Transparency and Takings:
Applying an RD Design to Measure Compensation

Christine Poulos and V. Kerry Smith[*]

This paper reports an analysis of the impact of a new interstate highway on property values in a neighborhood bisected by the road. A with/without analysis suggests the roadway reduced real property values by 16 to 20 percent. To develop these estimates a regression discontinuity design was used with a repeat sales property analysis. The research considers the effect of the temporal and spatial dimensions of the natural experiment permitting the measurement of with/without property values.

JEL Classification Numbers: R52, C30, H54

Keywords: Repeat sales, highway impacts, regression discontinuity design

---

Transparency and Takings:
Applying an RD Design to Measure Compensation

I.   Introduction

This paper uses a regression discontinuity (RD) design to estimate the

compensation required for private landowners due to the negative effects of a new

highway development.[1] In addition to illustrating the information needed to use property

values to estimate the effects of a new regulation (or public infrastructure), we address

two related methodological questions. The first considers how the temporal and spatial

dimensions of a natural experiment satisfying the RD design criterion can influence the

results. The second uses Rosenbaum and Rubin's [1983] propensity scores to evaluate

judgments about the spatial delineation in housing sales that serve as controls in isolating

the effect of a policy affecting property values.

In practice, it is often difficult to be confident that a temporal or spatial distinction

isolates the desired policy effect.   While we might observe housing sales before and after

some policy has been implemented, we rarely know the information that was available to

buyers and sellers at the time of their transactions. Under these circumstances, there is

the prospect for endogeneity between the building and purchase (or sale) decisions and,

in our case, the decision to locate the roadway. Equally important, sometimes a land use

such as a highway conveys benefits to some (i.e. increased convenience with improved

---

[1] Regression discontinuity designs have a reasonably long history in economics. Heckman's [2001] Nobel
lecture provides a detailed overview of the issues in identifying treatment effects. RD designs have
received renewed attention in economics with recent applications by Angrist and Lang [1999], Black

access) and losses (congestion and noise) to others. Distinguishing these separate effects for individual properties can be a difficult identification problem. Our study uses detailed information about the history of the route for the highway, along with a complete record of housing sales, geo-located in relation to the roadway, to overcome the limitations to earlier studies estimating the separate negative impact of a multi-lane highway.[2] The combination of a clear temporal information record, along with the ability to isolate the affected properties and to evaluate the effects of spatially delineated controls assured we could distinguish the negative effects of the new highway.

Our analysis considers the impact to properties located nearby the right-of-way for a new interstate urban loop north of Raleigh in Wake County, N.C. Written correspondence between the N.C. Department of Transportation and a residential developer led the existing homeowners (and new buyers) to believe that the road in question *would not* bisect a residential neighborhood. This information regime persisted for at least three years. During this time, the efficient <u>ex</u> <u>ante</u> bids for properties in this neighborhood would be consistent with the assumption that they would not border a major four-lane interstate. Beliefs changed in 1989 when the Draft Environmental Impact Statement indicated that the land initially reserved for the roadway would be used for that purpose, and the highway *would* bisect the subdivision. We propose to estimate the loss in property values in the area bisected by the highway.

[1999], and Holmes [1998]. See Hahn, et al. [2001] for discussion of some other aspects of the past literature.

[2] Examples of this earlier work on undesirable land uses facing comparable problems to these include McClelland, et al. [1990], Michaels and Smith [1990], Kolhase [1991], and Kiel and McClain [1995]. There is policy recognition of the importance of the information set on property values. Some time ago EPA analysts conducted an analysis of the effect of leaking RCRA landfills for hazardous wastes on nearby residential properties. The objective was to evaluate the benefits from cleanup of these sites as part of the benefit-cost analysis of a new rule requiring that cleanup. However, the specific locations included in the

Under ideal conditions with transparent policy design, compensation would <u>not</u> be efficient (Blume, Rubinfeld, and Shapiro [1984]). However, when the policy process is not transparent, Miceli and Segerson [1996] demonstrate that a conditional compensation rule can be efficient. In the case relevant to our example, when landowners make efficient <u>ex ante</u> decisions and then the policymaker acts <u>ex post</u> to affect the value of their land uses, efficiency requires that the policymaker face the full cost of his/her actions. This requirement implies that compensation for the unanticipated change in the landowners' properties will assure efficient policy. Miceli and Segerson's definition of conditional compensation provides one way to operationalize Justice Oliver Wendell Holmes' "diminution in value" test for regulatory takings.

Our application indicates that the required compensation can be significant. A repeat sales analysis, controlling for selection effects and depreciation, indicates an average loss of 15.5 to 19.5 percent in the real value of the residential properties affected by the roadway. Using the sales prices for the 42 homes that were directly impacted, the average loss ranges from $38,000 to $48,000 per property (in 1998 dollars), depending on the model used to estimate the effect of the highway.

Section two develops a brief overview of the literature on compensation and takings and outlines an amendment to the Miceli-Segerson [1996] framework to match our application. Section three describes the requirements for a RD design in relation to the situations most likely to arise with hedonic studies of regulatory policy or siting decisions. Section four describes the details of the highway case and our data. Section five presents our findings and section six discusses their implications.

---

study were not revealed for fear the existence of the study would be interpreted as an informational signal to the market. See Palmquist and Smith [forthcoming] for discussion.

II.   Compensation and Takings

The Fifth Amendment requires just compensation when the federal government is involved in the taking of private property. In 1922, case law expanded the scope of compensation from situations involving the actual takings of private property to regulatory actions that reduce the value of property. In what has come to be known as the "diminution in value" test, Justice Holmes asserted that government actions constitute a taking if they go too far in reducing property values. While Justice Holmes recognized that compensation for all public action would paralyze government, he also noted that without some compensation rules the government would tend to act until all private property disappeared.

Economic analyses have focused on the conditions under which compensation is efficient. For the most part, the Blume, Rubinfeld, and Shapiro [1984] demonstration that unconditional compensation is rarely efficient has been widely accepted as the primary conclusion of economic analyses.[3] The economic models of compensation presume a clear delineation of the timing of private citizens' and government regulators' (or highway developers') actions. We adapt Miceli and Segerson's model of continuing land uses with fiscal illusion to illustrate how their conclusions apply to our case.

Consider two land uses R (low density residential) and D (high density residential) with an unchangeable land use commitment for two time periods – 0 and 1.

---

[3] As Miceli and Segerson [1996] suggest, Blume, et al. has been the most influential paper on the incentive effects of compensation rules for the decisions of both developers and regulators.

The value (V) of the allocation to R must exceed the value to D for a landowner to commit to R, as illustrated in equation (1).[4]

$$V_R^0 + V_R^1 > V_D^0 + V_D^1 \tag{1}$$

Suppose a policymaker considers a land use for nearby property that would make D the only feasible activity on privately allocated land. The policymakers' efficient selection of this nearby parcel for the public project in period 1 requires that the incremental gain to the public objective per parcel, G, must exceed the private loss, $(\Delta V = V_R^1 - V_D^1)$. As Miceli and Segerson suggest, in practice we may know $V_R^1 - V_D^1$ and not G. Thus, it is reasonable to consider the policymaker's decision as uncertain and to define the probability of selecting the adjoining parcel as $Prob(G > V_R^1 - V_D^1) = p$. If landowners knew there was some prospect the policy would change land values, then efficient landowner choice of R assures that equation (2) is satisfied.

$$(1-p) \cdot (V_R^1 - V_D^1) \geq S \tag{2}$$

S corresponds to $((1-p)(V_D^0 - V_R^0))$ plus any initial costs of selecting R. The condition is derived by comparing the ex ante returns of each land allocation.[5]

    If the probability the policymaker selects the adjoining site depends on the amount of compensation paid to landowners, then it becomes more difficult to assure

---

[4] Since the role of time is not explicitly modeled, we assume that the value of each land allocation in period 1 is discounted to period 0.

[5] The basic structure of the model compares:

$$(1-p)(V_R^0 + V_R^1) + p(V_D^0 + V_D^1) \geq V_D^0 + V_D^1 + r$$

where $r =$ the initial cost of selecting the allocation to R.

