JOURNAL OF URBAN ECONOMICS 23, 307–326 (1988)

The Market Effects of Zoning Undeveloped Land: Does Zoning Follow the Market?*

NANCY E. WALLACE

Schools of Business Administration, University of California at Berkeley, Berkeley, California 94720

Received June 25, 1985; revised May 22, 1986

I

Through restricting the uses of land, zoning gives local governments the power to modify the dictates of unconstrained land markets. Many believe, however, that the regulations tend to coincide with, or anticipate, the market solution rather than modify it. The validity of this assertion is an empirical question and has not yet been adequately addressed. The purpose of this paper is to develop and apply an empirical test of the null hypothesis that zoning follows the market. The null hypothesis can be rejected only if land uses under zoning differ from what they would be without zoning.

The first task in deriving the test is to differentiate between constrained and unconstrained land market allocations. It is shown that parameter estimates for hedonic price functions, when markets are zoned, cannot be interpreted to imply that the regulations constrain market prices. It is also shown that sample selection bias is a likely source of specification error in estimating hedonic price functions for zoned land markets.[1] Since the sample selection of parcels to zone designations is unlikely to be random, the selection process must be accounted for in ordinary least-squares (OLS) estimation of zone constrained hedonic price functions. Fortunately, the "correction" for the specification error also provides a test for the market effects of zoning.

A bid-price interpretation of hedonic price functions for land is presented in Section II. The sample selection bias issues are also discussed in that section and a formal test for the market effects of zoning is introduced.

*Preparation of this article was undertaken while the author was an Assistant Professor at Washington State University and a Visiting Research Fellow in the Department of Agricultural and Resource Economics at the University of California at Berkeley. I thank Dan Rubinfeld, Jon Sonstelie, and John Quigley for their helpful comments on a draft of this article.

[1]Sample selection bias is also shown to be a probable source of misspecification bias in hedonic price analyses that are not explicitly accounting for zoning constraints or do not reflect zoned markets. This result suggests that corrections for sample selection bias should be accounted for in previous studies such as Wheaton [24, 26] and Polinsky and Shavell [19].

0094-1190/88 $3.00

The statistical model including "correction" terms for specification bias is developed in the next two sections. In Section V, the statistical procedure is applied to a data set drawn from the complete population of all 1978 arm's length transfers of vacant land in King Country, Washington.[2]

II

Following the multiproduct model of competition developed by Griliches [7] and Rosen [20], it is assumed that the commodity actually traded in an urban land market is a composite of location-specific attributes, such as access to roads, represented by a vector Z. The total unit payment made for land reflects the sum of transactions made in implicit attribute markets. Thus, land price is specified as a hedonic price function in which the total per unit payment for land V is a function of a vector of location-specific attributes Z,

$$V = V(Z). \tag{1}$$

Zoning fixes the supply of land with given attributes for different uses. For example, commercial allocations are often assigned to parcels with attributes such as location on an arterial intersection with sewer and water service. Since the existing or planned urban service infrastructure is fixed over long time horizons, it is reasonable to assume that the supply of land with any given bundle of attributes is fixed over the period the firms make land and attribute choices. If the allocation leads to a fixed supply of location-specific attributes for given land uses, the equilibrium attribute prices would be purely demand determined. Thus, the bids per unit of land in zoned markets would be determining the partial of the land price function in the process of generating equilibrium in the land market.

In unregulated land markets, competitive bidding and the potential for free market entrance assure that land attributes are purchased by the highest bidder [2, 25]. Thus, the outer envelope of firm bids for land exactly represents the hedonic price function at which demands balance existing supplies. A hypothetical equilibrium for two industries, residential housing producers ($j = 1$) and agricultural producers ($j = 2$), is portrayed in Fig. 1 and it is assumed that land price is a function of one attribute, distance to the city center.

As shown in Fig. 1, the bid-price function for each industry is tangent to the land market price function. Given different production technologies, we would expect the marginal bids and thus the gradient of the land price

[2] The same analysis was carried out for the 1975 parcels. The results of that analysis were quite similar; however, the 1978 sample contained large numbers of observations for each zone category.

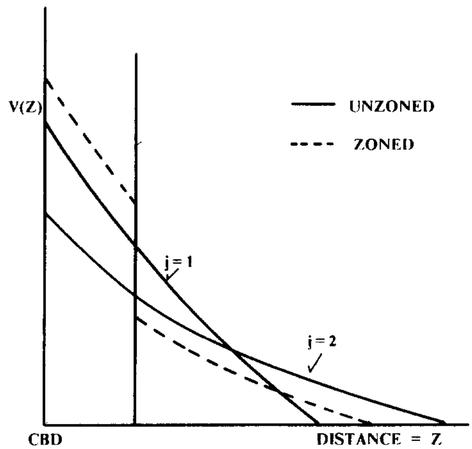

FIG. 1. Hypothetical equilibrium hedonic price functions for residential ($j = 1$) and agricultural ($j = 2$) uses.

function to differ between uses 1 and 2. If the hedonic price function for land was estimated econometrically and the estimated equation included slope or intercept dummies for industry, or use, we would expect to find statistically significant differences in the parameter estimates for those dummy variables. Similarly, we would expect statistically significant differences across equations if hedonic price functions were estimated separately for each industry. In this case, parameter estimates would reflect the portion of the exogenously determined hedonic price function that was tangent to the industry's bid-price function.

Now assume that the jurisdiction represented in Fig. 1 applies binding allowable use and minimum lot size constraints on 1 and 2. Under the constraints, the competitive bidding process leads to the discontinuous land price function shown. Clearly, the estimation technique above and a single cross-sectional data set, the predominant method in empirical zoning analyses (see [1, 3, 14, 16–18, 21, 22]) and hedonic price estimates that include zoning or similar policy variables [6], would lead to results that would be indistinguishable from those found in the unconstrained market. These results, therefore, could not be used to imply that zoning allocations had an effect on market prices. Instead, the issue that should be resolved in such analyses is whether the land use allocation would change if the constraints were abolished. That is, does the constrained land use allocation reflect highest marginal bids per unit of land or would the observed constrained use be outbid if zoning regulation was removed.

