	0.0285**
		[0.0015]	[0.035]
and change>.05 miles	EdgChDum2	0.0289***	0.0237*
		[0.0022]	[0.073]
and change>.1 miles	EdgChDum3	0.0286***	0.0220*
		[0.0026]	[0.094]
Distance Changes > than:			
.25 miles	AlEdgChDum25	0.0003	-0.0079
		[0.9586]	[0.407]
.5 miles	AlEdgChDum50	-0.0018	0.0018
		[0.7864]	[0.861]
.75 miles	AlEdgChDum75	0.0102	0.0481***
		[0.1805]	[0.003]
1 mile	AlEdgChDum100	0.0261***	0.0496**
		[0.0034]	[0.03]
	Observations	121854	

Robust p values with clustering at the property l
*** p<0.01, ** p<0.05, * p<0.1
Year controls not shown.

Table 6: Open Space Proximity is Significant Across a Range of Specification.
(Repeat Sales Regression Estimates of the Correlation between Price Ratios and Open Space Proximity.)

		1	2	3	4
		Base Specification	+False Treatment Dummy	Lagged Capitalization	Time Period Effects
Variable Defintion	Dependent Variable: Log of Price Ratio Model: General Linear model with log link.				
	Pseudo R2	0.3958	0.3958	0.3958	0.3959
Distance Change	AlEdgChVa	0.0119** [0.0208]	0.0117** [0.0230]	0.0132* [0.0799]	
False Treatment	AlFlseTrtDum		-0.0082 [0.2521]		
2nd sale minus preserve designation year	AlEdgChYrsVa			-0.0004 [0.8520]	
Time Period Coefficients					
1988-1998	RCIP0_edge				0.0129** [0.0127]
1999-2001	RCIP1_edge				0.0097* [0.0797]
2001-	RCIP2_edge				0.0014 [0.7931]
	Constant	0.1269*** [0.0000]	0.1274*** [0.0000]	0.1269*** [0.0000]	0.1263*** [0.0000]
	Observations	121853	121853	121853	121853

Robust p values with clustering at the property level in brackets
*** p<0.01, ** p<0.05, * p<0.1
Year controls not shown.