Rearranging terms we have:

$$(1-p)(V_R^1 - V_D^1) \geq (1-p)(V_D^0 - V_R^0) + r$$

Let $S = (1-p)(V_D^0 - V_R^0) + r$ and we have equation (2).

efficient decisions from the perspectives of both landowner and policymaker. In this case, Miceli and Segerson show that landowner decisions will be efficient, but the policymaker's choices will not.   However, conditional compensation can induce the efficient outcome (i.e., given ex ante efficient choice by landowners, this case is equivalent to Miceli and Segerson's proposition 1). The conditional compensation rule is that compensation should equal ($V_R^1 - V_D^1$) if ΔV exceeds a threshold, T, and zero otherwise. The threshold is determined by T=S/(1-p) from the ex ante efficient landowner choice in equation (2). This rule is established before the gain, G, from using the adjoining land is known. This compensation rule aligns the policymaker's probability of selecting the adjoining land with the efficient behavior.

In practice, measuring the value differential can be difficult and controversial. Expectations about the likelihood that a project will be undertaken and its effects on nearby residential properties will be capitalized into those residential property values. Thus, reconstructing the time profile of information available to private landowners is critical for interpreting changes in property values. To the extent landowners believe they can affect policy choices by raising the costs of the government's action through their private investments, there is the potential for moral hazard. This possibility, in turn, creates incentives for policymakers to conceal information. As a result, the time profile of information is ambiguous and efforts to reconstruct, retrospectively, the set of information available to private landowners over time are rarely successful.[6] It is

---

[6] Planning documents such as Environmental Impact Statements and Section 6f documents (required under the 1966 Department of Transportation legislation for federally funded highway projects) describe conditions at the time of each draft. As they are finalized they often remove information about the process used to establish consensus opinion and facts. Thus, they do not provide an historical record of either the issues that were resolved or the timing of those resolutions. See Smith, et al. [1999] for a discussion of the types of environmental regulations impacting federally funded highways.

reasonable to expect that the timing and content of information about a project in relation to adjoining land will be correlated with unobserved characteristics of these nearby properties. Thus, the degree of capitalization can be endogenous.

While compensation is often estimated by comparing property values before and after government action, a Congressional Budget Office guidance document [1998] notes that the relevant differential is between the property value with the adjoining use (V_D^1) in comparison to without it (V_R^1). The relationship between this "with and without" measure and a "before versus after" comparison depends on the information about the risk of the government's use of nearby land and the extent to which this risk is capitalized into property values.

Our case study overcomes the information problems by identifying an information time line describing what was known about the location of a highway in relation to a residential subdivision in an area north of Raleigh, N.C. This subdivision, known as Shannon Woods, was bisected by land set aside for the roadway. Uncertainty in the early eighties about the use of this land after some homes were built in the subdivision was resolved in a 1984 letter from the N.C. Department of Transportation (NCDOT) to the developer. This letter created an information regime in which it was believed the highway would *not* bisect the subdivision. This regime changed abruptly five years later when the Draft Environmental Impact Statement (DEIS) unambiguously established that the route would bisect the neighborhood. This discontinuous change in information provides the basis for using the RD design to measure $V_R^1 - V_D^1$. It offers a natural experiment in which the change in information about the path of the roadway can be considered a quasi-random influence on the housing market.

The bisected subdivision appears to have been the primary one impacted by a change in the highway's route. Other land areas around this section of the roadway were developed after this subdivision. Nonetheless, we test for the possibility of a more geographically extensive impact by considering alternative definitions of the control area.

III.   RD Design and Hedonic Property Value Models

Hahn, et al. [2001] have recently demonstrated how discontinuities in the treatment assignment mechanism (i.e., in natural experiments) can be exploited to identify and estimate the effects of those treatments. With an RD design the probability of receiving the treatment can be assumed to change discontinuously as a function of one or more underlying variables. Hahn, et al. discuss the two discontinuity designs most commonly considered in practice – sharp and fuzzy. If h_i is the treatment effect and z_i the observable variable giving rise to a known (non-stochastic) difference in h_i, then a sharp design assumes the deterministic function relating h_i to z_i is discontinuous at a known point.[7]

In one recent example relying on the RD logic, Black [1999] uses a hedonic property value model to compare houses in the same neighborhoods but on opposite sides of the geographic lines that determine the school a child attends within a school district. Test scores measuring school performance make discrete jumps at these boundaries while neighborhoods change in a smooth manner. Thus, the RD logic allows her to isolate how

---

[7] A fuzzy discontinuity design assumes h_i is not a deterministic of z_i. In this case it is a random variable, whose conditional probability ($P(h_i|z_i)$) is discontinuous at a known point.

test scores affect home prices and, through those differentials, the incremental household willingness-to-pay for improvements in educational performance.

In our application there are two features that delineate the roadway treatment: (a) the timeline of information about its location, which is established by the dates of availability of both the NCDOT letter and the Draft Environmental Impact Statement; and (b) the geographic boundary of the impacted subdivision.[8] The abrupt changes in information, described by the timeline, satisfy the conditions of an RD design. With z_i interpreted as the time period in which a property is sold, it indicates the information regime, h_i, under which a property transaction occurred. h_i equals zero if the information indicated that the roadway would not bisect Shannon Woods and h_i equals one when the information indicated that the roadway would bisect Shannon Woods. The treatment effect is measured by the difference in sales prices of properties that sold once when $h_i=0$ and a second time when $h_i=1$. Thus, the repeat sales methodology is appropriate for measuring the treatment effect.

Figure 1 uses a three-dimensional diagram to illustrate how the temporal and spatial attributes of our problem contribute to the definition of our treatment and control groups. On the vertical axis we plot the year of the first sale of each property in the treatment and control areas. On the horizontal axis we plot the year of the second sale of these properties. The third axis (going into the page) plots the radial distance (m) from the center of the subdivision.

To experience the with/without information treatment associated with learning that the highway would bisect the subdivision, the first sale had to take place between

---

[8] The geographic boundary of the subdivision was established with the GIS map for Wake County which identifies the lot and subdivision boundaries.

January 1, 1985 (allowing time for the October 31, 1984 letter to be made available to homeowners) and September 1987 and the second sale had to take place after July 1989. Repeat sales satisfying these conditions correspond to cases where households buy *after* the NCDOT letter was sent and then sell *after* the DEIS was released July 1989. Records of public hearings associated with the development of the Environmental Impact Statement about the roadway indicate that information about the NCDOT reconsideration of the route bisecting the neighborhood along with other alternatives was available beginning in September 1987.[9] Thus, initial sales between January 1, 1985 and September 1987 (prior to the first of these public hearings) in the figure are the only ones that can be assumed, unambiguously, to be associated with a "no bisecting highway" information regime. Sales between September 1987 and July 1989 (the date of the release of the DEIS), associated with the shaded area in the figure, relate to times when the information regime is not clear. During this time, the route bisecting the neighborhood, as well as other alternatives, were under consideration. Thus, it is reasonable to assume that there was uncertainty among buyers in this period about the ultimate disposition of the roadway. Accordingly, we deleted all transactions falling in this interval from our analysis. The area "T" corresponds to the treatment group. In the figure we use a box defined by initial sales from 1984 to 1987 and second sales of the same property after July 1989. It is bounded on the third dimension (m) by the subdivision boundary (m*), which defines the geographical extent of our treatment group.

---

[9] The next specific correspondence between the NCDOT and the subdivision's home owners association was in August 1988. A series of meetings beginning in September 1987 presented information about alternative routes. However, there was no definite information about the likely final route until the Draft

The spatial boundary of the treatment effects (m^*) was defined <u>ex</u> <u>ante</u> when the developer arranged to coordinate construction of homes in a single subdivision. The urban loop road passes through multiple land areas near this sub-division, but beyond m^*. We propose that sales of these other properties serve as "controls," provided (1) they can be assumed to have a clear and constant information concerning the location of the roadway in relation to their neighborhoods and (2) properties in the control group are not systematically different than properties in Shannon Woods. The information we assembled about the routes for the roadway suggests that during the study period (1985-1998), Shannon Woods was the only subdivision in the immediate area for which there was changing information about the route of the roadway relative to the neighborhood. The uncertainty about the route through Shannon Woods arose because the developer had infringed on a predefined right-of-way for the roadway, as we discuss in the next section.

Like the treatment repeat sales, the control repeat sales are delineated in the temporal and spatial dimensions. Control repeat sales fall in the areas labeled "C" and "D" in Figure 1. The areas labeled "C" correspond to "temporal controls." These are repeat sales taking place in the subdivision completely before or completely after the change in information about the highway's route. Repeat sales labeled "D" in Figure 1 correspond to properties outside the subdivision ($m>m^*$), regardless of the timing of their two sales. As the cross-hatching indicates we eliminated the sales involving the period of uncertainty (September 1987 through July 1989) from our control as well as our treatment group observations.

