Endogenous Open Space Amenities in a Locational Equilibrium

Randall P. Walsh[1]
University of Colorado

Comments Welcome.

February, 2004

[1]The author thanks Helen Ladd, Tom Nechyba, Holger Sieg, and especially V. Kerry Smith for their help and support. The author also thanks H. Spencer Banzhaf, Mark Coppejans and Roger Von Hafen for their help and comments. Dissertation support from Resources For the Future and the Lincoln Institute of Land Policy is gratefully acknowledged. The views presented in the paper

Abstract

Little is known about the equilibrium impact of open space protection and growth control policies on the entire metropolitan landscape. This paper is an initial attempt to evaluate open space policies using an empirical approach that incorporates the endogeneity of both privately held open space and land conversion decisions in a locational equilibrium framework. The analysis yields four striking results. First, when one allows for endogenous adjustments in privately held open space, increasing the quantity of land in public preserves may actually lead to a decrease in the total quantity of open space in a metropolitan area. Second, different strategies for spending the same amount of money to purchase open space have markedly different welfare implications. Third, partial equilibrium welfare calculations are extremely poor predictors of their general equilibrium counterparts. And finally, the analysis suggests that while a growth ring strategy is most effective in reducing total developed acreage in the metropolitan area, this reduction in developed acreage is associated with a large net welfare loss.

In addition to its policy relevance, The paper makes two methodological contributions to the locational equilibrium literature. First, the analysis considers a Nash equilibrium with endogenous public goods where these goods arise ‘naturally’ as a result of land market

land protection policies and growth restrictions influence their impacts. First, open space amenities are inherently spatial. An acre of land protected at location A is not equal to an acre of land protected at location B. Second, non-marginal land protection policies will directly impact the land market equilibrium – leading households to make different location and lot size choices. Third, to evaluate how land protection policies affect the market equilibrium it is necessary to take account of the role of differences in the suitability of different locations for development. Finally, households' adjustments in location and lot size in response to land protection policies create new patterns of development implying that some open space amenities will be endogenous.

This paper develops a general equilibrium (GE) residential land market model. Household preferences are estimated using an extension of the empirical locational equilibrium model initially proposed by Epple and Sieg (1999). Differentiation in the suitability for development is incorporated using a spatially explicit land supply model developed in Walsh (2004). These two components are combined to generate the land market equilibrium model.[2] Combining these two empirical components into a single computational model makes it possible to evaluate spatially delineated open space policies.

The remainder of the paper is organized into six sections. section two develops a theoretical model of land market equilibrium with endogenous landscape amenities. Section three presents the implementation strategy for the locational equilibrium model. Section four describes the data used in the analysis. Section five presents estimation results. Section six develops the policy simulations and presents general equilibrium computations for the policy experiments. Finally, section seven summarizes the conclusions.

[2] Sieg, Smith, Banzhaf, and Walsh (2003) adapt the Epple-Sieg (1999) model to a G.E. framework based on constant elasticity housing supply functions. This work extends their methodology by endogenizing the location specific amenities (open space), allowing for more complex substitution patterns between public goods by introducing augmented prices, and developing a more detailed description of the supply side of the model.

2 The Link Between Household Choices and Open Space

This section considers a selection of the empirical literature on open space amenities and outlines a formal model for household preferences to reflect some of the features identified in that literature.

2.1 Background

Most studies dealing with open space amenities use hedonic housing price models[3] and can not be used to evaluate non-marginal policies.[4] Weicher and Zerbst (1973) present one of the earliest examples, considering amenities from neighborhood parks in Columbus, Ohio that are assumed to be captured by a set of dummy variables describing immediate adjacency to protected open space. They find positive price effects only for houses which directly face protected open space. For Boulder Colorado, Correll, Lillydahl, and Singell (1978) suggest that open space is both a public good which benefits everyone in the Boulder area and a 'quasi-public good' due to distance based exclusion of some protected parcels. They find that distance to open space as an indicator of a reduction in the 'quasi-public' aspect of open space leads to a $4.20 reduction in expected housing price. Recent analysis by Geoghegan,

[3]Contingent valuation methods have also been used to study consumer preferences over open space amenities. For example Halstead (1984) provides estimates of Massachusets' resident WTP for protecting agricultural land and Ready, Berger, and Blomquist (1997) provide measures of the willingness to pay of Kentucky residents for a program that would provide incentives for horse breeders to locate in the state. These studies recover consumers' willingness to pay for a movement from one state of the world to another. Successful policy analysis therefore requires a complete understanding of the price and amenity levels that will be available after the policy is implemented. Thus, the methodology is of limited use in this application.

[4]There are three qualifications to this point. First, Palmquist (1992) demonstrates that in the case of localized externalities (for example highway noise in a single neighborhood), reduced form hedonic models are sufficient for welfare measurement of non-marginal changes. Second, Bartik (1988) provides a methodology for identifying bounds on the WTP of non-marginal changes based on hedonic models. Finally, in recent work, Ekeland, Heckman, and Nesheim (2002) demonstrate that in general, even in a single market, the hedonic model is non-parametrically identified. They propose an approach which makes it possible to recover technology and preferences in a separable version of the hedonic model. They further argue that identification strategies incorporating cross-market data are based on economically implausible assumptions about why hedonic functions vary across markets.

Wainger, and Bockstael (1997) confirms these effects for landscape amenities.[5]

2.2 Modelling Open Space

The analysis begins by marrying a model of household preferences for residential lots and open space with a spatially delineated static representation of lot conversion decisions to yield an equilibrium model of land markets incorporating the endogenous determination of privately held open space. Household preferences over spatially delineated neighborhoods incorporate heterogeneity in income and tastes and consider two distinct types of open space amenities. The decision to develop individual lots, based on prices and lot characteristics, are aggregated to yield neighborhood specific residential land supply functions.