Rephrasing so that the question can be posed in a statistical framework, firm marginal bids for land attributes are observed by the econometrician as an attribute price on the hedonic price function. The hedonic price is observed with error,

$$V_{ij} = Z_i\gamma_j + u_{ij}, \tag{2}$$

where V_{ij} is the observed land price per unit for the ith parcel in use j and Z_i is a vector of land attributes. In the world of two land uses shown in Fig. 1, a parcel i with a a given distance characteristic Z_i will be observed in use 1, residential uses, if the maximum bid for the parcel by use 1 firms exceeds the maximum bid by use 2 firms, $V_{i1}(Z) > V_{i2}(Z)$. Thus, the observed distribution of parcel prices for use 1 is truncated at $V_{i1}(Z) > V_{i2}(Z)$ and the expected land price bid for use 1 parcels would be

$$E(V_{i1} | V_{i1}(Z) > V_{i2}(Z)) = Z_i\gamma_1 + E(u_{i1} | V_{i1}(Z) > V_{i2}(Z)). \tag{3}$$

Since the market sorts parcels according to their comparative advantage in given uses, the conditional expectation for the errors, u_{i1}, given the use 1 is highest and best would be positive. That is, the market would tend to reward positive misperceptions by bidders so that parcels sorted into use 1 would tend to have higher than average use 1 bids. In addition, higher than average bids in use 1 would also reflect attributes such as views or shade trees that were unaccounted for by the econometrician and yet make the parcels particularly suitable for residential uses.

If the conditional expectation for u_{i1} given the market allocation was not included in Eq. (3), the parameter estimates for γ_1 would be biased due to misspecification error unless the Z_i vector exactly reflected the attributes leading to the market allocation. If the Z_i vector does not precisely reflect the attributes determining the market allocation, the parameter estimates for γ_1 would proxy the effects of the omitted conditional expectation u_{i1}. Measuring a conditional expectation of this type requires prior information on the attributes that led to the market allocation and assumptions about the distributions of the errors in that allocation process.

Hedonic price functions in zoned markets are also observed with error. These errors are observed conditional on the community's zone allocation which is unlikely to be random. The allocation reflects rules that are at least partly specified by the zone ordinance or, in an increasing number of communities, by the comprehensive plan. The community zone allocation criterion, its preferences for zoning parcels to given uses, can be represented as function of attributes and land-use management objectives and an error term,

$$U_i{}^* = X_i\beta_j + e_i. \tag{4}$$

Since the $U_i{}^*$ is a preference ranking it is never observed, but it can be represented by an indicator variable I. In the case of the zoned world of Fig. 1, the indicator I is related to U_i^* as follows:

$$I = 1 \quad \text{iff } U_i^* > 0$$
$$I = 0 \quad \text{otherwise.}$$

Again, the distribution of zoned land prices for use 1 parcels is truncated since V_{i1} is never observed unless $U_i^* > 0$. Therefore, the expected zoned price bid for parcel i in use 1 would appear as

$$E(V_{i1}|I = 1) = Z_i\gamma_1 + E(u_{i1}|I = 1). \tag{5}$$

If the conditional expectation of the u_{i1} given the community zone allocation rule is not included, attributes in Eq. (4), which might not belong in the hedonic Eq. (2), might appear significant because they proxy the excluded expectation. The parameter estimates for γ_j would thus be biased due to the misspecification error. The conditional expectations in Eq. (5) can be estimated given suitable assumptions about the distribution of the zone choice errors and the errors of the observed zoned parcel price, conditional on the community zone allocation (see [4, 5, 10, 11, 15]).

If the estimate for the conditional expectation of u_{i1} given the community allocation rule is found to be positive, the market effects of zoning would be ambiguous since the expected truncation in an unconstrained market is also positive. However, if the estimate for the conditional mean of u_{i1} is negative and statistically significant we could conclude that zoning had a market effect. This would be true because an unregulated market allocation would never lead to negative truncation of the observed parcel price distributions. A negatively truncated distribution for a zoned land-use price implies that the observed prices for those parcels given their attributes on average fell below the price obtainable for the same attributes if they were randomly assigned to any other use. Negative truncation would also imply that the allocation would change if the zoning constraints were removed. In this way the correction for the misspecification bias arising from the nonrandom sample also provides a test for the market effects of zoning.

III

It is beyond the scope of this paper to develop a positive model of land-use planning at the local level. Thus, it will be assumed that the comprehensive plan may itself specify the zone allocation criteria for a community. This assumption is in keeping with the recent trend away from "planless zoning" in land-use planning law.

In the analysis area for this study, the urban county of Seattle, the "King County Comprehensive Plan 1964" establishes policies for differing land uses given the actual and forecasted supply of infrastructure in the urban area. These policies are then interpreted by the county council in its determination of the legally binding zoning ordinance. Thus, agricultural uses are to be assigned to parcels located at greater distances from urban and suburban areas and outside sewer districts, etc., whereas residential multiple uses must be assigned to parcels abutting intersections of arterials that are within sewer and water districts. Once the parcels are zoned, the market determines the price of land and the price can be represented by a hedonic price function.

Given the policies of the comprehensive plan, the county council has a preference ordering over location-specific attributes of land. Council utility can be presented by a function of observable land attributes and policies of the comprehensive plan and an error term representing the contribution of unobservable parcel attributes and legislative errors. The function is assumed linear and additively separable in its parameters,

$$U_{ij} = X_i\beta_j + V_{ij}\eta_j + e_{ij}. \tag{6}$$

The U_{ij} is the relative ranking attached by the King County Council to zoning parcel i to the jth zone alternative, X_i is a vector of exogenous parcel-specific attributes and policy objectives of the "King County Comprehensive Plan 1964," the β_j are the weights attached to these predetermined variables. V_{ij} is the market price if a parcel i were assigned to zone j and it reflects the supply constraints on land uses. The weight attached to price is η_j. The error term, e_{ij}, is a random variable that is assumed to be distributed Weibull. The specification of the function is derived from the comprehensive plan.

