Valuation of Education and Crime Neighborhood Characteristics Through Hedonic Housing Prices

Robin A. Dubin
Wayne State University

Allen C. Goodman
The Johns Hopkins University

This article studies households' valuation of neighborhood amenities through analysis of housing as a bundle of structural and neighborhood characteristics. Using statistical techniques, hedonic prices, or incremental values in the market, can be imputed to each element of the bundle. Twelve variables describe neighborhood crime; twenty-one variables describe neighborhood schools. Together with structural variables, they significantly and substantially explain house prices in the Baltimore metropolitan area. The technique is also shown to provide a market framework for approximating the benefits of localized neighborhood improvements.

One of the thornier problems in recent urban economic analysis has been the treatment of *neighborhood*. In a discipline that seeks an easily describable good produced by a well-specified process, neighborhood provides neither. A household can move into a

---

Robin A. Dubin is Assistant Professor of Economics at Wayne State University, and Allen C. Goodman is Associate Professor of Political Economy and Research Scientist at The Johns Hopkins University. The authors wish to thank David Puryear, Ralph Taylor, and the two anonymous referees for their helpful comments on earlier drafts of this paper. We also thank Central Maryland Multiple Listing Service, the Baltimore City and Baltimore County Police Departments, and the Johns Hopkins Center for the Social Organization of Schools for house price, crime, and education data. None of the people or organizations listed is in any way responsible for the analyses or conclusions presented. Requests for reprints should be addressed to Dr. Allen C. Goodman, Center for Metropolitan Planning and Research, The Johns Hopkins University, Baltimore, MD 21218.

Population and Environment, Vol. 5(3), Fall 1982
0199-0039/82/1500-0166$02.75

new residence and renovate the house and yard, yet have little or no control over the neighborhood surrounding it. Moreover, the social, racial and public service components of neighborhood generally defy any easy characterization in one, or even several, dimensions. [For an up-to-date collection of economists' conceptions of neighborhoods, see Segal (1979).]

This article presents a study of households' valuations of neighborhood characteristics through the analysis of a house as a bundle of structural and neighborhood characteristics. Using statistical techniques, an implicit valuation or *hedonic price* can be attributed to each characteristic of a bundle. This hedonic price is interpreted as the added value *in the market* of a unit of the characteristic.

The particular strength in this study is the unusually detailed information that is available on two major facets of neighborhood quality, crime and education. Twelve variables describe neighborhood crime and twenty-one variables describe neighborhood schools. These variables are shown to exhibit significant and substantial power in explaining house prices in the Baltimore metropolitan area.

The article begins with a brief discussion of the hedonic technique and its application to housing markets. Specifying the pertinent structural and neighborhood characteristics is crucial in this analysis, and the following section discusses the methods used with a data set of house sales in the Baltimore area. After treatment of each of the neighborhood characteristics, hedonic price regressions are estimated and interpreted. The concluding section examines the policy relevance of the analysis.

HEDONIC PRICE ANALYSIS

Consider the analysis of a consumer's demand for food. Several categories of goods are purchased at the supermarket, chosen according to prices, incomes and consumer tastes as the shopper wanders through the market with a shopping basket. The economist following the shopper through the market could enumerate his or her purchases, and through several observations, learn about the determinants of consumer demand. Suppose, however, that consumers were constrained to purchasing already filled shopping carts, varying in composition, but with only a total bundle price presented. It would be considerably more difficult, on the basis of observed behavior, to determine individuals' tastes for specific goods within the cart.

The purchase of a house presents exactly this problem. Purchasers do not go out to buy "housing"; rather they wish a large lot, two-car garage, hardwood floors or good schools, and an additional amount of one will not always compensate for a deficient amount of another. Clearly, an analysis of housing market behavior must account for the heterogeneous nature of the housing bundle.

Hedonic price analysis is based on the premise that a heterogeneous good (such as housing) can be completely described by a relatively small number of characteristics. In functional notation, for a housing bundle, the expression

$$P = f(s_1,...,s_j,n_1,...,n_k \tag{1}$$

implies that the selling price of a house (or rental value of an apartment), P, is a function of a vector of j structural characteristics and k neighborhood characteristics. The partial derivative of this function gives the incremental value, all else equal, to the bundle price, of a one unit increase of the characteristic in question. In order to get the true hedonic price of each characteristic, a very large and varied sample is necessary. [This is rather heuristic definition and description of the hedonic price technique. For a more rigorous treatment, see Rosen, (1974).]