Environmental Impact Statement was issued in July 1989. The other subdivisions near the rerouting did not have homes affected. See Holbrook [2000].

Hahn et al. [2001] note that the RD model assumes that the outcome measure in the absence of treatment (in our case, a measure of price with information that the roadway will <u>not</u> bisect Shannon Woods) is constant within an arbitrarily defined range around the threshold defining the treatment variable. For this condition to be satisfied we need to control for three other sources of differences amongst properties in the treatment and control groups: (1) changes over time that are unrelated to the treatment; (2) changes over space that are unrelated to the treatment; and (3) changes in a property's characteristics.

First, in any area slated for a major highway it is reasonable to expect that other, residential, commercial, or retail development would be being planned to take advantage of the improvements in access to the rest of the metropolitan area. Our data set spans 14 years (1985-1998) and price changes in the data set are likely to reflect other changes also taking place in the area over this long time interval. Thus, we include year specific fixed effects to control for them.

Second, expansion in the geographic boundary defining the control transactions (m in Figure 1) risks mixing the effects caused by proximity to these other uses. It also raises other questions about our ability to assume the smooth (continuous) change in the other characteristics of residential properties and therefore the comparability between residential properties in Shannon Woods and properties in the expanded geographic area. In Section IV, we describe the use of propensity scores to evaluate selections for the geographic boundary for the control group. These methods indicate that a geographic boundary 1.5 miles from the center of Shannon Woods is appropriate. Nonetheless, we

also consider spatial boundaries beyond 1.5 miles to check whether our findings were sensitive to this assumption. They were not.

Third, our use of the repeat sales framework helps control for heterogeneity in property characteristics over time. To develop this argument more completely, consider a simple hedonic price model, expressed as a semi-log equation, with the log (p_i) a function of time invariant characteristics, c_{ik}, time variant characteristics, x_{jt}, and the treatment effect, h_i, as in equation (3).[10]

$$\ln p_{it} = a_0 + \sum_{k=1}^{K} a_k c_{ik} + \sum_{j=1}^{L} b_j x_{jt} + h_i + \varepsilon_{it} \tag{3}$$

$$\varepsilon_{it} = \text{random error}$$

ε_{it} is a function of three components, as in equation (4a).

$$\varepsilon_{it} = \mu_i + v_t + e_{it} \tag{4a}$$

where μ_i represents time invariant, property-specific unobservables, v_t represents time-variant unobservable heterogeneity, and e_{it} represents random variation.

We can use a repeat sales model to control for both the observed and unobserved property-specific heterogeneity. The effect of the change in information about the roadway is isolated by identifying the properties in the subdivision with two sales that span the information change in comparison to those properties that sold twice with a constant information regime or those selling twice outside the subdivision. The latter

---

[10] The Cropper, et al. [1988] evaluation of functional specifications for hedonic price equations generally favored a linear Box-Cox form when the correct specification for the attributes of the house and its location was known. For cases where there were omitted independent variables from the hedonic price equation, the semi-log was found to be among the most robust based on the properties of its estimates of marginal effects.

would also have constant information regime because these property values remained

unaffected by the NCDOT letter.

To illustrate how this process would work, consider the pairs of hedonic models

that contribute to the repeat sales model under different time and location conditions.

The first pair comprises an initial sale taking place after the NCDOT letter suggests the

roadway will not bisect the neighborhood but before the new information (i.e., between

January 1985 and September 1987), say time t. The initial price model is shown in (4b).

Since h_i equals zero, it is dropped from the model.

$$\ln(p_{it}) = a_0 + \sum_{i=1}^{K} a_k c_{ik} + \sum_{j=1}^{L} b_j x_{ijt} + \varepsilon_{it} \tag{4b}$$

If the second sale for this property takes place in period t+s, after t^* when the new

information associated with the DEIS is available (i.e., after July 1989), then the hedonic

price model will include the effect of the new information for properties in the

subdivision, as reflected in equation (4c).

$$\ln(p_{it+s}) = a_0 + \sum_{k=1}^{K} a_k c_{ik} + \sum_{j=1}^{L} b_j x_{ijt+s} + h_i + \varepsilon_{it+s} \tag{4c}$$

Taking the difference between equations (4c) and (4b) provides one way to isolate the

change in h_i, control for observable features of heterogeneous properties, as well as

"difference-out" time invariant, property-specific unobservables, $ì_i$. For repeat sales of

Shannon Woods properties in which both transactions occur either before or after t*,

differentiating will eliminate the effect of the information about the roadway (h_i). Sales

outside the subdivision, regardless of the timing of the sales, will not have the h_i effect

in the repeat sales equation either because the change in information about the route does

not impact their prices. These different types of control cases correspond to blocks "C"

and "D" in Figure 1.

Our application is fairly unique in the ability to reconstruct the time profile of

information using the NCDOT letter and the release of the DEIS. Because the change

affects primarily one subdivision we can use this time interval and the geographic

distinction in impacts to meet the requirements for a sharp RD design and estimate the

impact of the highway route. The measure of impacts is equivalent to the compensation

measure in Miceli and Segerson's model.[11]

Equation (5) illustrates the general form of the repeat sales model derived from

equations (4a) through (4c). Since the treatment effect is defined in two dimensions, the

product of two variables, D_1 and D_2, provides the method to measure the treatment effect,

$h_i = è$. D_{1i} is a qualitative variable identifying properties in Shannon Woods (=1 if in the

subdivision; =0 otherwise) and D_{2i} is a qualitative variable identifying the timing of sales

---

[11] This can be shown by recognizing that a property's price is the discounted stream of time-period-specific property values. By substituting (1) into (3), the price of a low-density residential use property can be expressed as:

$$p_{it} = \sum_t V_{iR}^t = V_{iR}^0 + V_{iR}^1$$

where period 1 could represent either a single future period or a discounted stream of future time periods. We also assumed the effect of infrastructure is a constraint proportion of price by using a semi-log specification. p_0, the property value without infrastructure, and p_1, the property value with infrastructure can be rewritten as:

$$p_{0it} = V_{iR}^0 + V_{iR}^1 \quad \text{without infrastructure in Shannon Woods}$$
$$p_{1it} = V_{iR}^0 + V_{iD}^1 \quad \text{with infrastructure in Shannon Woods}$$

The repeat sales model implements Hahn et al.'s treatment effect estimator by taking the difference between the outcome if interest with and without the treatment. This treatment effect is equivalent to the value differential necessary for the efficient outcome in Miceli and Segerson's model.

$$p_{1it} - p_{0it} = \left(V_{iR}^0 + V_{iD}^1 \right) - \left(V_{iR}^0 + V_{iR}^1 \right) = -\left(V_{iR}^1 - V_{iD}^1 \right)$$

16

that imply a change in information between sales (=1 if the first sale occurs between

January 1, 1985 and September 1987, and the second sale occurs after July 1989; =0

otherwise). è reflects the fact that we are assuming a constant proportionate effect on the

property values independent of each house's distance to the roadway.[12]

$$\ln(p_{it+s}) - \ln(p_{it}) = \sum_{j=1}^{L} b_j (x_{ijt+s} - x_{ij_t}) + \theta \cdot D_{1i} \cdot D_{2i} + (v_{t+s} - v_t) + (e_{it+s} - e_{it}) \quad (5)$$

As noted, transactions from September 1987 through July 1989 are dropped from the

sample because the public meetings associated with alternative route selections began at

that time. The resulting uncertainty would be reflected in sale prices for housing in this

subdivision.