The market-determined hedonic price function, given the supply constraints on location-specific attributes, is the envelope of industry-specific bid-price functions. The hedonic price function for land is assumed to be exogenous to individual firms in each use. Thus, to recreate the information available to firms, the hedonic price function should be expressed as a function of location-specific attributes of the parcel designated to that use. The hedonic price function for use categories is

$$V_{ij} = Z_i\gamma_j + u_{ij}. \tag{7}$$

In Eq. (7), Z_i is vector of parcel-specific characteristics affecting the price of land per unit and a constant term. The Z_i vector also includes an institutional variable, whether the parcel was platted, that was excluded from the X_i vector.[3] Parcels can be platted only after they are zoned and they reflect

[3] If the vector of exogeneous variables X_i does not exclude at least one variable from the Z_i vector, the η_j parameter in the preference ordering is identified only by the nonlinearity of the correction terms introduced to account for the conditional covariance.

an important holding cost of the land. The other attributes are the same for X_i and Z_i. The u_{ij} are structural disturbances that are parcel specific.

Substituting Eq. (7) into Eq. (6) gives the reduced form for the utility function

$$U_{ij} = X_i\beta_j + Z_i\delta_j + \varphi_{ij}, \tag{8}$$

where

$$\delta_j = \gamma_j\eta_j$$
$$\varphi_{ij} = \eta_j u_{ij} + e_{ij}.$$

As previously discussed, zoned parcel price can be observed only if the jth zone designation is chosen. To account for this, let I be a polychotomous indicator variable taking values 1 to J, the number of zone designations, and $I = j$ if the jth zone designation is chosen. Thus,

$$I = j \quad \text{iff } \left[\left(X_i\beta_j + Z_i\delta_j\right) - \left(X_i\beta_k + Z_i\delta_k\right)\right] > \left[\varphi_{ik} - \varphi_{ij}\right] \tag{9}$$
$$\text{for all } k = 1, 2, \ldots, J. \quad k \neq j$$

The statistical model is a polychotomous-choice model with J-1 binary decision rules and follows the approach used by Hay [9] and Dubin and McFadden [4].

To simplify the notation, define $\omega_{ijk} = [\varphi_{ik} - \varphi_{ij}]$ and $Y_{ijk} = [(X_i\beta_j + Z_i\delta_j) - (X_i\beta_k + Z_i\delta_k)]$ so that condition (9) becomes $\omega_{ijk} < Y_{ijk}$. As discussed, OLS estimation of Eq. (7) given I would not be appropriate.

Assuming that the error terms ω_{ijk} follow the multivariate logistic distribution and the conditional expectation of the u_{ij} given ω_{ijk} is linear, the expected value of the zoned price, V_{ij}, given i is most preferred for zone j is

$$E\left(V_{ij} | \omega_{ijk} < Y_{ijk}\right) = Z_i\gamma_j + \sum_{\substack{k=1 \\ k\neq j}}^{J} \alpha_j E\left(\omega_{ijk} | \omega_{ijk} < Y_{ijk}\right), \tag{10}$$

where

$$\alpha_j = \operatorname{cov}(u_{ij}, \omega_{ijk})\left[\operatorname{var}(\omega_{ijk})\right]^{-1}.^4$$

[4] This formulation is a straightforward extension of results shown in Johnson and Kotz [13] and Maddala [15]. It also draws upon the additional restriction that the correlation between the u_{ij} and ω_{ijk} does not vary across j. Since each ω_{ijk} is the difference between i.i.d. Weibull error terms, it is not unreasonable to assume that they all have the same correlation with u_{ij}. Thus α_j reflects an average covariance.

Omission of the conditional expectation of u_{ij} given the zone selection rule would lead to misspecification bias if the conditional covariance, α_j, is nonzero. The covariance interpretation of α_j provides both a test for sample selection bias, a simple test of statistical significance, and an estimate for the sign and magnitude of the covariance of the u_{ij} and the ω_{ijk}, the vector of differenced error terms from the zone allocation decision. A statistically significant negative covariance estimate implies negative selection bias, or negative truncation, arising from the zone allocation. This finding indicates that zoning does not follow the market. A positive estimate for the α_j would lead to an ambiguous result, since positive truncation is also expected in unconstrained markets.

Parameter estimates for the sample correction regressors under a polychotomous choice can be obtained using techniques developed by Hay [9] and Dubin and McFadden [4].

$$E\left(V_{ij} \mid \omega_{ijk} < Y_{ijk}\right) = Z_i\gamma_j + \sum_{\substack{k=1 \\ k \neq j}}^{J} \left[\alpha_j^* \lambda_{ik}^j\right], \tag{11}$$

where

$$\lambda_{ik}^j = \sum_{\substack{k=1 \\ k \neq j}}^{J} \left[\sqrt{6}/\pi\right]\left[\frac{P_k \ln P_k}{1 - P_k} + \ln P_j\right] \qquad (j \neq k)$$

and

$$\alpha_j^* = \operatorname{cov}\left(u_{ij}, \omega_{ijk}\right).$$

The P_k and P_j terms are the estimated probabilities of observing the ith parcel in the kth or jth zone. The conditional logit model can be used to estimate the parameters of the zone choice that are needed to calculate the variables with coefficients α_j^* in Eq. (11). OLS estimates can then be obtained for the conditional expectation corrected equation written as

$$V_{ij} = Z_i\gamma_j + \sum_{\substack{k=1 \\ k \neq j}}^{J} \left[\alpha_j^* \lambda_{ik}^j\right] + \xi_{ij} \tag{12}$$

to get estimates of γ_j and α_j^*. The procedure is repeated over all zone designations $j = 1, \ldots, J$.