2.2.1 Household Preferences

Preferences are defined over two distinct measures of open space, O^p, a measure of the distance from a given lot location to the nearest protected parcel of open space, and O^n, a measure of the percentage of a given lot's neighborhood which is in open space (both protected and unprotected). For each lot location, O^p is assumed to be determined as the result of exogenous land protection policies.[6] O^n on the other hand is endogenous to the model and arises from the aggregation of development decisions and the exogenously determined land protection policies.[7]

[5]The studies cited here only scratch the surface in terms of the valuation of open space. See for instance Lee and Linneman (1998), Li and Brown (1980) and Greenwood and Hunt (1989) on the valuation of public parks; Halstead (1984), Ready et al. (1997) , Kline and Wichelns (1994), Bergstrom, Dillman, and Stoll (1985), Bolitzer and Netusil (2000) and Mahan, Polasky, and Adams (2000) on privately owned open space; Acharya and Bennett (2001) on the value of land cover surrounding housing; and Tyrvainen and Miettinen (2000), Rodriguez and Sirmans (1994), and Benson, Hansen, Schwartz, and Smersh (1998) on the value of natural views.

[6]As discussed below, in order to implement the model, this measure is aggregated to the neighborhood level.

[7]For tractability, the levels of non-residential development are treated as exogenous and are not formally modelled.

Household $i \in I$ maximizes its utility by choosing neighborhood $j \in J$. Each neighborhood is characterized by its land price P_j, open space amenities O_j^p and O_j^n, and controls for additional spatially delineated amenities A_j. Households are characterized by their income y_i and taste for neighborhood amenities α_i. Implementation of the model is facilitated by adopting the indirect preference specification given in equation 2.1.

$$V(P_j, O_j^p, O_j^n, A_j | y_i, \alpha_i) = \left[\frac{1}{1-\nu} y_i^{1-\nu} - \frac{1}{1+\eta} B \left(\frac{P_j}{1+(O_j^p)^\lambda} \right)^{\eta+1} \right] G($$

The price of land, P_j, is assumed to equal the average 1992 land assessment per square foot annualized following Poterba's (1992) approach for incorporating tax and appreciation effects. The model assumes that the privately capitalized open space component O_j^p is captured by including the average distance from a home in neighborhood j to a protected parcel of open space. The neighborhood or endogenous component of open space, O_j^N equals the percentage of the land area in zone j which is undeveloped. Permanent income and heterogeneity in the taste for the locational attributes are introduced through y_i and α_i respectively. The distribution parameters for these variables are not directly observed and are assumed to follow a bivariate log-normal distribution.

2.2.2 Supply Model

Price-induced supply responses are incorporated using an empirical model of the conversion of land from undeveloped to residential use. The estimates from this model provide for each parcel a probability distribution of the reservation price at which the parcel will be converted (see Walsh (2004)). Based on these estimates, the land supply function maps neighborhood specific residential land prices P_j to the supply of residential land in each neighborhood S_j as in equation 2.4.

$$S_j(P_j) = \overline{L}_j + \sum_{k \in j} F_{kj}[P_j] \times \text{AREA}_{kj} \tag{2.4}$$

$F_{kj}[.]$ is the CDF for the reservation price of parcel k in zone j, AREA_{kj} is the area of parcel k and $\overline{L}_j$ is the area of land in residential use in neighborhood j as of 1984.[10] Figure 1 presents a graphical representation of equation 2.4 under the assumption of a logistic distribution of reservation prices.

Figure 1 also illustrates how government purchases of land affect the market. New land protection has two effects on supply in each zone. First, protection reduces the aggregate

[10] In the land market equilibrium model, land developed prior to 1984 is treated as irreversibly developed.

Figure 1: Impact of Land Protection on Residential Land Supply Model

supply of land available for residential development from Lu to Lu . In addition, depending on the distribution of reservation prices for the protected parcels, removal of parcels from the land market will cause a displacement of the supply curve. The distribution of the reservation prices of parcels identified for protection under each of the different policies will depend on the attributes of the parcels selected for protection. Figure 1 describes how two policies for protecting an identical acreage of land can have different effects on supply. The first policy results in purchases of parcels with relatively low reservation prices while the second policy purchases parcels with high reservation prices. The second policy will have little effect on the land market until the demand reaches point A while the first policy has an impact as soon as demand increases above L_l. This example demonstrates how

2.2.3 Equilibrium in the Land Markets

The land market outcome described by individual choices of location and lot size $\{j_i, d_i\}_{i \ I}$ arises from the interaction of supply and demand in the residential land market. Neither side of the market internalizes the externalities that arise through the neighborhood character component of the metropolitan landscape, $O_j^n\left(\{j_i, d_i\}_{i \ I}\right)$. This externality complicates characterization of equilibrium. As consumers respond to exogenous changes in the market, not only will prices adjust, but changes in their locational choices and land demands will lead to new values for the endogenous open space measures. These open space changes then in turn imply revised consumer land demands.

Given a finite set of location choices, J and households I, a Nash equilibrium is characterized by equations 2.5 thru 2.8.

$$i \quad j_i = \arg\max_{j \ J} \left(v\left(P_j, O_j^p, O_j^n, A_j \mid y_i, \ \alpha_i \right) \right) \tag{2.5}$$

$$i \quad d_i = - \frac{\partial \ v[P_j, O_j^p, O_j^n, A_j \mid y_i, \ \alpha_i]}{}$$

3 Estimation of Household Preferences

This analysis extends the Epple-Sieg framework for estimating preferences based on the properties of locational equilibrium by allowing open space to have two effects on individual preferences.[11] Access to protected public land, O^p, influences demand for lot size directly, while the neighborhood quality measure O^n acts at the extensive margin, affecting community choice. As a result it influences demand for lot size indirectly. This specification allows for heterogeneous tastes for the index of public goods (G) via the taste parameter α_i. The estimation strategy recovers four sets of parameters: the parameters of the joint distribution of income and tastes for location specific amenities; the parameters of the indirect utility function; the parameter of the function mapping open space to the augmentation parameter (λ); and, the parameters of the public good index.