Neighborhood characteristics can include racial and income variables, levels of crime and school quality, location relative to workplaces, and local property tax rates. The hedonic price of a neighborhood characteristic can be interpreted as the incremental value of an increase of one unit of the characteristic (school quality, for example), which should presumably increase the value of those houses having access to it. It need not reflect solely the cost of supplying school quality; indeed, it does not say that those who are paying this incremental amount are reasonable in paying it. It can be treated, however, as a valuation that someone is willing to pay, which faces those who would buy or sell within the neighborhood.

Although the expressions have often been estimated in linear form, this is overly restrictive. It implies, for example, that the eighth room in the house provides the same marginal valuation as does the fourth, irrespective of the number of bathrooms, the amount of floor space and so on. More recent work has emphasized the amount of floor space and soon. More recent work has emphasized the importance of interactive forms such as semi-log or log-log, and some of the most recent work applies a transformation suggested by Box and Cox (1964).

$$\frac{P^{\lambda} - 1}{\lambda} = \alpha_o + \Sigma_j \beta_j S_j + \Sigma_K \delta_K n_K \varepsilon \tag{2}$$

This procedure has both the linear (λ equals 1) and semi-logarithmic (λ equals 0) forms as special cases and provides a more flexible format for the estimation of functional form. [Goodman, (1978), applies this technique to housing market analysis.]

STRUCTURAL AND NEIGHBORHOOD VARIABLES

Estimation of the hedonic prices of the various characteristics requires enumeration of the entire set of characteristics valued by households. Omission of one or more variables will generally bias the coefficients of the included variables; thus the sample used must be large and varied. For example, although a basement may be important, if all houses have them, it will be impossible to estimate a hedonic price coefficient. If patios are positively correlated swimming pools, but the latter variable is omitted, the coefficient of "patio" will overestimate its value.

A well-specified set of structural characteristics should include variables describing both quantity and quality of housing. Examples of the former are number of rooms and bathrooms, number of garage spaces or lot size. Examples of the latter (although neither set is necessarily exclusive of the other) are the presence of a pool or patio, dwelling type (attached versus detached, for example) or the availability of air conditioning. Several of these are customarily modeled as "dummy" variables ("1" if the characteristic is present; "0" otherwise) and represent the incremental bid given the presence of the characteristic.

The data set used in this study is composed of a sample of 1,765 houses sold through the Multiple Listing Service in the Baltimore Metropolitan Area in the first week of each month of 1978. It provides a good, although not excellent, set of house characteristics.Most observations exclude a usable value for square feet of living space, for example. Also, a realtor sample probably undersamples housing in the inner cities and units serving the poor. The list and mean values of structural components are enumerated in Table 1.1-1.3 and are discussed in more detail in the empirical section of the paper.

Neighborhood components present a more difficult problem in model specification. For a characteristic such as racial composition to carry a meaningful hedonic price, buyers must observe it in a market setting. More specifically, the proper level of aggregation must define an entire neighborhood, rather than just the block upon which the house is located or the block face that faces it. Goodman (1977) discusses the differences among socioeconomic census data aggregated at the census tract, block group and block levels and opts for the block groups on "goodness of fit" criteria.

The census-defined neighborhoods are clearly arbitrary, however.

Conceptually, this exercise requires a rigorous delineation of individual neighborhoods. Within and among neighborhoods, a set of meaningful characteristics should be compiled and measured. These measures would serve as arguments in the hedonic price function representing neighborhood quality. Even these would probably require supplementing; school quality, for example, measured for the neighborhood grade school, junior high school and high school may transcend normal neighborhood boundaries.

The above task is an enormous undertaking. Existing data sources, consisting of census data and neighborhood measures of crime, education, and public services, are customarily used, appealing to Schelling's assertion that "a resident can be conscious of different neighborhoods in which he lives, the definitions corresponding to work, play, school, travel, social activities, civic and cultural identity, sentimental attachment, and (especially for homeowners) financial interest" (see Schelling, 1972, p. 176). How well they measure true neighborhood quality is indicated by the results of the analysis. Plausible values of neighborhood coefficients imply that the variables used are either important in and of themselves, or are at least good proxies for what really is important.

The available 1980 Census counts provide only racial data; these are assigned to each house at the tract level. Both 1980 percentage black and change in black share (1970 to 1980) are used. (The precise variable used is (BLACK 80—BLACK 70)/(TOTAL 70 + TOTAL 80)/2.) However, no income, education or employment data are currently available, so it is probable that use of racial variables alone imparts an omitted variables bias. This is discussed further in the interpretation of the hedonic price regression coefficients.