While OLS estimates of Eq. (12) would be consistent, the conditional variances of the error terms would be expected to be heteroscedastic since the correction factor is estimated. Additionally, as has been shown by Heckman [10, 11] and Maddala [15], the OLS standard error formula in this case is correct only if there is no selection bias. Consistent unbiased estimates can, however, be obtained from an asymptotic approximation of Eq. (12) using a Taylor's series expansion of $\hat{\lambda}^j_{ik}$ around the unobserved λ^j_{ik}. To simplify notation define the vector $\Psi = (\beta \;\; \delta)$, the vector of parameters in the reduced form Eq. (8), and define $\hat{\Psi}$ as a consistent estimate for this vector. The Taylor's expansion can then be written as,

$$\hat{\lambda}^j_{ik} - \lambda^j_{ik} = \sum_{\substack{k=1\\k\neq j}}^{J} \alpha^*_j \left[\hat{\lambda}^j_{ik}(\hat{\Psi}) + \partial\hat{\lambda}^j_{ik}/\partial\hat{\Psi}(\hat{\Psi} - \Psi)\right.$$
$$\left. + (\text{Higher Order Terms})\right]. \tag{13}$$

Reorganizing Eq. (13) and using asymptotic results, it is shown in Wallace [23] that OLS can be performed on

$$V_{ij} = Z_i\gamma_j + \sum_{\substack{k=1\\k\neq j}}^{J} \alpha^*_j \left[\hat{\lambda}^j_{ik}(\hat{\Psi}) + \left(\partial\hat{\lambda}^j_{ik}/\partial\hat{\Psi}\,\hat{\Psi}\right)\right]$$
$$+ \sum_{\substack{k=1\\k\neq j}}^{J} \alpha^*_j\left(\partial\hat{\lambda}_{ik}/\partial\hat{\Psi}\,\Psi\right) + \xi_{ij}, \tag{14}$$

where $\partial\hat{\lambda}^j_{ik}/\partial\hat{\Psi}$ is a function of the estimated probabilities.

The test for the market effects of zoning is a test on the sign and statistical significance of α^*_j, since the variance of the multivariate logistic distribution is a constant, $\pi^2/6$. The estimate for the bias appears alone on the first correction term $(\hat{\lambda}^j_{ik} + \partial\hat{\lambda}^j_{ik}/\partial\hat{\Psi}\,\hat{\Psi})\alpha^*_j$ but it is interacted with the unknown true choice equation parameters, Ψ_j, on the second term $(\partial\hat{\lambda}^j_{ik}/\partial\hat{\Psi}\,\Psi)\alpha^*_j$. Thus, the test for the direction of the conditional covariance, or selection bias, is a test on the parameter estimate for the first correction term of a α^*_j given the interaction effects on the second term.

The estimation of the zone choice and the market-determined price of zoned vacant land in King County is carried out in two stages. The first stage of the estimating procedure considers whether the King County Council exercises a consistent decision rule in the assignment of land parcels to zone designations. In particular, this first stage determines whether the observed zone choices for King County parcels reveal a consistent pattern of rank orderings based on the attributes of the land parcels and the objectives of the "King County Comprehensive Plan 1964."

The second stage consists of estimating the sample selection corrected hedonic price functions for zoned parcels. Sample selection biases can be tested on the null hypothesis of zero covariance between the error term of the zoned parcel price equation and the zone allocation rule, the α_j^*. In addition to the test for misspecification bias, the sign and magnitude of the estimates for α_j^* are an indication of the truncation of the land price distribution of given attributes when zoning allocates land to uses. A statistically significant negative bias effect, or truncation, implies that zoning does not follow the market.

IV

Even in the case where $X_i = Z_i$, the parameters β_j and η_j are identified because the correction terms are nonlinear functions of X_i and Z_i. In this case, as previously discussed, overidentifying restrictions on the parameters of Eq. (6) have been imposed by excluding elements of X_i from the Z_i vector. Eq. (7) is in reduced form, so that the structural parameters, γ_j, are identified.

The variables used in the estimation of the King County Council zone preference ordering include amenity, transportation, and environmental characteristics. These characteristics were specified by the "King County Comprehensive Plan 1964" as factors that determine whether land is used intensively or extensively. The amenity variables include:

NOSEWH = No sewer service and septic hazardous soils.

WATSP = Water service and soil percolates slowly.

FIRE = Rating (1 to 8, where 8 is the poorest service level) of availability of fire fighting equipment and water hydrants.

More intensive zone types would be expected to have both sewer and water service or be located on soils where wells or septic tanks could be used. Sewer, water, and fire service are considered predetermined variables because the State planning enabling legislation requires that land-use infrastructure be planned prior to the zoning ordinance. The fire rating variable is determined by the water service infrastructure and hydrant availability.

The transportation variables include:

CBDDIST = Actual road distance measured in kilometers from the Central Business District of Seattle.

UGCDIST = Actual road distance in kilometers from the nearest suburban center; these are smaller cities and towns in King County.

RDWIDTH = Width in feet of the abutting road.

LFSHLDER = Width in feet of the left shoulder (roads with no shoulders have curbs and sidewalks).

The comprehensive plan requires that more intensive uses locate closer to wider roads with curbs and sidewalks and closer to suburban centers. Residential land is expected to be located to facilitate commutes to Seattle or suburban centers, but on secondary or tertiary roads. Agricultural and forestry zones are to be located at greater distances from urban centers.

The road width and left shoulder width variables are predetermined in the council zone deliberations because road network design decisions are made by the Puget Sound Council of Government Transportation Plan and the Washington State Department of Transportation in accordance with traffic movement objectives for the county and state.

The environmental characteristics include:

DCORO = Dummy variables measuring the degree of soil corrosivity.

DSEISMIC = Dummy variable indicating whether the parcel is located on a fault area where seismically related slippage is a risk. This variable also measures slope.

Planning objectives in King County suggest that more intensive uses should be located on more corrosive soil, due to the expense of site preparation. Less intensive uses may be located in seismic hazard areas.

The variables for the price Eq. (7) are similar for the amenity and transportation variables; however, as discussed above, an additional regulation variable is included.

DPLAT = A dummy variable for whether the parcel has been platted.

Land parcels are typically not platted until they are zoned, because the zone determines the minimum allowable lot sizes. Many developers assert that a major component of land price is the cost of moving the plat application through the various levels of county bureaucracy. The variable is therefore included as an indication of the holding cost of the land.

Following established urban land rent theory [2], it is expected that parcel price would decrease with distance from the central city or suburban centers for residential uses. Parcel price would be expected to increase with distance for agricultural uses if proximity to urban uses presented externalities to farming production or for commercial/manufacturing uses if those uses were characterized by production technologies requiring low capital/land ratios. Parcels with amenities, curbs, or sidewalks would be expected to have higher parcel prices as would parcels with wider roads and more fire fighting services. More corrosive soils would be expected to reduce parcel price as would the presence of seismic hazards on a site.