The parameters of the household utility model are estimated by taking advantage of the conditions for a locational equilibrium.[12] Following Epple and Sieg (1999) a two-stage simulation-based procedure is used to estimate the model's parameters. The first stage recovers the heterogeneity parameters, indirect utility parameters, and the parameters of the

[11] The approach taken here is related to two additional empirical approaches to estimating household sorting models that have been developed recently. Bayer (2000) extends the differentiated product model of Berry, Levinsohn, and Pakes (1995) to estimate an equilibrium sorting model of residential and schooling decisions of households with elementary school-aged children in California. In a more recent application of this approach, Bayer, McMillan, and Rueben (2002) use restricted access census data that links household demographics to characteristics of the actual residence and census block to estimate a model of household choice in the greater San Francisco Bay area. Their analysis adopts a probabilistic notion of housing market equilibrium over a fixed set of houses with fixed characteristics. In equilibrium, for each house, the sum across individuals of the probability of occupying said house is equal to one. The second approach is based on the computable equilibrium model of Nechyba (2000) and has been developed by Ferreyra (2001). She estimates an empirical model that jointly determines school quality and household residential and school choices within an economy composed of multiple public school districts and private schools. Equilibrium under Ferreyra's model involves assignment of househeholds to a fixed stock of houses with fixed characteristics such that each house is occupied and no household can be made better off relocating to a different house.

Each of these two approaches are variants of the basic assignment model which treats the quantity and characteristics of the housing stock in each region as fixed. This assumption is not problematic for the types of analysis undertaken by Bayer et al. (2002) and Ferreyra (2001). However, because of the critical connection between changing development patterns and open space provision, the assumptions regarding the supply side of the equilibrium model makes these approaches inappropriate for the current analysis.

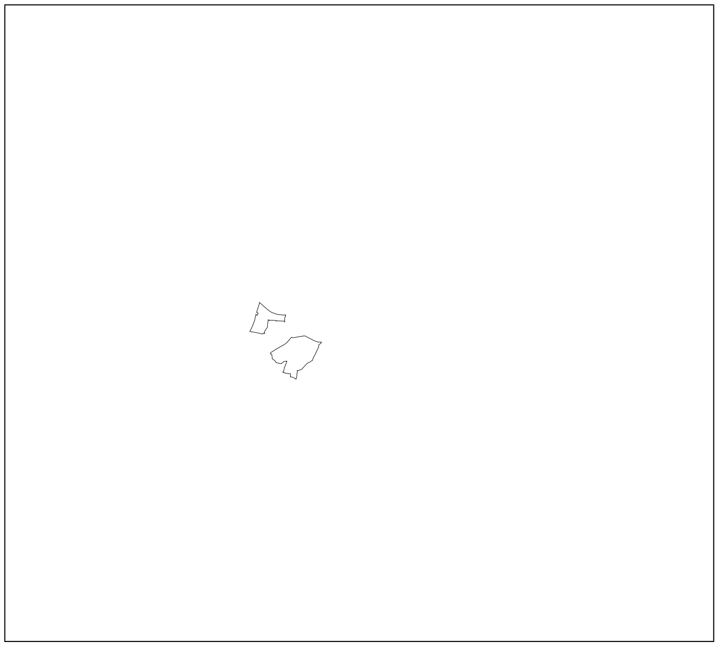
[12] As with all of the recent advances in estimating locational equilibrium models, this methodology requires the assumption of costless mobility.

augmentation function. The algorithm for the first stage begins with starting values for these first stage parameters. A simulation algorithm maps these parameter values into a unique vector (normalized to a relative scale) of local amenity levels (G-index) which implies a sorting across zones that matches the predicted zone populations to actual zone populations. Each iteration of the first stage parameter vector together with the implied relative levels of the public good indices leads to a unique sorting of households (characterized by income and taste for locational amenity levels) across zones. This sorting is used to recover a predicted distribution of land demands for each zone. Quartiles of the predicted land demand distribution for each zone are differenced from those actually observed in each zone. These differences form the basis of a Minimum Distance Estimator (MDE). A numerical optimization routine is then employed to find the set of parameters which minimize the value of the MDE's objective function. The specifics of estimation are presented in the Appendix.

4 Data

The study area for this project is Wake County, North Carolina. The county includes the state capital and a portion of the Research Triangle Park. It has experienced rapid development and contains significant areas of protected and unprotected open space. The empirical model requires dividing the county into a set of spatially distinct choice alternalocal jurisdictional boundaries, major roadways and school attendance boundaries. Figure 2 shows the boundaries of the 91 zones.

Figure 2: Map of 91 Zones and Protected Open Space



Shaded areas represent parcels of protected open space.

Information on lot size, 1992 tax assessments (distinguished into separate land and structure assessments) and current land use were assembled from GIS parcel data and tax records supplied by the Wake County Assessor's Office. This data set contains information on approximately 230,000 parcels of land in Wake County. Collectively the parcels cover 510,677 acres of land and account for 95% of the area of Wake County, with the remaining 5% comprised mainly of roads and road right of ways. Each parcel is identified as having one of 24 land use codes. Based on these codes, each parcel is placed into one of the 5 categories presented in table 1.[14]

[14] Privately held open space is comprised of the following land uses: agricultural uses, vacant, cemetery, golf course, single family residential greater than 10 acres, and all parcels that were developed after 1992.

Table 1: Land Use Summary from Parcel Maps

Land Use	Acres	Percentage of Total
Business/Commercial	16,694	3.27%
Residential	102,897	20.15 %
Protected Open Space	81,084	15.88%
Privately Held Open Space	295,566	57.88%
Other	14,436	2.83%
Total	510,677	100%

Two open space measures are developed for use in the locational equilibrium model. The first is the endogenous or neighborhood component of open space, O_j^n which is proxied for using the percentage of each zone's total land area which is in open space. This measure is comprised of a mixture of publicly protected land and privately held land in open space uses such as agriculture. The second measure is the exogenous component of open space, O_j^p. This measure captures the average distance to protected open space for each home in a given zone. In order to construct this measure for each zone, a unified GIS description of

Table 2: Summary Statistics for 91 Residential Zones

	Mean	Std. Dev.	Min.	Max.
Average Assessed Lot Price ($ per ft^2)	0.165	0.128	0.013	0.553
Lot Expenditure 25th Quartile	15741.44	15630.43	1742.4	88862.4
Lot Expenditure 50th Quartile	23450.6	23254.7	3920.4	131551.2
Lot Expenditure 75th Quartile	43703.6	78702.25	5662.8	652528.8
Residential Lot Count	1005.978	757.86	3	3026
Population Share	0.011	0.008	0	0.033
Ratio of Commercial to Residential Acres	0.2428	0.3225	0	1.9812
Percent of Zone in Open Space (O^n)	0.61	0.215	0.17	0.965
Average Distance to Protected Open Space (feet)	3004.69	3065.06	320.31	20326.39
Private Open Space Measure (O^p)	0.856	0.15	0	0.983
Average Distance to CBD (feet)	40298.07	25625.36	2018.619	96833.3

control for commuting distances, the average distance to several employment centers, including the state capitol and Research Triangle Park (RTP) was calculated for each zone. Additionally, indicator variables for each of the 13 local jurisdictions are used to control for other unobservable attributes. Summary statistics for the 91 zones are contained in Table 2.