School and crime variables are appended geographically by elementary school district (21 variables) and by crime reporting area (12 variables). Although they do not provide an exhaustive set of non-racial neighborhood components relevant to the individual buyer, they provide insights into categories that are thought to be important.

Neither type of variable is novel in this type of analysis, but the breadth of coverage is unusual. Many studies have used educational variables (see Freeman (1979) for a partial review). Only three studies presently available treat crime specifically. Kain and Quigley (1970) find little impact on individual house values. Hellman and Naroff (1979) and Frisbie, et al. (1979) look at the determinants of grouped house values. Although the latter study uses several crime rates, it provides very weak controls on structural variables and the grouping suppresses within-group variation, artificially increasing explanatory power.

Data reduction procedures are used on these characteristics for two reasons. First, it is apparent from the correlation matrices that severe multicollinearity exists and could hamper the estimation of individual coefficients if the entire sets of characteristics are included in the hedonic price regressions. (These matrices are available from the authors on request.) Even more important, it is unlikely that households evaluate either the school or the crime situation using twenty or thirty different measures. Rather, it is more likely that fewer dimensions, summarizing the many, enter the household's decision making. Kain and Quigley (1970) use factor analysis to reduce a set of 39 housing characteristics to five factors. King (1974) and Little (1976) use principal components analysis in more explicit treatment of neighborhood characteristics.

Estimation of a set of principal components for a category of neighborhood quality does not preclude evaluating the impact of a given neighborhood characteristic. Consider a very simple example of a linear hedonic price model in variables x_1 and x_2,

$$P = \beta_1 X_x + \beta_2 X_2 \tag{3}$$

in which x_1 is formed from the principal components equation

$$x_1 = \Sigma \gamma_i y_i \tag{4}$$

An increase of one unit of y_1 leads to an increase λ_1 units of x_1, which leads to an incremental increase of $\beta_1\gamma_1$ dollars in house price. This is a simple use of the "chain rule" of differentiation and allows the imputation of hedonic prices for all characteristics of the bundle, if desired, rather than simply the several principal components that are developed.

DATA

Several authors have argued that housing market functions should be stratified into separate compartmentalized submarkets within urban areas. (Straszheim, 1974, was among the first to discuss this hypothesis.) This is especially important if housing types vary significantly from area to area, or if buyers are constrained by workplace, income, racial discrimination or search costs from participating in all housing submarkets. The Baltimore metropolitan area is unusual in that it is composed almost entirely of only two municipalities, Baltimore City (henceforth City) and Baltimore County (henceforth County) which surrounds but does not contain

TABLE 1.1

Neighborhood Variables--Education

	City		County	
	Mean	Standard Deviation	Mean	Standard Deviation
ENROLMNT	663.07	271.37	532.12	147.95
PSRATIO	18.52	2.57	18.41	1.61
ATTEND	93.32	2.10	96.43	.77
TEACHEXP	10.68	2.12	10.54	1.81
ADMINEXP	26.12	6.08	23.62	5.42
PMASTERS	19.65	6.11	41.11	10.48
SAS3	92.61	5.95	106.37	4.68
VOCGE3	3.31	.68	4.07	.43
READGE3	3.11	.55	4.10	.44
LANGGE3	3.59	.55	4.53	.51
MATHGE3	3.19	.45	4.07	.36
SAS5	93.45	5.69	108.22	4.78
VOCGE5	4.64	.73	5.50	.59
READGE5	4.48	.59	5.70	.51
LANGGE5	4.88	.63	6.06	.55
MATHGE5	4.73	.52	5.91	.51
VGEDIFF	1.33	.57	1.44	.44
RGEDIFF	1.37	.45	1.60	.29
LGEDIFF	1.28	.43	1.53	.32
MGEDIFF	1.54	.36	1.84	.30
SASDIFF	.84	4.05	1.85	3.62

TABLE 1.2

Neighborhood Variables--Crime

	City		County	
	Mean	Standard Deviation	Mean	Standard Deviation
MURDER	.32	.71	.02	.15
RAPE	.94	1.44	.14	.44
ROBTOT	14.71	14.49	.69	1.37
AGGASS	9.86	11.33	3.04	4.55
BRGTOT	26.63	24.12	7.31	8.93
LAACC	10.43	10.28	5.34	7.47
LFAUTO	9.86	8.39	2.02	2.82
LSHOP	7.23	29.58	2.52	15.11
LPURSE	2.41	3.34	.06	.64
LBIKE	3.87	4.94	1.75	2.60
ATHEFT	9.97	11.03	2.21	3.71
LOTHER	22.90	23.16	5.93	7.44
	N = 589		N = 1178	