King County zoning designations can be grouped into five identifiable use categories: (A) agricultural uses, (G) general uses, (R) residential family uses, (RM) residential multiple uses, and (CM) commercial/manufacturing

TABLE 1
1978 Data: Logit Estimates for Preference Ordering
(Dependent Variable: Zone Choice)

Variable	Zone cat.	Logit coefficient	Standard error	Variable	Zone cat.	Logit coefficient	Standard error
Constant	A	−21.5990*	2.1768	Constant	R	−2.2390	1.2000
DCORO	A	0.1317	0.6861	DCORO	R	−0.5338	0.3547
NOSEWH	A	1.2560*	0.6080	NOSEWH	R	−0.2944	0.3368
FIRE	A	1.9637*	0.2965	FIRE	R	0.7039*	0.1789
CBDDIST	A	0.2075*	0.0352	CBDDIST	R	−0.0938*	0.0200
UGCDIST	A	0.4034*	0.1072	UGCDIST	R	0.2650*	0.0732
RDWIDTH	A	−0.1507*	0.0358	RDWIDTH	R	−0.0034	0.0113
LFSHLDER	A	0.2376	0.1515	LFSHLDER	R	0.0559	0.0742
WATSP	A	−0.4811	0.5695	WATSP	R	1.4373*	0.3528
DSEISMIC	A	−0.5617	0.5913	DSEISMIC	R	−0.5609*	0.3450
DPLAT	A	0.5815	0.5894	DPLAT	R	2.1621*	0.3245
Constant	G	−4.7761*	1.1931	Constant	RM	−3.8142*	1.8010
DCORO	G	0.2273	0.3373	DCORO	RM	−0.2566	0.5188
NOSEWH	G	0.8116*	0.3202	NOSEWH	RM	−2.5761*	1.0981
FIRE	G	0.8256*	0.1750	FIRE	RM	0.9414*	0.2441
CBDDIST	G	0.0116	0.0194	CBDDIST	RM	−0.2291*	0.0292
UGCDIST	G	0.4419*	0.0709	UGCDIST	RM	0.2803*	0.1072
RDWIDTH	G	−0.0476*	0.0114	RDWIDTH	RM	0.0426*	0.0169
LFSHLDER	G	0.1621*	0.0682	LFSHLDER	RM	0.0253	0.1074
WATSP	G	−0.0250	0.3253	WATSP	RM	3.0824*	0.8664
DSEISMIC	G	−1.2113*	0.3359	DSEISMIC	RM	−2.5863*	0.7121
DPLAT	G	1.3565*	0.3216	DPLAT	RM	0.6448	0.4677

Auxiliary statistics	At convergence	At zero
Log likelihood	2024.68	3424.08
χ^2 40, 0.05	1399.40	
Percent correctly predicted	77.33%	20%
N	1606	

*Significant at 0.05 level.

uses. Each use category represents a predetermined percentage of total county land. The zone categories span a continuum of density and minimum lot size constraints; agricultural zones are least intensive and commercial zones are most intensive.

V

Table 1 gives the results of the logit estimation of the reduced form preference ordering, Eq. (8), for King County zoning decisions in 1978. From the χ^2 test, the null hypothesis that the council does not exercise consistent choice criteria in its zone allocations can be rejected at the 0.05 level of significance. Furthermore, the decision rule generally follows the policy guidelines of the "King County Comprehensive Plan 1964." For example, the odds of selecting agricultural or general zones increases if parcels are located farther from urban areas (UGCDIST and CBDDIST) or for sites with no sewers and hazardous soils (NOSEWH). The odds of assigning these zones falls as the road width (RDWIDTH) increases. On the other hand, the odds of residential and residential multiple zone assignments falls for parcels farther from urban areas or on sites without sewers or with septically hazardous soils. The odds of residential multiple zone assignments increases with wider roads.

Given these estimation results, it is reasonable to assume that the observed market price of undeveloped land parcels in King County is conditional on the preference orderings exercised by the King County Council. Since the Council appears to consistently translate the land-use planning objectives of the "King County Comprehensive Plan 1964" into zone allocation criteria, the market price of a zoned parcel is observed if and only if the council most prefers the parcel for its zone designation given its characteristics and the objectives of the plan. Thus, the conditional covariance terms in Eq. (14) cannot be assumed to be zero a priori.

Table 2 compares the use of OLS on samples stratified by zone which do and do not include conditional covariance correction terms. Estimation of the sample selectivity corrected price Eq. (14) was carried out using Hinkley's [12] heteroscedastic-robust jackknife technique to estimate the variance/covariance matrix. The corrected parameter estimates were obtained from a single OLS estimation of Eq. (14) using the zoned samples in blocked form.

The results from the uncorrected OLS estimation of the agricultural/forest zoned parcels are not very precise, and only the parameter estimate for RDWIDTH is both of the anticipated sign and significantly different from zero at the 0.05 level. The general zone parameter estimates are generally of the expected sign and are consistent with a priori expectations. The regression fit for the residential single family parcel price equation shows that distance to the nearest suburban center (UGDIST), road width (RDWIDTH), no sewer/septic hazards (NOSEWH), left shoulder width (LFSHLDER), and plat status (PLAT) all have significant effects and are of the correct sign. The results are less precise for the other parameter estimates for generally zoned parcels. The parameter estimates for residential multiple parcels show only DCORO, UGCDIST, and PLAT to have a significant effect on parcel price per square foot at the 0.05 level. All the parameter estimates for the commercial/manufacturing parcels were of the expected sign, except for the variables WATSP, RDWIDTH, and DSEISMIC.