Finally, the model requires that land prices be converted to annual rental values. Using the calculation suggested by Poterba (1992), annualized rents are given by equation 4.9.

$$R = [(1 - \tau)(i + \tau_p) + \beta + m + \delta - \pi] \; P_H \tag{4.9}$$

The specific variables and their values are given in Table 3.

analysis, measures of individual school quality were never found to be significant. This is likely because the presence of magnet schools and frequent changes in school assignments further serve to weaken the link between location choice and school quality. Crime is a second concern. Unfortunately, the only data available at a sufficiently disaggregated level is for homicides. Due to the small number of homicides in the County and their lack of spatial variation, crime data are excluded from the analysis.

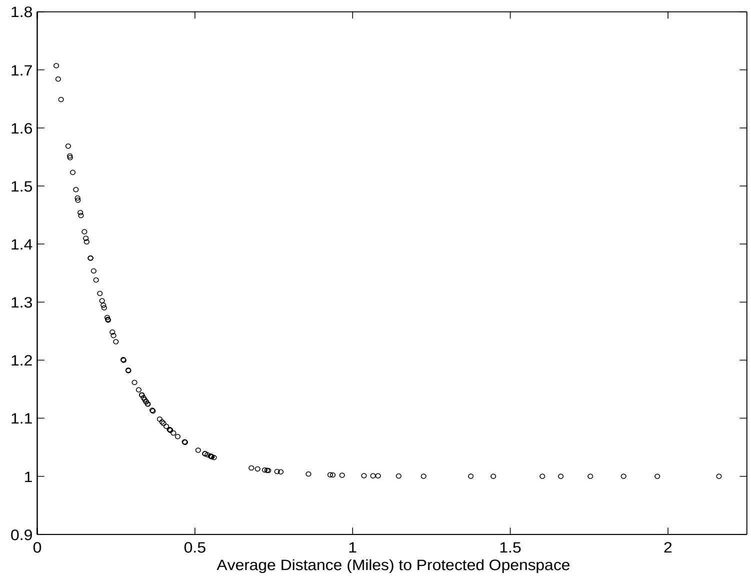
Table 3: Values for Poterba Calculation

τ	owner's marginal tax rate	15%
τ_p		

Table 4: Household Location Model's First Stage Parameter Estimates

	Parameter Estimate	Estimated Asymptotic S.E.
Variance of Ln (α)	0.2981	(0.0033)
Variance of Ln (y)	0.2823	(0.0061)
Price Elasticity (η)	-0.6174	(0.0166)
Income Elasticity (ν)	0.7474	(0.0844)
Demand Parameter (B)	1.5545	(0.0051)
α-y Correlation (ρ)	-0.0196	(0.004)
Augmentation Parameter (λ)	21.8257	(0.0001)

Figure 3: Effective Price Function - Household Location Model



* For better scaling, one observation corresponding to an average distance of 3.8 miles has been omitted.

open space are substitutes. Figure 3 illustrates what the relationship implies for how the average distances to protected open space (on the x-axis) translates into values for the price augmentation factor $1 + (O_j^p)^\lambda$ (on the y-axis). As the average distance moves from half a mile to one quarter of a mile, the augmentation factor increases by 17.5%. Evaluated at the mean (across zones) of the 50th percentile of lot expenditure ($23,451) this change in the augmentation factor roughly corresponds to a one time incremental willingness to pay for the change in average distance of $4,104. This estimate is less than that of Correll et al. (1978). Using their marginal willingness to pay measure of $4.20 per foot of walking distance to protected open space leads to an estimate of $12,250 as the value of decreasing the distance to protected open space for a given house by a quarter mile.[20]

In addition to the parameters discussed above, first stage estimation recovers the index of local amenities, G_j

Table 5: Decomposition of Augmentation Model's G-index

Model	Open Space[2]	Pct. Bus/Com	Mean Emp. Dist.	Ln G1	K	Jur. Ind.	I.V.	Critical Point
I	-1.5221			-0.5839	0.04733			0.3285
	(0.2638)			(0.3289)	(0.0187)			
II	-1.6472	0.6986		-0.7311	0.0388			0.3035
	(0.3607)	(0.4124)		(0.4460)	(0.0177)			
III	-1.4892	0.6032	-0.0439	-1.3758	0.0274			0.3357
	(0.2461)	(0.3735)	(0.0151)	(0.5210)	(0.0090)			
IV	-1.5255	0.2628	-0.0014	-0.4159	0.0675		X	0.3277
	(1.0205)	(0.5698)	(0.0082)	(1.0058)	(0.1021)			

is increasing in open space percentage over an initial range of values and then beyond some critical point a negative relationship holds.[21] The values for this critical point are included in the table. For the simple specification of Model I, the value of this critical point is 33%. Model II adds as an explanatory variable the percentage of each zone that is in a business or commercial usage. This measure is treated as exogenous. The coefficient on this variable is positive in all four specifications, but is statistically insignificant, in all but model III which controls for distance to employment centers but does not instrument for the endogeneity of open space percentage. Models III through V incorporate the mean distance to eight different employment centers as explanatory variables.[22] All three models result in the expected negative coefficient on this variable. However, the coefficient is not significant for instrumental variable estimates.

With the endogeneity of each zone's open space percentage a potential concern, models IV and V instrument for open space percentage.