Our implementation of this model also accounts for two influences on the

observations included in a repeat sales analysis: selection effects and time variant

characteristics. First, for any house to sell, the bid offered must exceed the seller's

reservation price. To compose a sample with two sales of each house, this must happen

twice in the time span with sales price information. It is certainly possible that this

double, sequential, selection effect, if ignored, would bias our estimates of the effects of

the highway on property values in the area, both those in the subdivision and in the

control areas. To estimate the first selection effect requires information on properties in

---

[12] If we assume that $P_i = \tilde{P}_i \cdot (1+\alpha)^{D_i}$ where $D_i = D_{1i} * D_{2i} = 1, 0$ and defines when the effect is present and α is the percentage impact on Y of a change in D_i from 0 to 1, then in a repeat sales framework that isolates $(1+\alpha)^D$ through differencing we can interpret the parameter è as first discussed by Halvorsen and Palmquist [1980]

$$\hat{\theta} = \ln(1 + \hat{\alpha})$$

Solving for $\hat{\alpha}$ we have: $\hat{\alpha} = e^{\hat{\theta}} - 1$ with $\hat{\theta}$ the estimated parameter. In our estimates we include the Kennedy [1981] proposed adjustment for bias due to the nonlinearity in the transformation. This approach calls for $\hat{a} = \exp(\hat{q} - 1/2 \operatorname{var}(\hat{q})) - 1$, where $\operatorname{Var}(\hat{q})$ is the variance in the estimate of q.

the area that are not included in the set of sales. Most hedonic studies are limited to the sales.

Gatzlaff and Haurin [1997] discuss this framework with a first sale observed only if a buyer's offer price, p_i^O, exceeded the seller's (or builder's) reservation price, p_i^R. To observe a second sale, the house must have sold once <u>and</u> a second buyer's offer price must have exceeded the seller's reservation price. Thus, the two selection rules (s_{1i} and s_{2i}) are defined by the fact that (s_{1i}, s_{2i}) is observed if $p_i^{O1} - p_i^{R1} \geq 0$ and $p_i^{O2} - p_i^{R2} \geq 0$. The second superscript on prices refers to the sale order. We follow Gatzlaff and Haurin and use Tunali's [1986] generalization to the Heckman [1979] two-step correction for sample selection effects and compute the associated inverse Mills ratios.

Second, a set of variables, corresponding to ($x_{ijt+s} - x_{ijt}$) on the right side of equation (5), attempts to control for other time variant effects on the properties and the area. Two sets of variables are included to control for time effects and land use change effects. The first set of these controls is a set of time dummies for the pairs of years involved in each repeat sale. The year is coded as –1 for the initial sale date and +1 for the second sale date and zero otherwise. These are akin to Black's boundary fixed effects. The second set measures the change in land use in the area. The public records on land classifications allow identification of current land uses. Vacant land (i.e., privately owned, undeveloped land) is estimated for each year of sales using a backward recursion method. Beginning with land parcels that have been classified as vacant lots with no improvement in the most recent year of our data (1998), we add the parcel areas

18

for each newly constructed house (based on the year built of the homes involved) to

estimate the stock available in the preceding year. Thus, to estimate the stock of vacant

land in 1997 we add the land for all the parcels that had homes built in 1998 to the stock

of areas identified in 1998 as vacant. This process is repeated for 1996 and so forth back

to 1980. This allows us to impute a historical record of the amount and location of

undeveloped private land in our study area.[13] Unfortunately, the data sets do not contain

sufficient information to define a historical record for other types of land uses.

The newly defined set of reported and imputed vacant land in each year is then

used to identify the closest vacant parcel to a residential property sold in a given year.

We measure the difference in this distance to the closest vacant parcel as a proxy for land

use changes leading to denser development. We hypothesize that property owners have

preferences over the density of development surrounding their property and, thus, the

change in the distance to the closest vacant land ($\ddot{A}$dvacant$_i$) is used to capture this effect

for each property between the two sales.[14] Our estimating equation is given in equation

(6).

$$\ln(p_{it+s}) - \ln(p_{it}) = \beta_0 + \sum_{k=1}^{13} \beta_k \, year_{ik} + \beta_{14} \Delta dvacant_i \qquad (6)$$

$$+ \gamma_1 \lambda_{1i} + \gamma_2 \lambda_{2i} + \theta \cdot D_{1i} \cdot D_{2i} + \tau \cdot D_{2i} + u_i$$

The terms in equation (6) are different from equation (5) because: (a) we have been

explicit about the time variant variables (i.e., the fixed effects for the years of sales and

the change in the distance to vacant land between the two sales, reflecting the increased

---

[13] The idea for this strategy to construct a measure of vacant land is due to Walsh [2000].

development pressure in the area as well as open space); (b) we introduce the inverse

Mills ratios ($\ddot{e}_1$ and $\ddot{e}_2$) for the double sequential selection effects[15]; and (c) we include a

term for the separate effects of the time interval with the information effect ($\tau \cdot D_{2i}$).

The net result of these changes will be reflected in the properties of our error. u_i is

different from the error term in equation (5) because it incorporates, implicitly, the fact

that selection effect terms we used to estimate the model were separately <u>estimated</u>

inverse Mills ratios following the logic of a two-step Heckman approach. Substituting

for the $\ddot{e}$'s in equation (6) with random variables, $\hat{\lambda}$'s, we expect that the induced error in

the resulting model will be heteroscedastic. As a result, we compute the standard errors

---

[14] Some of the properties could not be assigned information on the distance to the nearest vacant parcel (531 of 2917 in our initial sample). They were assigned the average value of distance to a vacant parcel in the year they sold.

[15] Our repeat sales difference in the log of house prices is observed if $p_i^{O1} - p_i^{R1} \geq 0$ and $p_i^{O2} - p_o^{R2} \geq 0$.
If we specify a selection model that describes the prospects we observe $p_i^{O1} > p_i^{R1}$ as:

$$y_{1i} = x_{1i}\gamma_1 + U_{1i}$$

and a comparable model for the second sale $p_i^{O2} > p_i^{R2}$ (which for a repeat sale will be conditional on observing the first sale) we have:

$$y_{2i} = x_{2i}\gamma_2 + U_{2i}$$

Our difference in selling prices model is conditional on observing both prices. Thus, the error's expectation is conditional on $U_{1i} > -x_{1i}\gamma_1$ and $U_{2i} > -x_{2i}\gamma_2$. The Tunali generalization of Heckman uses the conditional expectation of a trivariate normal to derive:

$$E(u_i | U_{1i} > -x_{1i}\gamma_1, \ U_{2i} > x_{2i}\gamma_2) = \rho_{13} \frac{f(x_{1i}\gamma_1)F(C_{2i})}{G(x_{1i}\gamma_1, x_{2i}\gamma_2, \rho)} + \rho_{23} \frac{f(x_{2i}\gamma_2)F(C_{1i})}{G(x_{1i}\gamma_1, x_{2i}\gamma_2, \rho)}$$

where f( ), F( ) are the standard normal density and distribution functions for a univariate normal random variable. G (a, b, ñ) is standard bivariate normal distribution function with ñ the correlation function.

$$C_{1i} = \frac{x_{1i}\gamma_1 - \rho \cdot x_{2i} \cdot \gamma_2}{\sqrt{1 - \rho^2}}$$

$$C_{2i} = \frac{x_{2i}\gamma_2 - \rho \cdot x_{1i} \cdot \gamma_1}{\sqrt{1 - \rho^2}}$$

The relevant inverse Mills ratios in our two-step estimator are then computed as:

$$\lambda_1 = \frac{f(x_{1i}\gamma_1) \cdot F(C_{2i})}{G(x_{1i}\gamma_1, x_{2i}\gamma_2, \rho)}$$

$$\lambda_2 = \frac{f(x_{2i}\gamma_2) \cdot F(C_{1i})}{G(x_{1i}\gamma_1, x_{2i}\gamma_2, \rho)}$$

for our models with Huber robust methods. In addition, to gauge the potential impact of these issues we report several different models in Section V and focus on the sensitivity of our measure of the treatment or information effect to these judgments.

IV.   Background and Data

The area impacted by the new roadway, in the northern portion of Wake County, N.C, has experienced rapid growth in recent years because it is convenient to the state capital, Raleigh, as well as to the nearby Research Triangle Park.[16] Planning for the roadway began long before the surge in population over the past decade and a half. Initial discussions began in the 1960's. Its approximate location was identified in 1967 and the route for the northern route was reported in more definitive terms in the Raleigh Thoroughfare Plan issued in 1978. Thus, the decision to construct this outer loop and its overall location preceded the property development and sales covered in our sample period. The primary area of concern is a single subdivision, known as Shannon Woods, which is bisected by the right-of-way for the roadway.