TABLE 2
1978 Data: Comparison of Sample Selectivity Bias Corrected and Uncorrected Ordinary Least-Squares Estimates
(Dependent Variable: Parcel Price per Square Foot)

Corrected				Uncorrected			
Variable	Zone cat.	Parameter estimate	Standard error	Variable	Zone cat.	Parameter estimate	Standard error
Constant	A	0.4082	0.7466	Constant	A	0.0725	0.1105
DCORO	A	0.0323	0.0311	DCORO	A	−0.0361	0.0232
NOSEWH	A	0.0969	0.0566	NOSEWH	A	−0.0116	0.0386
FIRE	A	−0.0761	0.0701	FIRE	A	−0.175	0.0207
CBDDIST	A	0.0049	0.0052	CBDDIST	A	0.0021	0.0016
UGCDIST	A	0.0025	0.0060	UGCDIST	A	−0.0008	0.0035
RDWIDTH	A	−0.0014	0.0029	RDWIDTH	A	0.0045*	0.0016
LFSHLDER	A	0.0020	0.0074	LFSHLDER	A	−0.0031	0.0059
WATSP	A	−0.0516	0.0671	WATSP	A	−0.0051	0.0248
DSEISMIC	A	0.0284	0.0263	DSEISMIC	A	0.0028	0.0247
DPLAT	A	0.1366*	0.0453	DPLAT	A	0.0402	0.0265
CONCOV1	A	−0.0232	0.0441	R^2 =			0.2684
CONCOV2	A	−0.0062	0.0177	F ratio =			1.94
				N =			64
Constant	G	0.1706	0.1715	Constant	G	0.4038*	0.1381
DCORO	G	−0.212	0.0378	DCORO	G	−0.0682	0.0460
NOSEWH	G	0.0559*	0.0259	NOSEWH	G	−0.2226*	0.0291
FIRE	G	0.0010	0.0156	FIRE	G	0.0142	0.0173
CBDDIST	G	−0.0005	0.0018	CBDDIST	G	−0.0058*	0.0019
UGCDIST	G	−0.0098	0.0066	UGCDIST	G	−0.0126*	0.0055
RDWIDTH	G	0.0051*	0.0019	RDWIDTH	G	0.0119*	0.0019
LFSHLDER	G	−0.0046	0.0062	LFSHLDER	G	−0.0312*	0.0074
WATSP	G	0.0459	0.0244	WATSP	G	0.0255	0.0313
DSEISMIC	G	−0.0342	0.0353	DSEISMIC	G	0.0033	0.0404
DPLAT	G	0.1088*	0.0224	DPLAT	G	0.3399*	0.1381
CONCOV1	G	−0.1245*	0.0491	R^2 =			0.2470
CONCOV2	G	−0.0063	0.0100	F ratio =			30.97
				N =			955
Constant	R	2.512*	0.8867	Constant	R	0.6056	0.3194
DCORO	R	−0.4828*	0.1571	DCORO	R	0.1632	0.1298
NOSEWH	R	−0.6314*	0.1245	NOSEWH	R	−0.2335*	0.1075
FIRE	R	0.0216	0.0815	FIRE	R	−0.0870	0.0491
CBDDIST	R	−0.0170	0.0112	CBDDIST	R	−0.0094	0.0070
UGCDIST	R	−0.0110	0.0230	UGCDIST	R	−0.0793*	0.0191
RDWIDTH	R	−0.0071	0.0062	RDWIDTH	R	0.0292*	0.0038
LFSHLDER	R	0.0254	0.0263	LFSHLDER	R	−0.0421*	0.0187
WATSP	R	0.0589	0.1668	WATSP	R	0.5215*	0.1332
DSEISMIC	R	0.4722*	0.1547	DSEISMIC	R	−0.0773	0.1074
DPLAT	R	0.2556	0.1636	DPLAT	R	0.7238*	0.0949
CONCOV1	R	0.1502*	0.0751	R^2 =			0.3618
CONCOV2	R	0.0142	0.0224	F ratio =			24.72
				N =			447

TABLE 2–*Continued*

Corrected				Uncorrected			
Variable	Zone cat.	Parameter estimate	Standard error	Variable	Zone cat.	Parameter estimate	Standard error
Constant	RM	−3.2719	2.2066	Constant	RM	−0.5063	0.1245
DCORO	RM	0.0061	0.3166	DCORO	RM	0.7450*	0.3553
NOSEWH	RM	−4.4360*	1.1925	NOSEWH	RM	−0.0490	0.8806
FIRE	RM	0.2296	0.2494	FIRE	RM	−0.2234	0.2018
CBDDIST	RM	−0.2730*	0.0770	CBDDIST	RM	0.0032	0.0196
UGCDIST	RM	0.3963*	0.1223	UGCDIST	RM	0.2977*	0.1060
RDWIDTH	RM	0.0008	0.0228	RDWIDTH	RM	0.0204	0.0121
LFSHLDER	RM	−0.0641	0.1145	LFSHLDER	RM	−0.0699	0.0915
WATSP	RM	5.1538*	1.4966	WATSP	RM	0.5902	0.6446
DSEISMIC	RM	−2.2497*	0.6465	DSEISMIC	RM	0.0882	0.5347
DPLAT	RM	−2.6775*	0.9039	DPLAT	RM	1.3764*	0.4807
CONCOV1	RM	−1.4752*	0.4327	R^2 =			0.6904
CONCOV2	RM	0.2209*	0.0626	F ratio =			11.37
				N =			62
Constant	CM	−3.0793*	1.4682	Constant	CM	4.4047*	0.9849
DCORO	CM	0.0846	0.2969	DCORO	CM	0.3861	0.2923
NOSEWH	CM	−0.8207*	0.2875	NOSEWH	CM	−0.1138	0.3554
FIRE	CM	0.2661	0.2278	FIRE	CM	−0.2841*	0.1227
CBDDIST	CM	0.0942*	0.0265	CBDDIST	CM	−0.0395*	0.0191
UGCDIST	CM	0.1544	0.1166	UGCDIST	CM	−0.0532	0.0739
RDWIDTH	CM	−0.0864*	0.0322	RDWIDTH	CM	0.0116	0.0120
LFSHLDER	CM	0.0927	0.0686	LFSHLDER	CM	0.1271*	0.0489
WATSP	CM	0.6963*	0.2891	WATSP	CM	−0.2263	0.2864
DSEISMIC	CM	−0.5002	0.3231	DSEISMIC	CM	0.2692	0.2789
DPLAT	CM	0.6258	0.5008	DPLAT	CM	−0.1315	0.2670
CONCOV1	CM	−0.7394*	0.2575	R^2 =			0.32
CONCOV2	CM	0.1581*	0.0627	F ratio =			3.08
R^2 =			0.66	N =			78
F ratio =			47.3				
N =			1606				

*Significant at 0.05 level.

The most important differences between the corrected and uncorrected parcel price equation estimates are the changes in magnitude and precision for the corrected coefficient estimates on the no sewer/septic hazard variable (NOSEWH), the distance to downtown Seattle and suburban centers (CBDDIST and UGCDIST), and the road width and left shoulder width measures (RDWIDTH and LFSHLDER). These variables were also significant components in the logit estimation of the probabilities of a given zone type. It appears, therefore, that exclusion of the sample selectivity correction factor does confound the structural parameters of the price equation with those of the zone preference ordering.