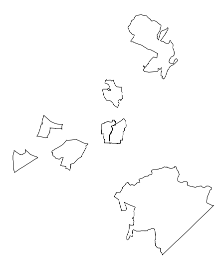
points in the range of 0.30 to 0.33. These empirical estimates suggest that local amenities are maximized when the level of open space provision lies between 30 and 33 percent of a given neighborhoods land area. Because it includes instruments for the open space measures and indicator variables for the local jurisdictions, model V is selected for use in the general equilibrium policy simulations.

6 Policy Simulations

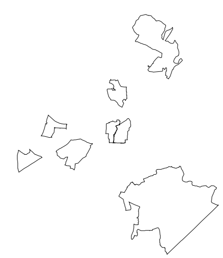
This section of the paper utilizes the estimated household model to evaluate the impacts of different policy interventions. Three policies are analyzed using the model, two value is accounted for by a property's development potential. Based on these estimates, the analysis assumes $15 million available for protection will protect approximately $22.5 million (*i.e.* 3/2 of $15 million) worth of residential land. The parcels protected under the two land protection policies are shown in Figure 4.

The first four rows of Table 7 summarize the implementation of the policies. The land protection policies are realized in the model by removing the newly protected parcels from the set of potentially developable parcels in the land supply functions and recalculating the open space access measures to reflect the additional parcels of protected open space. These calculations are then input into the model and new equilibrium values for household location choices, neighborhood land prices and neighborhood open space percentages are

Figure 4: Policy Simulations

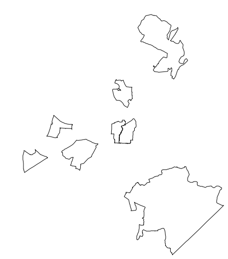


Land Protection Policy I



Land Protection Policy II

Shaded areas represent the additional protected land under each policy.



Growth Ring Policy

Shaded areas represent zones where development is frozen at 1984 levels.

zones at their 1984 levels by fixing the model's land supply function for these zones at their 1984 level of residential development. Because this policy is designed to replicate zoning or growth restrictions there is no increase in access to publicly owned open space, O^P. The zones chosen for inclusion in this policy (identified in Figure 4) are those closest to the center of the county with a 1984 density of development less than .1 residential households per acre.

Implementation of the final policy differs from that of the land protection policies. First, no new protected open space is created under the growth ring policy so no adjustments are made to the open space access measure, O^p, for the different zones. Second, the land supply adjustments for the policy are an order of magnitude greater than those for the land protection policies.

basic algorithm is as follows. The vector of zone specific endogenous land amenities (open space percentages) is fixed at an initial level and a vector of prices is identified which clears the land markets.[26] This market clearing process leads to a change in the level of privately held open space in each zone away from the initial level. This occurs for instance when households move into a zone and reduce the stock of private vacant land. To account for this endogeneity, following the land market clearing, new amenity levels are calculated (new levels of the public good index G which account for the endogenously deterimined open space percentages) for each community. A new vector of land market clearing prices, based on the updated G vector, is then computed. This process is iterated until a fixed point is identified. In practice, this process typically takes approximately 10 iterations.

6.3 Baseline Equilibrium

The general equilibrium model assumes that the spatial distribution of households that we observe in Wake county can be perfectly characterized by a Nash Equilibrium arising from the household choice and land supply functions specified above. To provide a baseline for comparison to the three policy simulations, the G.E computation is performed using the observed level of endogenous landscape amenities (open space percentages) as starting values and holding the distribution of protected open space fixed. If the parameter estimates, structural specification, and Nash equilibrium assumption are the "true model" that gave rise to the observed actual location choices then this computation should yield equilibrium values for the endogenous open space measure that correspond exactly to the observed values. Calculating this initial equilibrium provides an appropriate baseline with which to compare the equilibrium outcomes under the various policy scenarios. The baseline equilibrium reasonably replicates the observed distribution of open space. In all but 4 of the zones, the computed baseline general equilibrium open space percentages were within 20 percentage points of the actual observed value with the majority differing by much less.

6.4 Welfare Measurement

One goal of the analysis is to evaluate the welfare implications of the different policies. In the context of the model, welfare measurement must account for four distinct adjustments (relative to the baseline equilibrium) that are associated with the counterfactual equilibria. First, under the counterfactual, household i's new location choice j may differ from its original location choice j^0. Second, as households relocate and adjust their optimal lot-size, the open space percentages in each zone adjust, leading to new levels of the public good index. These new levels are denoted by G'. Because O^n enters the G-index in a quadratic form, the direction of the change in G depends on the change in O^n and the initial level of O^n. Above the critical point of 32% increases (decreases) in O^n reduce (increase) the level of G. Below the critical point, the relationship is reversed. Third, household location adjustments and the supply effects associated with the additional land protection cause

6.5 Zone-Specific Decomposition of Policy Outcomes

To illustrate the interactions that occur in the model, consider the impact of the two land protection policies on specific neighborhoods. These Policies have both direct and indirect impacts that affect the new equilibrium. I begin by considering the direct effects. First, the protection of additional parcels in a given zone increases O^P, the measure of access to protected open space.[30] This increase makes the given zone more attractive relative to other zones that do not experience an increase in protected open space. Also, because O^P is a substitute for land in household preferences, *ceteris paribus*, this increase in O^p decreases per household land demand. The immediate impact of the increase in protected open space (decreases) in this amenity level will increase (decrease) the number of households choosing

Table 6: Decomposing the Policy Impacts

Policy Evaluated	Policy I		Policy II	
Sample Zone	A1	B1	A2	B2
Initial Open Space Percentage	86.9%	22.4%	63.4%	75.2%
Acres Protected	342.7	0	931.3	0
Initial Augmentation Factor	1.026	1.301	1.079	1.096
Change in Augmentation Factor	0.063	0	0.262	0.015
Avg. Partial Equilibrium WTP	$66.37	$0.00	$315.10	$29.02
G.E. Open Space Percentage	86.9%	22.3%	71.4%	73.5%
Change in G - index				