In the early 1980's a developer acquired land for the subdivision and had begun the development process. Prior to October 1984, the homeowners in the development became aware that the developer had constructed homes on one street, named Bantry Court, which encroached on the right-of-way reserved for the roadway, leaving an insufficient area for the four-lane interstate. Figure 2 illustrates the right-of-way, the road

---

[16] The recent Center on Urban and Metropolitan Growth [2000] reports documents that North Carolina's rapid growth is affecting resources and externalities associated with the quality of life in the region. The report notes that North Carolina was ranked fifth among states in the number of acres of land developed between 1992 and 1997. 781,500 acres were developed over this period. The annual rate means an area about the size of the city of Charlotte, North Carolina was developed each year.

encroaching on the roadway, and the houses that were affected by the action (i.e., with the lighter shading of the lots). The resulting controversy led to considerable confusion as to the ultimate location of the road in relation to this neighborhood, as well as the disposition of the land in the right-of-way if the outer loop's route did not use it.

Figure 3 reproduces the exact text of the letter from the North Carolina Department of Transportation, dated October 31, 1984, to the developer indicating the right-of-way would no longer be used in highway construction.[17] It suggested that the developer dedicate the area to Wake County. This letter, which was shared with the Homeowners Association for the subdivision, implied that the path of the loop road would be changed, leaving existing right-of-way land available for other uses, including other residential development or open space.[18]

Starting nearly three years after the letter was sent, in September 1987, the North Carolina Department of Transportation held a series of public hearings about the impacts of the roadway as part of the activities required for the preparation of the Environmental Impact Statement. These meetings raised concerns that the NCDOT's position, expressed in the letter, would be reversed and the DEIS, released in July 1989, confirmed them. The DEIS identified four routes as options in this area of the roadway and the route (identified as N) bisecting the subdivision was identified as the preferred

---

[17] The N.C. Department of Transportation provided us complete access to their records. Without this access we would not have been able to reconstruct this history. We present the exact text of the letter because the Department has suggested that the letter was intended to inform the developer that a larger buffer area was needed. Our interpretation seems to be the one, based on the public hearings, that homeowners in Shannon Woods and others derived from the letter.

[18] We are especially grateful to Michelle Holbrook who researched the files of the N.C. Department of Transportation to document the disposition of every home sale in this area following the development. This process involved a time consuming search of archived records in old warehouses. She reviewed minutes of all Homeowners Association meetings around the dates of the controversy and discovered indirect evidence of the existence of this letter. She uncovered this letter in the archived records and, as a

alternative. To make segment N sufficiently wide for highway construction, the DEIS

indicated that the Bantry Court roadway – including eight residences on that street –

would be removed. Though the Shannon Woods Homeowners Association opposed the

alignment, it remained the preferred alternative and was included in the Final EIS, which

was released in 1990. Construction of segment N was to be completed in December

2000. As noted earlier, the time profile defined by the letter and the DEIS provide the

basis for the RD design, and the boundary of the subdivision is the spatial boundary of

the treatment group.

Four sets of information were combined to form our database. First, the preferred

alignment of the roadway was mapped into a geographic information system (GIS) made

available by the National Highway Planning Network and the Federal National Highway

Administration. Second, GIS data for Wake County, current as of February 2000, were

obtained from the Wake County GIS Department. The GIS data include: a property

identification number; the area of each parcel; the land classification, as well as other

information on boundaries; right-of-way; and subdivisions. Third, property sales data for

Wake County are drawn from a commercially available data set on property transactions.

The data were purchased from TransAmerica Intellitech, a company that uses tax records

to assemble data on transaction prices, transaction dates, and a limited set of structural

characteristics of the houses involved. Fourth, data from the Wake County Tax

Assessors office were used to develop a more detailed set of property characteristics.

These data also allow us to include properties that were not for sale during the time of our

---

result, completed the record associated with the history of events proceeding the DEIS. See Holbrook
[2000] for a summary of the record.

study. They contribute to the initial selection model as well as to the development of the model for the propensity score analysis discussed below.

Table 1 lists the number of residential property sales by year in our study area, which comprises the Shannon Woods subdivision along with all the residential properties in an area 1.5-miles from the center of that development. The average sales price is deflated using the Consumer Price Index (base = 1998). The reported number of sales in the table is a count of the number of pairs of repeat sales in the study period. They are listed by the year of second, or most recent, sale. The time span begins with the period for the first sale, which we take to be January 1, 1985 (allowing time for the letter to be made available to homeowners) through August 31, 1987. The repeat or second sale must take place after release of the DEIS and is taken to correspond to August 1, 1989 to December 31, 1998.

The selection of the geographic boundary for the control group, or the properties outside the subdivision that are comparable to Shannon Woods properties, is more subjective. The base analysis uses a 1.5 mile boundary, but we considered three variations in the 1.5-mile radius, each an increase of 0.5 miles. The sample of comparison or control properties changes as we select m, or the geographic boundary for are "D" in Figure 1. Table 2 identifies these changes by illustrating how the sample composition changes with this difference in the definition of the boundary for the control properties.

Three aspects of this comparison are important. First, notice that the number of properties with repeat sales in Shannon Woods (42), the treatment group, does not change. Second, the number of sales dropped from the sample for different reasons does

change. These include: (a) sales between September 1, 1987 and July 31, 1989 (the

period of uncertainty about the route); (b) properties with a building permit associated

with an unspecified improvement [19]; (c) transactions involving land sales; (d) transactions

likely to be other than "arms-length" sales (and outliers) identified by prices below or

above the first and 99th percentiles, respectively; and (e) sales of relocated properties in

Shannon Woods (i.e., the eight properties that were removed from the right-of-way).[20]

Finally, we see that the composition of the study sample also affects the samples relevant

for the selection models, as indicated in the lower panel in the table, where we identify

residential properties that never sold, as well as those in the one sale and two sale

categories.

Before turning to the results, we consider how the geographic boundary of the

study area would affect our results. The RD design assumes the other features of

treatment and control samples change smoothly, so that the only reason for a change in

the log price difference can, in our case, be attributed to the effect of the roadway

bisecting the subdivision. Past applications have generally relied on analyst judgment to

define these boundaries and evaluate exceptions to them. Black's study, for example,

imposed limits on the number of elementary schools with overlapping grade levels in a

district, required absence of intra-district choice programs, and dropped cases where the

attendance districts were poorly defined or unavailable. All of these choices seem

---

[19] Our data sets do indicate whether a building permit was issued. However, there are incomplete records on the date the permit was issued and no information on the permitted activity. As a result, we do not know which of the housing attributes changed between sales. For this reason we exclude these pairs of transactions from the sample. Table 2 provides a record of the number involved.

[20] The Department of Transportation purchased these homes and the prices associated with these transactions do not reflect the values of these properties. They reflect the cost of acquiring comparable housing for the displaced homeowners.

motivated by assuring a sharp RD design. Another consideration she uses is potentially

more idiosyncratic, noting that:[21]

> "Because of concerns about neighborhood differences in opposite
> sides of an attendance district boundary, I was careful to omit boundaries
> from my sample if the two attendance districts were divided in ways
> that seemed to clearly divide neighborhoods; attendance districts
> divided by large rivers, parks, golf courses, or any large stretch of land
> were excluded" (p. 582).

To offer a more systematic approach for evaluating these choices, we use the

logic underlying most matching estimators (see Smith and Todd [2000]). That is, we use

the Rosenbaum-Rubin [1983] propensity score to gauge the probability other homes

around the Shannon Woods subdivision would be judged "similar enough" to the

Shannon Woods homes to have been classified with them in a neighborhood

classification based on observable characteristics.[22] The independent variables in this

probit model include square footage of living space, travel distance to the center of

Raleigh, distance to I-540, distance to a golf course, and the year built. Table 3 reports

the average propensity scores, selected demographic characteristics of these census tracks

(based on the 1990 Census), and the mean values for a few descriptive characteristics of

the properties in each zone. The propensity score for each property is predicted and these

scores are summarized by dividing the region into six concentric rings, each one-half

mile wide (see top row of Table 3).

---

[21] Holmes [1998] reports a similar set of judgments in his analysis of right to work laws, noting that he used counties with 25 miles of state borders that differ in right-to-work laws. However, he excludes Western states from the analysis because of the large size of the counties, arguing that:
> "A key step in my method is to accurately measure the distance of observed
> manufacturing activity from the policy change border. The coarseness of the
> geographic information in the western states makes accurate measurements of
> distances relatively difficult to make" (p. 681).

[22] The appendix to this paper reports the probit model used to estimate the propensity scores. Predictions for different intervals were used to gauge neighborhood similarity.