As a further check on the estimation results, a Hausman test [8] was performed on the null hypothesis of no specification bias comparing the OLS with the conditional expectation corrected estimates. The estimated value of the test statistic was 124.85. Since the asymptotic distribution of the test statistic is χ^2 with a critical value of 61.75 at the 0.05 level, this result is further evidence of misspecification of the OLS estimation of zoned subsamples.

As discussed under Sections II and III, the conditional expectation corrected parameter estimates reported in Table 2 represent unbiased estimates of the true attribute prices of zoned land under random assignment. Without a test on the estimated conditional expectations of the hedonic error structures, however, it would not be possible to conclude that zoning per se affects the market price of land. It would be expected that land in different uses would lead to different attribute prices under a market allocation.

The test for the market effects of zoning is a significance test on the conditional expectation correction term CONCOV1. As previously discussed, this is a test for the magnitude and sign of the selectivity bias introduced by the council allocation of parcels to zone designations. As shown in Table 2, the coefficient estimates for the selectivity variable CONCOV1 are precise for all the zone categories except agriculture. The estimate is positive and significant for the residential single family zone and negative and significant for the general, residential multiple, and commercial zones. The interaction terms on CONCOV2 are positive and significant at the 5% level for the commercial and residential multiple zones, so the conditional expectations of the error terms are difficult to interpret.

The positive and statistically significant estimate on CONCOV1 for residential zoned land is an ambiguous result because a market allocation would also lead to such effects. A χ^2 test on the null hypothesis that the CONCOV1 and CONCOV2 terms were jointly zero for agriculture zoned land was accepted at the 0.05 level. The negative bias effects can be assumed for general zones since CONCOV1 is negative and statistically significant and CONCOV2 is also negative, though not statistically significant. The general zone designations require a minimum lot size from 35,000 to 217,800 ft^2 for residential uses.

To interpret these results for the general zone, consider a sample of parcels with characteristics similar to those zoned general. Negative selection bias implies that the average price per square foot for parcels of given characteristics that were zoned general is less than the price per square foot that would have been observed for parcels with the same characteristics had they been assigned at random to other zones. This result would never be expected from a market allocation. The test on the conditional expectation, in the case of general large lot zones, can be interpreted as strong evidence

TABLE 3
Predicted Number of Parcels that Would Have Attained Their Maximum Price under Alternative Zone Designations

	Predicted number that attain maximum price under alternative zone designations				
Actual zone for parcels:	A	G	R	RM	C/M
Agricultural/forest zone ($N = 64$)	0	0	0	0	64 (100%)
General zone ($N = 955$)	0	0	7 (0.7%)	0	948 (99.3%)
Residential single Family zone ($N = 447$)	0	0	3 (0.7%)	1 (0.2%)	443 (99.1%)
Residential multiple zone ($N = 62$)	0	0	0	0	62 (100%)
Commercial/manufacturing zone ($N = 78$)	0	0	26 (33%)	0	52 (67%)
Totals	0	0	36	0	1569

that large lot zoning does have market effect. This result indicates that on average, holding all other parcel attributes constant, King County zone allocations tend to decrease parcel price for large lot zones. This would imply that the legislative decision rule leads to an oversupply of land zoned to large minimum lot sizes.

VI

A marginal analysis was carried out to determine, in an informal sense, what the price of a parcel with known characteristics would have been if it had been zoned differently. The corrected price equations were also used to determine the zone category that would allow the parcels in the sample to attain their highest valued use. As discussed under Section III, we would not expect to find the observed market price for a given land use to be less than the price the same parcels could attain in other land uses. If the observed market prices for parcels of given zone designations never attain their highest value zoned as they are, it is quite likely that zoning regulations do not follow market allocations. However, no formal test for this conclusion will be attempted.

The results in Table 3 indicate that parcels actually zoned to general or agricultural uses would not be expected to attain their highest value when zoned to those uses, given their parcel characteristics. This can be interpreted as further evidence of the negative selection bias introduced by the King County zoning allocation for large lot zoning. Residential single

TABLE 4
Comparison of Actual Price Per Square Foot and Predicted Average Price Per Square Foot across All Zone Descriptions

	Actual mean priced as zoned	Predicted mean under alternative zone designations				
		A	G	R	RM	C/M
Agricultural/forest (N = 64)						
Predicted mean	\$.10	\$.19	\$.33	\$1.16	−\$15.40	\$6.74
Standard deviation	\$.08	\$.08	\$.42	\$.42	\$4.70	\$1.27
General (N = 955)						
Predicted mean	\$.40	\$.23	\$.30	\$1.70	−\$9.60	\$5.59
Standard deviation	\$.48	\$.11	\$.12	\$.41	\$4.60	\$1.61
Residential (N = 447)						
Predicted mean	\$1.03	\$.20	\$.43	\$2.04	−\$4.98	\$6.07
Standard deviation	\$1.02	\$.10	\$.10	\$.48	\$3.11	\$1.80
Residential mult. (N = 62)						
Predicted mean	\$2.02	\$.11	\$ .55	\$1.92	\$1.21	\$6.91
Standard deviation	\$.30	\$.27	\$.16	\$.46	\$.61	\$2.25
Commercial/manuf. (N = 78)						
Predicted mean	\$1.32	\$.43	\$.32	\$1.06	−\$12.29	\$1.37
Standard deviation	\$.98	\$.13	\$.11	\$.42	\$5.41	\$.53

family uses attained their highest value zoned either as they were or when assigned to residential multiple or commercial/manufacturing uses. Although the estimated results for commercial/manufacturing uses are difficult to interpret, the results of Table 3 indicate that most land in the sample would achieve its highest predicted value if zoned to commercial/manufacturing. This implies that the positive interactive effect of the choice parameter and the conditional covariance appears to outweigh the negative pure conditional covariance effect.