($66.37 per household). The general equilibrium calculation identifies additional changes. First, new location and lot size changes offset the newly protected open space yielding virtually no change in the percent of open space in the zone. Therefore there is no change in the G-index. The increase in protected land shifts up the land supply function for the zone as some parcels are removed from the set of potentially developable parcels. This upward influence on the zone's price combines with the upward demand shift associated

These changes yield an average G.E. WTP measure for the zone's households of -$15.39. In contrast to zone/policy pair 'A1', accounting for the change in property values does not completely offset this negative WTP and yields an aggregate benefit measure of -$1.79 per household.

are expressed for the region as a whole on a per household level. The first four rows of the table present the specifics of the two land protection policies. Because the household preference model incorporates access to open space through the price augmentation function, the change in the price augmentation factor measures how much the new land protection increased access to protected open space. Row four reports the average change in the price augmentation factor under the two land protection policies.[35] The difference in this measure across the two land protection policies demonstrates that the spatial distribution of land protection is important for determining the direct benefit of the policy. Policy II which focuses protection in more densely populated areas and protects many small parcels leads to a greater increase in access to protected open space (averaged across zones).

The table's fifth row reports the change in residential development, relative to the baseline. For the two land protection policies there is a reduction in total residential development. However, the reduction in development from land acquisition is much less than one for one. Policy II has the larger reduction in development, but only results in a .12 reduction in acres developed per acre protected. A key point to note here is that even though a shift from policy II to Policy I more than doubles the amount of protected open space (an increase of over 5000 acres) endogenous adjustments in privately held open space lead to a *net loss* in open space of more than 200 acres under policy I relative to policy II.

The development freeze (growth ring) policy leads to much larger changes in development. This is because on average the policy forces households out of low amenity (G) low price zones where lot sizes are large into relatively higher priced higher G areas. The net reduction in development under this policy is 2,668 acres. Row six reports the relationship between the number of acres protected and the reduction in development for the two Land Protection Policies. Row seven reports the change in average lot size and provides another measure of the overall impact on developed land. The growth ring policy has the largest

[35] Recall that the development freeze policy leaves this factor unchanged.

impact, reducing average lot size by 4.2%.

Row eight of the table reports the average (across zones) change in endogenous public good level. For some summary statistics such as this change in public good level, it unclear what is the appropriate unit of observation over which to calculate averages. Averaging across zones holds the zone constant, but due to sorting doesn't necessarily reflect the change actually experienced by households initially located in these zones. Averaging across the change experienced by households corrects for this problem, but no longer holds the zone constant. This distinction is most important for the change in public good levels. Row eight of the table reports the average across zones, which is negative for all three policies. Row nine reports the change actually experienced by households which is positive under all three policies – reflecting that on average, in equilibrium, households relocate to zones that are closer to the open space 'bliss point'. Row ten reports the average change in zone land prices. Rows eleven and twelve report the low and high estimates of capitalization changes.[36] These figures reflect a combination of general equilibrium effects and the capitalization of changes in relative public good levels into residential land prices.

The capitalization changes are tightly bounded for the land protection policies. This is not the case for the development freeze policy. The large difference stems from the problem that no lower bound (above $0.00) on the reservation price is available for those parcels that were developed under the baseline scenario, but are blocked from development by the proposed policy. For the development freeze policy, under which development is blocked in large sections of the county, the lower bound reservation price of $0.00 leads to an "extreme lower bound". Therefore, the upper bound is likely to be the more closely reflect the actual measure. Without data on 1984 prices, it is impossible to provide definitive evidence of this relationship.

The results demonstrate that there is heterogeneity in the effects of the different policies

[36] For each household, these are calculated as discussed above in footnote 29.

on land capitalization. The Growth Ring Policy leads to positive capitalization effects and Land Protection Policies I and II lead to negative capitalization.

Row thirteen reports bounds for the average Compensating Variation(CV) for each of important role in the overall welfare measurements.

7 Conclusions

Open space protection and "anti-sprawl" policies are proliferating in the U.S. This analysis advances the work on empirical locational equilibrium models to provide an initial analysis of these policies in a framework that incorporates the endogeneity of both privately held open space and land conversion decisions. The results highlight the importance of these adjustments for understanding the impacts of land market interventions.

From a policy perspective, four key results emerge. First, increasing the quantity of protected open space may not reverse a trend toward low density development. Accounting empirical locational equilibrium models, the analysis incorporates an empirically estimated supply model into the locational equilibrium framework. These methodological contributions are central to the resulting policy analysis.

APPENDIX: Estimation of Locational Equilibrium Model

A two-stage process is used to estimate the parameters of the household model. The first stage isolates the heterogeneity parameters, indirect utility parameters, and the parameter of the augmentation function. These are labelled the first stage parameters. The basic algorithm for the first stage begins with an initial guess for the values of these parameters. A simulation algorithm maps these parameter values into a unique vector (normalized to a relative scale) of local amenity levels, the vector G, which implies a sorting across zones that matches the predicted zone populations to actual zone populations. Each iteration of the first stage parameter vector together with the implied relative levels of the public good indices leads to a unique sorting of households (characterized by income and taste for locational amenity levels) across zones. This sorting is used to recover a predicted distribution of land demands for each zone. Quartiles of the predicted distribution of lot sizes for each zone are differenced from those actually observed in each zone. These differences form the basis of a Minimum Distance Estimator (MDE). A numerical optimization routine is then employed to find the set of parameters which minimize the value of the MDE's objective function. Finally, the impact of different location amenities are on the levels of G are identified using a moment-based estimator to decompose the vector into a systematic component and idiosyncratic shocks.

The specifics of the model are as follows. Each household can choose to live in one of ninety-one discrete zones within a predefined metropolitan area. Conditional on choosing zone j, household i's indirect utility function is given by equation 2.1. P_j is assumed to equal the average 1992 land assessment per square foot annualized using Poterba's (1992) approach for incorporating tax and appreciation effects. Permanent income and heterogeneity in the taste for the locational attributes are introduced thru y_i and α_i respectively. These variables are not directly observed but are assumed to follow a bivariate log-normal distribution.