The results indicate that the average propensity score is highest for the properties within one-half mile of Shannon Woods. They drop off sharply outside the 1.5-mile radius. Thus, these findings support our judgment that our study area (1.5-mile radius) is sufficient for capturing a similar set of treatment and control properties.

V.   Results

Table 4 reports our basic results for the repeat sales model. To implement the model, we follow Palmquist's [1980] proposal and adjust the price differential for depreciation based on a separate estimate of this effect.[23] The first four columns provide models within our primary definition of the study area (1.5 miles from the center of Shannon Woods). They differ in whether the specification accounts for the double selection effects with the relevant inverse Mills ratios, and whether the model includes the separate fixed effects for the first and second years of the sales involved in the repeat sales. Regardless of the specification selected, the term used to isolate the effect of the roadway bisecting the neighborhood, labeled *Shbrac* (*Shbrac* $= D_1 * D_2$, for properties in the Shannon Woods subdivision $D_1 = 1$, and for repeat sales that bracket the required interval $D_2 = 1$), is negative and significantly different from zero. The treatment effect ranges from –0.168 (with no fixed effects for the years of the sales or selection effect terms) to –0.216 (with both). In all cases, the treatment effect indicates that the new information about the highway significantly reduced the sale prices (in real terms). These conclusions use the Huber robust standard errors to reflect the heteroscedasticity

---

[23] The depreciation effect was estimated using a hedonic property value model for all residential sales in Wake County over the period 1965 to 1997 as a semi-log function with fixed effects for each year, age, square feet of living space, a lot size as independent variables. The estimated annual depreciation rate was

associated with our two-step approach to the selection terms. Using the Halvorsen-Palmquist-Kennedy interpretation of the treatment effect with a semi-log model implies a reduction in real prices from 15.5 to 19.5 percent using the 1.5 mile radius for the control neighborhood, or $38,196 to $48,092 in real loss.

Only one of the selection effects is significant using the 1.5-mile radius for controls. The estimated effect of the change in the distance to vacant land is a significant negative influence on the change in property values, indicating that households purchasing in this area appear to have a preference for lower density development. This interpretation follows because as the distance to open land increases over time the real property value differential is smaller and if it decreases, the real differential becomes larger.

The fixed effects for the years bracketing the sales take account of the time varying influences in the area, aside from the information about how the roadway affects the Shannon Woods subdivision. Because we have already used the CPI to measure housing prices in real terms, these terms are not price index measures as interpreted in Gatzlaff and Haurin [1997].

The estimated parameter for D_2 (the qualitative variable indicating that repeat sales bracket the release of the DEIS), indicates that the temporal dimension is not reflecting the treatment effect. The estimated coefficient for D_2 is only significant in our base sample when the fixed effects for the years of the sales are dropped. However, as we expand the control area from the 1.5 mile interval supported by the propensity scores

---

–0.2. Following Palmquist [1980] we adjusted the price series for depreciation before computing the log price differences.

to a 2-mile radius, D_2 is significant with or without the fixed effects.[24] In this case, D_2 is likely to be reflecting other changes in land use that are accompanying the development of the outer-loop roadway. These models are given in the fifth and sixth columns of Table 4. Both of the selection effects are also significant in these models, while only one is significant in our base models. The estimated treatment effect remains about the same size as in the base model.

To estimate the two selection effects we included properties that were never sold in the full area (i.e., the Shannon Woods subdivision along with the definition for the control neighborhoods) for the first selection terms as well as those that sold twice for the second. Table 5 reports our estimates of this sequential selection model (i.e., treating the first sale as a conditioning factor influencing the ability to observe a second sale). Because we use the information on the unsold properties, there are limits on the determinants we can consider. In general, structural attributes of the property and age are important determinants with effects that generally agree with our a priori expectations. Larger lots reduce the likelihood of an initial sale and smaller homes are more likely to sell twice. Based on the ability to distinguish different effects for property characteristics and lot size in each sale, it does appear that the effort to take account of the double sequential selection was warranted.

VI.   Implications

Natural experiments that conform to the regression discontinuity design can measure the appropriate compensation for a regulation or policy intervention that impacts

---

[24] The regression results in Table 4 reflect the base sample of properties within 1.5 miles of Shannon Woods. The last two columns also reflect the sample when the boundary is expanded to 2 miles. The results

land values. By adapting the Miceli-Segerson [1996] analysis, we found that for continuing land uses conditional compensation is warranted to assure that policymakers respond to the incentives for efficient choice. In particular, if their probability of taking an action is related to the amount of compensation paid and the likelihood of policy action with no compensation exceeds the efficient level, then a conditional compensation rule can be efficient.

The appropriate level of compensation corresponds to the difference between impacted properties' values without the policy and impacted properties' values with the policy (i.e., $V_R^1 - V_D^1$). Our review of the RD design criteria suggested that this difference can be difficult to isolate. It requires an unambiguous "information trail" that is rarely available in cases involving public policies that impact private land uses. Our specific application was able to uncover and use a clear record of the information available to the market to mimic the with/without conditions required for a compensation measure. Nonetheless, the process of developing our estimates identified an issue that has not been adequately treated in the literature on natural experiments. The focus of most of the discussion has been on defining the so-called treatment group. This process requires we assure that the behavioral response we are using to measure the treatment effect arises from a well-defined treatment group. To do so the analysis must specify both the treatment <u>and</u> the control groups. It is the specification of the control and its implications for the estimated impact of policy that has been underappreciated in the hedonic literature. We proposed that propensity scores used with matching estimators

for the 2.5 and 3 mile boundaries are available from the authors upon request.

can add information on the plausibility of alternative definitions for the control groups.[25]

In our application this strategy required that the temporal and geographic definitions for the treatment group be well defined by the policy intervention being evaluated, so that we could consider propensity scores to measure the similarity of properties unaffected by policy action in comparison to a well-defined set of homes that exogenous information allows us to hypothesize are impacted.

Finally, in the context of the specific application we considered, the impacts of highways on adjoining land uses, most of the attention in the past literature has been on the disamenity effects (see Langley [1976] for early example), such as noise or air pollution, and little specific evidence on the with/without change in property values. Most studies are done ex post and lack the ability to measure conditions that would meet the with/without standard. Fortunately, the available records for our application together with a complete history of property sales in the area allowed us to overcome these shortcomings in the past literature. For our application, compensation for losses due to the roadway in a neighborhood adjoining the right-of-way we found to be substantial, ranging from about 16 to 20 percent of the real property values.

---

[25] See Smith and Todd [2000] for further discussion and examples.

References

Angrist, Joshua D. and Victor Lavy, 1999, "Using Macmonides' Rule to Estimate the Effect of Class Size on Stochastic Achievement," *Quarterly Journal of Economics*, **114** (May): 533-576.

Black, Sandra E., 1999, "Do Better Schools Matter? Parental Valuation of Elementary Education," *Quarterly Journal of Economics*, **114** (May): 577-600.

Blume, Lawrence, Daniel L. Rubinfeld, and Perry Shapiro, 1984, "The Taking of Land: When Should Compensating Be Paid?," *Quarterly Journal of Economics*, **99** (February): 71-92.

Center on Urban and Metropolitan Policy, 2000, *Adding It Up: Growth Trends and Policies in North Carolina*, report prepared for the Z. Smith Reynolds Foundation, The Brookings Institution, July.

Congressional Budget Office, 1998, *Regulatory Takings and Proposals for Change*, December (http://www.cbo.gov).

Cropper, Maureen L., Leland B. Deck, and Kenneth E. McConnell, 1988, "On the Choice of Functional Form for Hedonic Price Functions," *Review of Economics and Statistics*, **70** (November): 668-675.

Gatzlaff, Dean H. and Donald R. Haurin, 1997, "Sample Selection Bias and Repeat Sales Index Estimates," *Journal of Real Estate Finance and Economics*, **14** (1): 33-50.

Hahn, Jinyong, Petra Todd, and Wilbert Van der Klaauw, 2001, "Identification and Estimation of Treatment Effects with a Regression Discontinuity Design," *Econometrica*, **19** (January): 201-210.

Halvorsen, Robert and Raymond B. Palmquist, 1980, "The Interpretation of Dummy Variables in Semi-logarithmic Equations," *American Economic Review*, **70** (3): 474-475.

Heckman, James J., 1979, "Sample Selection Bias as a Specification Error," *Econometrica*, **47** (January): 153-161.