Table 4 compares the actual price of land zoned as it was and the average predicted price if the parcels were zoned otherwise at the margin. The corrected estimates fall below the observed mean parcel price per square foot for the general, agricultural, and residential single family and overestimate them for the residential and commercial/manufacturing zones. The magnitude of the effects of large lot zoning are substantial. The general zoned parcels would achieve a predicted square foot of \$.30 zoned as they are and \$5.59 if they were zoned commercial/manufacturing. The negative valued mean estimates for residential multiple zones reflects the distance sensitivities of these estimates. Since most agricultural and general land is located at greater distances from urban centers and amenities, these parcels would have lower value if they were zoned to apartment uses. The corrected hedonic estimates for commercial use, on the contrary, rise with greater distance from the center of Seattle and the suburban centers.

These informal analyses suggest that King County zone designations serve to decrease the average land value of general zones. This result would not be expected as the result of a market allocation, implying that there is an oversupply of large lot zones in the urban fringe of King County. The results imply that residential and commercial/manufacturing zones are probably undersupplied.

VII

The preceding results are promising in several respects. First, they indicate that hedonic price functions estimated for zoned land markets are likely to be misspecified unless "correction" terms accounting for the nonrandom sampling framework are included. These terms, or instruments, reflect the conditional covariance between the zone allocation rule and the observed market price and are computationally tractable even for polychotomous zone choices. A second advantage of the technique is that the instrument for the conditional covariance provides a test for the null hypothesis that zoning follows the market. Finally, the econometric specification is compatible with both the theory of hedonic price functions for land and complementary bid-price interpretation of a competitive land market equilibrium.

These findings are particularly important because most previous analyses of the market effects of zoning have used uncorrected OLS regression techniques to estimate the hedonic price function parameters for zoned land markets. They also strongly suggest that the predominant statistical test for the market effects of zoning is not compatible with the theory of land market prices in equilibrium and that sample selection bias may lead to misspecification errors in many types of hedonic analyses of urban land markets.

The estimation results for zoned land in King County indicate that there is misspecification bias in uncorrected OLS regression estimates of hedonic price function parameters. The location-specific attributes most affected by the bias were those that were most significant in the council zone decision. The null hypothesis that zoning followed the market was accepted for all but the general zone, large minimum lot size, designation. The conclusion drawn from the statistically significant and negative covariance instrument was that large lot zones were probably oversupplied in King County. The informal marginal analysis suggests that commercial and manufacturing zones allocations were probably undersupplied.

REFERENCES

1. F. G. Adams, G. Milgram, E. Green, and C. Mansfield, Undeveloped land prices during urbanization: A micro-empirical study over time, *Rev. Economo. Statist.*, **50**, 248–258 (1968).

2. W. Alonso, "Location and Land Use", Harvard Univ. Press, Cambridge, MA (1964).
3. M. E. Avrin, Some economic effects of residential zoning in San Francisco, in "Residential Location and Urban Housing Markets" (G. Ingram, Ed.), Ballenger, Cambridge (1977).
4. J. Dubin and D. McFadden, An econometric analysis of residential electric appliance holdings and consumption, *Econometrica*, **52**, 345–362 (1984).
5. G. M. Duncan, Formulation and statistical analysis of the mixed continuous/discrete dependent variable model in classical production theory, *Econometrica*, **48**, 838–852 (1980).
6. D. M. Grether and P. Mieszkowski, Determinants of real estate values, *J. Urban Econom.*, **1**, 127–146 (1974).
7. Z. Griliches, Hedonic price indexes revisited: Some notes on the state of the art, in "Price Indexes and Quality Change" (Z. Griliches, Ed.) Harvard Univ. Press, Cambridge, MA (1971).
8. J. A. Hausman, Specification tests in econometrics, *Econometrica*, **46**, 1251–1271 (1978).
9. J. Hay, "Occupational Choice and Occupational Earning: Selectivity Bias in Logit OCS Model," Unpublished Ph.D. dissertation, Yale University, New Haven, CT (1980).
10. J. Heckman, The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models, *An. Econom. Soc. Meas.*, **514**, 475–492 (1976).
11. J. Heckman, Sample selection bias as a specification error, *Econometrica*, **47**, 153–162 (1979).
12. D. Hinkley, Jackknifing in unbalanced situations, *Technometrics*, **19**, 285–292 (1977).
13. N. L. Johnson and S. Kotz, "Distributions in Statistics and Continuous Multivariate Distributions," Vol. 3, Wiley, New York (1970).
14. G. D. Jud, The effects of zoning on single family residential property values: Charlotte, North Carolina, *Land Econom.*, **56**(2), 142–54 (1980).
15. G. S. Maddala, "Limited-Dependent and Qualitative Variables in Econometrics," Cambridge Univ. Press, Cambridge (1983).
16. S. M. Maser, W. Riker, and R. N. Rosett, The effects of zoning and externalities on the prices of land in Monroe County, New York, *J. Law Econom.*, **17**, 111–132 (1974).
17. L. L. Orr, "Income, Employment and Urban Residential Location," Academic Press, New York (1975).
18. G. E. Peterson, "Land Prices and Factor Substitutes in the Metropolitan Housing Market," Paper for 1974 COUPE Meeting, Resources for the Future, Inc., Nov. 8–9 (1974).
19. A. M. Polinsky and S. Shavell, Amenities and property values in a model of an urban area, *J. Publ. Econom.*, **5**, 119–129 (1976).
20. S. Rosen, Hedonic prices and implied markets, *J. Political Econom.*, **82**, 34–55 (1974).
21. L. B. Sagalyn and G. Sternlieb, "Zoning and Housing Costs," Rutgers University Center for Urban Policy, New Brunswick (1973).
22. F. Vaillancourt and L. Monty, The effect of agricultural zoning on land prices, Quebec, 1975–1981, *Land Econom.*, **61**, 36–42 (1985).
23. N. E. Wallace, "The Market Effects of Zoning and the Consistency Doctrine in King County, Washington," Unpublished Ph.D. dissertation, The University of Michigan (1982).
24. W. Wheaton, A bid rent approach to housing demand, *J. Urban Econom.*, **4**, 200–217 (1977).
25. W. Wheaton, Linear programming and locational equilibrium, *J. Urban Econom.*, **1**, 278–287 (1974).
26. W. Wheaton, Income and urban residence: An analysis of consumer demand for location, *Amer. Econom. Rev.*, **67**(4), 620–631 (1977).