Using Roy's identity, conditional on choosing community j, household i's optimal residential land demand is given by:

$$L^D_{i,j} = \frac{1}{1 + (O^p_j)^\lambda} B P_j^{-\eta} y_i^\nu \qquad \text{(A.1)}$$

η and ν are the price and income elasticity of demand for residential acreage respectively.

To illustrate how changes in G affect location decisions, consider the locus of households that are indifferent between community 1 and community 2, defined implicitly by the requirement:

$$\left[\frac{1}{1-\nu} y^{1-\nu} - \frac{1}{1+\eta} B P_1^{-\eta+1}\right] G_1^\alpha = \left[\frac{1}{1-\nu} y^{1-\nu} - \frac{1}{1+\eta} B P_2^{-\eta+1}\right] G_2^\alpha \qquad \text{(A.2)}$$

This expression can be solved for α:

$$\alpha = \ln \left[\frac{1}{1-\nu} y^{1-\nu} - \frac{1}{1+\eta} \right.$$

A.5 as follows.

$$C_{1\ 2|y_i} = \left\{ \alpha_i \left| \ln\left(\frac{G_2^{\alpha_i}}{G_1^{\alpha_i}}\right) < K_{12}^{i} \right. \right\} \qquad \text{(A.9)}$$

This expression suggests that only relative levels of the G-index vector can be identified. Less obvious is the second identification problem involving the mean of the log of the taste heterogeneity parameter α_i. The problem is as follows.

Define:

$$\alpha_i = \exp(\tilde{\alpha}_i) \qquad \text{(A.10)}$$

By assumption:

$$(\ln y_i, \tilde{\alpha}_i) \sim N(\mu, \Sigma) \qquad \text{(A.11)}$$

Each household can be represented by functions of draws on the random variables $R_{\alpha i}, R_{yi}$, where $R_{\alpha i}, R_{yi}$ are defined by:

$$\begin{pmatrix} R_{\alpha i} \\ R_{yi} \end{pmatrix} = \Sigma^{\frac{1}{2}} Z_i \qquad \text{(A.12)}$$

where Z_i is a random draw from the bivariate normal with mean 0, covariance 0 and variance 1 and Σ is the covariance matrix of $\ln y_i$ and $\ln \alpha_i$. Draws of $\tilde{\alpha}_i$ can be constructed based on draws from $R_{\alpha i}$:

$$\tilde{\alpha}_i = \mu_{\tilde{\alpha}} + R_{\alpha i} \qquad \text{(A.13)}$$

Substituting back into G^{α_i}:

$$G^{\alpha_i} = G^{\exp(\tilde{\alpha}_i)} = G^{\exp(\mu_{\tilde{\alpha}} + R_{\alpha i})} = G^{\exp(\mu_{\tilde{\alpha}})\exp(R_{\alpha i})} = \left(G^{\exp(\mu_{\tilde{\alpha}})}\right)^{\exp(R_{\alpha i})}$$

Substituting this result into the key expression from equation A.9 and generalizing to zone j yields equation A.14.

$$\frac{G_j^{\alpha_i}}{G_1^{\alpha_i}} = \left[\frac{G_j^{\exp(\mu_{\ln\alpha})}}{G_1^{\exp(\mu_{\ln\alpha})}}\right]^{R_{\alpha i}} = \tilde{G}_j^{R_{\alpha i}} \tag{A.14}$$

In the first stage, the vector of $J-1$ values of $\tilde{G}_j$ are identified. Equation A.14 indicates that the $J-1$ elements of the G vector are defined conditional on G_1 and $\mu_{\ln\alpha}$. Thus, for any arbitrary values of G_1 and $\mu_{\ln\alpha}$ there will exist a set of values for G_2 thru G_J which satisfy equation A.14. It is therefore necessary to choose a normalization for μ_{ln_α}.

In order to estimate the first stage parameters, including the vector of $\tilde{G}_j$ values, G_1 is set equal to one and $\mu_{\ln\alpha}$ is set equal to zero.[39] The value of G_2 is then defined implicitly by equation A.6. This implied value of G_2 is identified using a line search algorithm. The solution algorithm iterates through the communities to solve for the entire G vector. Once this vector has been identified, the income quartiles in each zone can be identified. These income quartiles are then substituted into the land demand function implied by Roy's identity in equation A.1 to derive the land demand quartiles that form the basis of the MDE.

The specific form of the Minimum Distance Estimator (MDE) is given in equation A.15.

$$\hat{\theta} = \underset{\theta \in \Theta}{\operatorname{argmin}}\ M(\theta)'\ \mathrm{W}\ M(\theta) \tag{A.15}$$

[39] These normalizations only serve to identify the first stage parameters, which are invariant to this normalization, and are *not* utilized for identification of the second stage parameters.

where θ is the vector of first stage parameters. The vector $M(\theta)$ is constructed as follows:

$$M(\theta) = \begin{bmatrix} \sqrt{\frac{N_1}{N}} \begin{bmatrix} \hat{Q}_1(\theta, .25) - Q_1(.25) \\ \hat{Q}_1(\theta, .50) - Q_1(.50) \\ \hat{Q}_1(\theta, .75) - Q_1(.75) \end{bmatrix} \\ \vdots \\ \sqrt{\frac{N_J}{N}} \begin{bmatrix} \hat{Q}_J(\theta, .25) - Q_J(.25) \\ \hat{Q}_J(\theta, .50) - Q_J(.50) \\ \hat{Q}_J(\theta, .75) - Q_J(.75) \end{bmatrix} \end{bmatrix}. \tag{A.16}$$

$\hat{Q}_j(\theta, .25)$ is the predicted value of the 25th lot size quantile in zone j when the model's first stage is evaluated at the parameter value θ and $Q_j(.25)$ is the observed 25th quantile in zone j. N is the total number of observations used to construct the quantile measures and N_j is the number of observations in aone j.