Heckman, James J., 2001, "Micro Data, Heterogeneity and the Evaluation of Public Policy: Nobel Lecture," *Journal of Political Economy*, **109** (August): 673-748.

Holbrook, Michelle, 2000, "An Assessment of the Planning and Development of the Northern Wake Expressway," unpublished CEnREP discussion paper, August 29.

Holmes, Thomas J., 1998, "The Effect of State Policies on the Location of Manufacturing: Evidence from State Borders," *Journal of Political Economy*, **106**

(4): 667-705.

Kennedy, Peter E., 1981, "Estimation with Correctly Interpreted Dummy Variables in Semi-logarithmic Equations," *American Economic Review*, **71** (September): 801.

Kiel, Katherine A. and Katherine T. McClain, 1995, "The Effect of Incinerator Siting on Housing Appreciation Rates," *Journal of Urban Economics*, **37** (   ): 311-323.

Kolhase, Janet E., 1991, "The Impact of Toxic Waste Sites on Housing Values," *Journal of Urban Economics*, **30** (January): 1-26.

Langley, C. John Jr., 1976, "Adverse Impacts of the Washington Beltway on Residential Property Values," *Land Economics*, **52** (February): 54-65.

McClelland, Gary H., William D. Schulze, and Brian Hurd, 1990, "The Effect of Risk Beliefs on Property Values: A Case Study of a Hazardous Waste Site," *Risk Analysis*, **10** (4): 485-497.

Miceli, Thomas J. and Kathleen Segerson, 1986, *Compensation for Regulatory Takings: An Economic Analysis with Applications* (Greenwich, CT: JAI Press, Inc.).

Michaels, R. Gregory and V. Kerry Smith, 1990, "Market Segmentation and Valuing Amenities with Hedonic Models: The Case of Hazardous Waste Sites," *Journal of Urban Economics*, **28** (September): 223-242.

Palmquist, Raymond B., 1980, "Alternative Techniques for Developing Real Estate Price Indexes," *Review of Economics and Statistics*, **60** (August): 442-448.

Palmquist, Raymond B. and V. Kerry Smith, 2001, "The Use of Hedonic Property Value Techniques for Policy and Litigation," forthcoming, *International Yearbook of Environmental and Resource Economics*, **6** (2001/2002) (Cheltenhen, U.K.: Edward Elgar, in press).

Rosenbaum, Paul R. and Donald B. Rubin, 1983, "The Central Role of the Propensity Score in Observational Studies for Causal Effects," *Biometrika*, **70** (April): 41-55.

Smith, Jeffrey and Petra Todd, 2000, "Does Matching Overcome LaLonde's Critique of Non-experimental Estimators?," unpublished paper, University of Pennsylvania, October 6.

Smith, V. Kerry, Roger H. Von Haefen, and Wei Zhu, 1999, "Do Environmental Regulations Increase Construction Costs for Federal-Aid Highways? A Statistical Experiment," *Journal of Transportation and Statistics*, **2** (May): 45-60.

Tunali, Insan, 1986, "A General Structure for Models of Double Selection and an Application to a Joint Migration/Earnings Process with Remigration," in *Research*

in Labor Economics, **8**, Part B, edited by R.G. Ehrenberg (Greenwich, CT: JAI Press, Inc.): 235-282.

Walsh, Randy, 2000, "Analyzing Open Space Policies in a Locational Equilibrium Model with Endogenous Landscape Amenities," unpublished paper, Duke University, November.

Table 1: Number of residential sales and average sales prices in study area, by year and number of sales[a]

Year	Residential Properties that Sold Once During the Study Period		Residential Properties that Sold More Than Once During the Study Period	
	Average Sales Price	Number of Sales	Average Sales Price	Number of Sales
1985	185,625	72	168,080	131
1986	174,305	101	174,693	149
1987	179,312	119	172,004	141
1988[b]	--	--	--	--
1989	180,982	54	173,073	69
1990	195,845	116	180,110	113
1991	186,393	126	187,560	126
1992	194,099	142	191,202	94
1993	199,169	172	179,188	72
1994	213,797	150	193,083	78
1995	228,193	137	213,900	50
1996	237,989	180	197,722	45
1997	251,629	201	236,385	26
1998	244,682	192	245,357	7
Full Study Period	212,221	1,762	183,095	1,101[c]

[a] The geographic extent of the study area is a 1.5-mile radius of the center of the Shannon Woods development.
[b] Sales in 1988 are excluded from they period because the information homeowners had on the path of I-540 was uncertain during that year.
[c] One observation was dropped due to missing values for some of the independent variables in the model.

Table 2: Repeat Sales Sample Exclusion Criteria and Selection Model Samples, for Study Area and Concentric Circles Around Shannon Woods

	Study Area (1.5 Mi. Radius)	Rings around Shannon Woods		
		0-2.0 mile	0-2.5 mile	0-3.0 mile
Total number of sales in study area	6622	12359	20616	28655
Number of residential properties that sold at least once	2989	5499	9125	12619
Number of land sales[a]	670	1138	1766	2336
Number of sales before 1985 or after 1998	1683	3343	6054	8882
Number of sales between 9/1/1987 and 8/1/1989	755	1362	2021	2756
Number of sales of relocated properties	7	7	7	7
Number of properties with unidentified improvements	492	761	1260	1569
Outliers	105	198	331	458
Number of sales remaining	2917	5550	9177	12647
Number of repeat sales	1100	2156	3577	4978
Number of sales in Shannon Woods	42	42	42	42
Selection Model Samples, for Study Area and Concentric Circles Around Shannon Woods				
Number residential properties that did not sell	27	42	75	94
Number of properties with single sales (*SONCE*)	1762	3394	5599	7726
Number of properties with repeat sales (*STWICE*)	811	1552	2609	3597

[a] This includes: (1) sales price less than land value, (2) sale occurs before construction of structure, and (3) first of two sales if time between those two sales is less than one year and second sales price is more than twice the first sales price. Order is 1, missing (in next footnote), 2, and 3.

Table 3: Propensity Scores and Population Characteristics for Study Area (1.5-mi. radius) and Concentric Rings Around Shannon Woods

	Distance from the Center of Shannon Woods						
	Study Area (0 to 1.5 mi.)	0 to 0.5 mi.	0.5 to1.0 mi.	.0 to 1.5 mi.	1.5 to 2.0 mi.	2.0 to 2.5 mi.	2.5 to 3.0 mi.
Propensity Score (Std Dev)	--	0.188 (0.004)	0.003 (0.000)	0.022 (0.000)	4.79e-6 (0)	5.93e-10 (0)	1.47e-5 (0)
1989 Median Annual Household Income (Std Dev)	68,566 (14,780)	82,577	72,513	63,691	56,130	54,238	53,982
% of Population Completing (Std Dev)							
High School	12 (2)	12	12	13	14	14	14
4-yr. College	36 (3)	36	36	36	35	35	35
Grad. or Prof. School	18 (2)	19	19	18	15	14	14
Pct of Population (Std Dev) that is:							
White	91 (3)	87	89	92	91	88	88
Black	5 (6)	7	5	5	7	9	9
Native American	0 (0)	0	0	<1	<1	<1	<1
Asian/Pacific Islander	4 (4)	6	5	3	2	2	2
Other	0 (0)	<1	<1	<1	<1	<1	<1
Total Sq. footage of property	2624	2,733	2,879	2,901	2,600	2,351	2,898
Heated Sq. Footage of Property	2656	2,933	2,778	2,641	2,406	2,179	2,176
Number of rooms	8	8	8	8	7	6	7
Number of bathrooms	3	3	3	3	2.8	2.6	2.6
Number of fireplaces	1	1	1	1	1	0.9	0.9
Acreage of lot	0.7	1.8	1.7	1	2.5	43.6[a]	2.6

[a] This mean reflects a few very large private properties that are in this band (i.e., 2.0 to 2.5 miles from the center of the subdivision). It is not a coding or typing error.

Table 4: Repeat Sales Model – Impact of I-540 on Residential
             Subdivision[a]

Model	1.5 Mile Radius				2.0 Mile Radius	
	1	2	3	4	1	2
intcpt	-0.270	0.104	-0.233	0.112	-0.176	0.113
	(-3.59)	(4.83)	(-3.42)	(8.35)	(-3.15)	(6.81)
Ädvacant	-0.127	-0.160	-0.132	-0.126	-0.189	-0.206
	(-4.90)	(-5.53)	(-5.51)	(-5.06)	(-8.04)	(-7.98)
D_2	-0.085	-0.102	-0.125	-0.157	-0.083	-0.104
	(-1.52)	(-1.76)	(-7.08)	(-8.95)	(-1.76)	(-2.20)
shbrac	-0.216	-0.173	-0.211	-0.168	-0.219	-0.167
	(-6.77)	(-6.43)	(-6.05)	(-6.10)	(-7.30)	(-7.26)
$ë_1$	4.789	--	4.915	--	1.098	--
	(5.57)		(5.44)		(3.27)	
$ë_2$	-0.046	--	-0.069	--	0.245	--
	(-0.50)		(-0.78)		(3.57)	
Year 85	0.082	0.149	--	--	0.128	0.154
	(1.23)	(2.16)			(2.36)	(2.82)
Year 86	0.124	0.148	--	--	0.173	0.182
	(1.92)	(2.25)			(3.22)	(3.43)
Year 87	0.076	0.121	--	--	0.097	0.119
	(1.16)	(1.75)			(1.77)	(2.17)
Year 89	0.108	0.139	--	--	0.150	0.174
	(2.13)	(2.51)			(3.71)	(4.16)
Year 90	0.010	0.134	--	--	0.103	0.131
	(1.83)	(2.23)			(2.56)	(3.07)
Year 91	0.062	0.086	--	--	0.048	0.072
	(1.09)	(1.37)			(1.11)	(1.59)
Year 92	0.064	0.090	--	--	0.060	0.074
	(1.13)	(1.43)			(1.44)	(1.66)

[a] The numbers in parentheses are t-statistics.

Year 93	0.078 (1.32)	0.091 (1.38)	--	--	0.090 (2.02)	0.104 (2.18)
Year 94	0.057 (0.95)	0.047 (0.70)	--	--	0.086 (1.94)	0.080 (1.68)
Year 95	0.189 (3.02)	0.187 (2.73)	--	--	0.169 (3.42)	0.160 (3.02)
Year 96	0.148 (2.22)	0.128 (1.76)	--	--	0.169 (3.33)	0.158 (2.90)
Year 97	0.160 (2.42)	0.138 (1.92)	--	--	0.185 (3.49)	0.159 (2.86)
Year 98	0.075 (1.10)	0.024 (0.33)	--	--	0.114 (2.11)	0.053 (0.95)
R^2	0.210	0.131	0.179	0.093	0.188	0.132
Number of obs.	1092	1100	1092	1100	2137	2156

Table 5: Double, Sequential Selection Models for Base Case and 2.0-Mile Radius Neighborhoods[a]

	Base = 1.5 Mile Radius	2.0 Mile Radius
Sold Once		
Year Built	0.0020	0.0021
	(72.96)	(76.44)
Acreage	0.1800	0.0801
	(10.60)	(4.93)
Sq. Ft. Living Area	-0.0001	-0.003
	(-8.82)	(-21.11)
Intercept	-1.536	-1.156
	(-50.19)	(-36.65)
Sold Twice		
Acreage	-0.3370	-0.3489
	(-3.64)	(-5.33)
Sq. Ft. Living Area	0.025×10^{-3}	0.0614×10^{-3}
	(0.41)	(0.97)
Age	0.0470	0.0592
	(5.18)	(9.84)
Baths	0.0392	0.0838
	(0.86)	(2.33)
Fireplaces	--	0.075
		(0.82)
Intercept	-0.3627	-0.730
	(-2.40)	(-5.11)
Number of obs.	1756	3381
Number of censored obs.	27	42
Rho	-0.842[b]	-0.747[b]
	(0.00)	(0.00)

[a] The numbers in parentheses below the estimated coefficients for the probit with first stage selection are the ratios of the coefficients to their asymptotic standard errors.
[b] The number in parentheses below the estimate of Rho is the p-value for the chi-square test of equality with zero. It indicates highly significant estimates in this case.
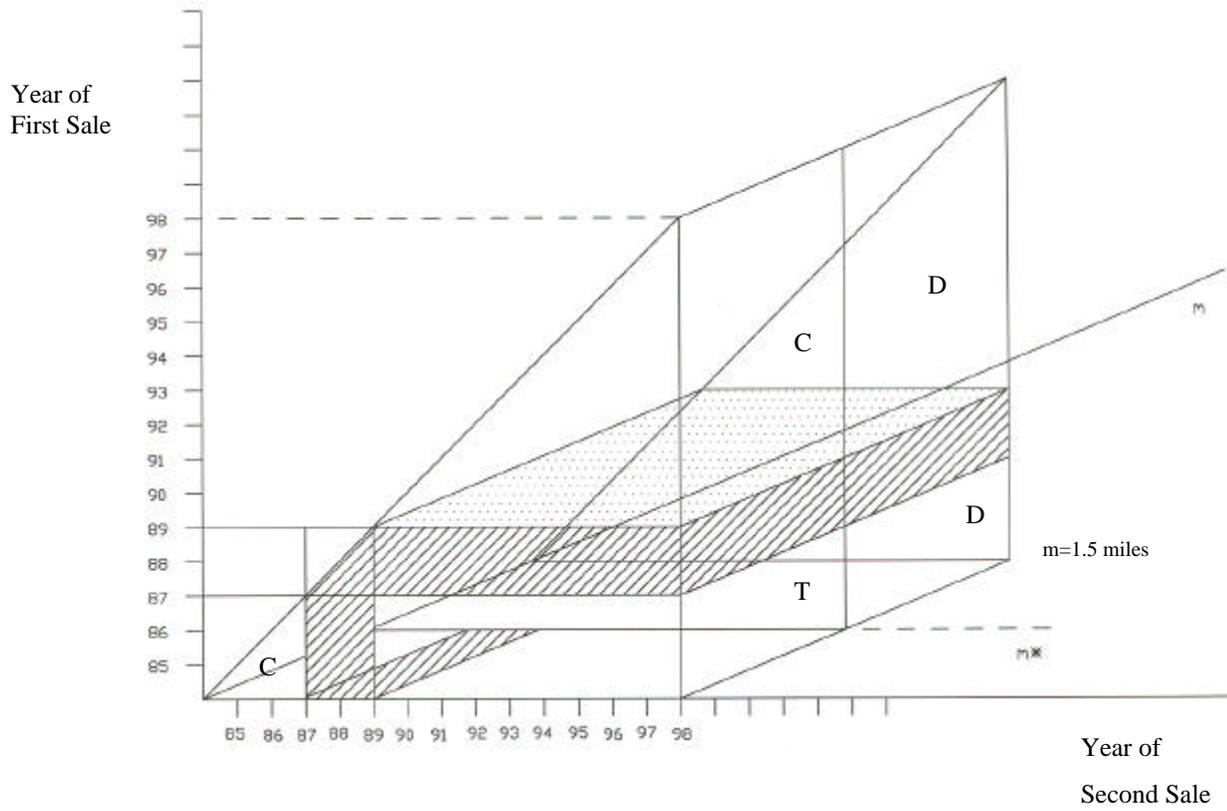
Figure 1

Figure 2

Bantry Court

Right-of-Way

Streets

Relocated
Houses

Shannon
Woods

Figure 3: Body of an October 31, 1984 Letter from the NCDOT to the Shannon Woods
Developer

"This is to confirm our telephone conversation of this date in which I advised that the Department of Transportation was not agreeable at this time to accepting the area reserved for a greenway or highway through the Shannon Woods Subdivision. This is due to the fact that the necessary environmental work has not been completed, as well as the fact that the 200-foot width is not sufficient to accommodate the facility proposed through this area. Also, no provisions have been made for control of access, and we note that an at-grade intersection is proposed for Chander Drive.

"It is our suggestion that the area be dedicated to the County if they are agreeable to accepting the same.

"Thank you for you interest and we regret that we cannot give you a more favorable response."

Appendix: Probit Model for Propensity Scores[a]

Independent Variables	Shannon Wood Sales (=1)
Intercept	-161.21
	(-4.24)
Square feet – living space	0.04×10^{-3}
	(0.23)
Travel distance (CBD)	2.55
	(4.53)
Distance to 540	-0.03×10^{-1}
	(-9.07)
Distance to golf course	0.41
	(5.08)
Year built	0.06
	(3.13)
Number of observations	38,486
Pseudo R^2	0.88

[a] The numbers in parenthesis are ratios of the estimated coefficient to the estimated asymptotic standard errors.