At the true value of the parameter vector θ_0 the only source of variance in $M(\theta_0)$ is sampling error in the estimates of the observed quantiles of the lot size distribution in each community. Therefore the Central Limit Theorem implies the limiting distribution for $M(\theta)$ presented in equation A.17:

$$\sqrt{N}\, M(\hat{\theta}) \xrightarrow{D} N(0, \Sigma) \tag{A.17}$$

where Σ is the block diagonal matrix:

$$\Sigma = \begin{bmatrix} \Psi_1 & & 0 \\ & \ddots & \\ 0 & & \Psi_J \end{bmatrix} \tag{A.18}$$

and Ψ_j is given by:[40]

$$\Psi_j = \begin{bmatrix} \frac{3}{16\, f_j^2(Q_j^{.25})} & \frac{1}{8\, f_j(Q_j^{.25})\, f_j(Q_j^{.5})} & \frac{1}{16\, f_j(Q_j^{.25})\, f_j(Q_j^{.75})} \\ \frac{1}{8\, f_j(Q_j^{.25})\, f_j(Q_j^{.5})} & \frac{1}{4\, f_j^2(Q_j^{.5})} & \frac{1}{8\, f_j(Q_j^{.5})\, f_j(Q_j^{.75})} \\ \frac{1}{16\, f_j(Q_j^{.25})\, f_j(Q_j^{.75})} & \frac{1}{8\, f_j(Q_j^{.5})\, f_j(Q_j^{.75})} & \frac{3}{16\, f_j^2(Q_j^{.75})} \end{bmatrix}. \tag{A.19}$$

$f_j\left(Q_j^{.25}\right)$ is the pdf of the distribution of lot size in the jth community evaluated at the .25 quantile, and $f_j^2\left(Q_j^{.25}\right)$ is the square of the pdf of the distribution of lot size in the jth community evaluated at the .25 quantile.[41] The asymptotic theory for MDE's presented in Newey and McFadden (1994) can be used to develop asymptotic estimates of the asymptotic covariance matrix of $\hat{\theta}$. Setting $\mathrm{W} = \Sigma^{-1}$ yields:

$$\hat{\theta} \overset{D}{\rightarrow} \mathrm{N}\left(\theta_0, \frac{1}{N}(\mathrm{G}\ \mathrm{W}\ \mathrm{G})^{-1}\right) \tag{A.20}$$

Where G is the $3J \times k$ matrix given by:

$$\mathrm{G} = \frac{\partial\, M(\hat{\theta})}{\partial\, \theta}. \tag{A.21}$$

One final issue related to the estimation of the first stage parameters is the treatment of the income in the model. The scale parameter of the land demand, B, and the mean of permanent income are not separately identified, estimation of the preference parameters requires an estimate of the mean permanent income in Wake County. Therefore, to estimate B a separate measure of the mean of permanent income in Wake County is introduced using Woods-Poole (1995). BEA's Personal Income Series was used to develop the measure ($56,123).

To decompose the G-index, in the second stage of the estimation procedure the locational attributes index is assumed to follow a semi-log function, $G_j = \exp(X_j\gamma + \varepsilon_j)$. G_1 is treated

[40] See Mood, Graybill, and Boes (1974) For a discussion of the asymptotic variance of order statistics.
[41] These values are estimated non-parametrically using kernel estimators.

as a nuisance parameter to be estimated as part of the second stage of the model. This yields the following linear model:

$$\frac{\ln(\hat{G}_j)}{\exp(\mu_{\ln\alpha})} = -\ln G_1 + X_j\gamma + \varepsilon_j \tag{A.22}$$

Two normalizations identify this model. First, the intercept of the semi-log model is assumed to equal $\ln G_1$. Second the coefficient on the first element of the local public good index is set equal to 1. Under this normalization, the coefficients on locational attributes can be interpreted relative to this first element and the value of $\mu_{\ln\alpha}$ reflects, on average, the relative importance of the locational attribute index in preferences.

A GMM procedure is then utilized to estimate the model's second stage parameters. The parameter vector is comprised of the parameters of the index γ,[42] $\ln G_1$ and $\exp(\mu_{\ln\alpha})$. The moments used in estimation are then of the form in equation A.23.

$$\frac{1}{J}\sum_{j=1}^{J}(\varepsilon_j)\,X_j \tag{A.23}$$

where ε_j as derived from equation A.22 and is given in equation A.24.

$$\varepsilon_j = \frac{\ln(\hat{G}_j)}{\exp(\mu_{\ln\alpha})} + \ln G_1 - X_j\gamma. \tag{A.24}$$

The set of explanatory variables X includes the percentage of the zone in open space. The square of this measure is also included to allow for non-linear open space effects. It is to be expected that this measure is correlated with the the error term ε. However, equation A.23 can be estimated using instrumental variables with a vector of instruments Z_j that includes all of the elements of X except the open space percentage measures and additional variables assumed to be correlated with open space percentage, but not correlated with ε. The Instrumental Variables Estimator is defined by equation A.25.

$$\frac{1}{J}\sum_{j=1}^{J}(\varepsilon_j)\,Z_j. \tag{A.25}$$

[42] Recall that the parameter on open space percentage is fixed at one.

References

Nechyba, T. (2000). Mobility, Targeting and Private School Vouchers. *American Economic Review*, *90*, 130–146.

Nevo (2001). Measuring Market Power in the Ready-to-Eat Cereal Industry. *Economietrica*, *69* (2), 307–342.

Newey, W. and McFadden, D. (1994). Large Sample Estimation and Hypothesis Testing. In *Handbook of Econometrics IV*. Elsevier North Holland.

Palmquist, R. B. (1992). Valuing Localized Externalities. *Journal of Urban Economics*, *31*, 59–68.

Polinsky, A. (1977). The Demand for Housing: A Study in Specification and Grouping. *Econometrica*, *45* (2), 447–461.

Poterba, J. (1992). Taxation and Housing: Old Questions, New Answers. *American Economic Review*, *82 (2)*, 237–242.

Ready, R., Berger, M., and Blomquist, G. (1997). Measuring Amenity Benefits from Farmland: Hedonic Pricing vs. Contingent Valuation. *Growth and Change*, *28*, 438–458.

Rodriguez, M. and Sirmans, C. (1994). Quantifying the Value of a View in Single-Family Housing Markets. *Appraisal Journal*, *62*, 600–603.

Sieg, H., Smith, V., Banzhaf, H., and Walsh, R. (2003). Measuring the Impact of Air Quality on Housing Markets and Residential Choices in Southern California. Discussion Paper.