TABLE 1.3

Structural Variables

	City		County	
	Mean	Standard Deviation	Mean	Standard Deviation
DWELCODE	.25	.43	.59	.49
NROOMS	5.19	1.11	5.18	.96
POOL	.006	.08	.02	.15
PATIO	.09	.28	.17	.37
AIRCOND	.10	.31	.36	.48
NSTORIES	2.10	.41	1.90	.59
GARAGE	.32	.79	.37	1.10
NBATHRMS	1.62	.74	1.59	.59
DIST	21.75	7.87	39.93	9.31
AGE	39.53	20.81	21.62	13.86
PBLACK	32.62	33.90	26.62	39.97
PCHANGE	11.86	19.36	4.39	11.56
	N = 888		N = 877	

or administer Baltimore City. (A third contiguous municipality, Anne Arundel County, shares only a small border and is not included in the analysis.) Baltimore City is considerably blacker and poorer than Baltimore County, has a more limited variety of housing stock and has double the property tax. Although formal t-tests are not displayed, they show significant differences in levels of many housing variables between the two municipalities, as noted in Table 1.

Houses selling in the City appear to have approximately the same number of rooms and bathrooms as do those in the County. They are considerably older than County houses (39.53 versus 21.62 years) and are considerably more likely to be attached to others (.59 versus .25). County houses are more likely to have the "quality" characteristics such as patio, pool and air conditioning.

Education variables also show differences. Enrollment in elementary schools is higher in the City, and percentage of faculty with masters' degrees is lower, although other administrative variables appear to be comparable. Test scores are generally higher in the County schools, as are the derived increments from the third to fifth grade levels. The crime measures are considerably higher in the City than in the County, although these are raw numbers, collected over areas that vary in size.

PRINCIPAL COMPONENTS

Separate principal component analyses were performed for the City and the County on both the crime and the school data. The 1,178 crime-reporting areas in the County tend to be smaller than the 589 areas in the City. Because the data set consists of frequency counts by reporting area, elements are not directly comparable. Therefore, separate principal component analyses were performed for the City and County crime data. The results of the analyses are presented in Table 2. (Components with eigenvalues of the unrotated matrix greater than 1 were retained.)

For the City, three factors account for 71% of the total variance. Table 2.1 shows that burglary, larceny of auto accessories, bicycle theft and auto theft load heavily onto C1CITY. Larceny from auto loads moderately on this factor. These are all nonviolent crimes involving property loss. Rape, robbery, and aggravated assault load heavily on C2CITY. Murder has a high loading while burglary and purse snatching have moderate loadings. C2CITY, then, appears to represent violent crimes. Shoplifting loads heavily on C3CITY with purse snatching and larceny "other" having high loadings. These crimes also involve property loss. One difference between the C1CITY and the C3CITY crimes appears to be that the latter may tend to occur most often in shopping centers and other commercial areas. This is probably not true of C1CITY crimes.

The three principal components for the County crime data (accounting for 65% of the total variance) are presented in Table 2.2. These factors are more difficult to interpret than those of the City. Aggravated assault, burglary, and larceny of auto accessories all load heavily on C1COUNT. Larceny from auto, bicycle theft, and auto theft have high loadings while rape and larceny "other" have moderate loadings. Shoplifting and purse snatching load heavily on C2COUNT. Larceny "other" has a high loading while robbery and auto theft have moderate loadings on this factor. Murder is the only crime to have a loading greater than .5 on C3COUNT. C2COUNT and C3CITY seem to be quite similar, especially if the moderate loadings are ignored. Thus C2COUNT may be interpreted as shopping center crime. C3COUNT is murder, and C1COUNT is all other crimes. All of the crime factors are expected to have negative coefficients in the hedonic regressions.

TABLE 2.1

Neighborhood Variable Factor Loadings

	City Crime		
	C1CITY	C2CITY	C3CITY
MURDER	-.03665	.67909	-.02734
RAPE	.27021	.71937	.01644
ROBTOT	.30523	.72776	.43884
AGGASS	.34663	.79277	.13487
BRGTOT	.71435	.50631	.14029
LAACC	.85375	.08069	.20877
LFAUTO	.58751	.36851	.36563
LSHOP	.13760	-.14746	.85899
LPURSE	.05821	.55667	.65169
LBIKE	.82001	.03777	.13894
ATHEFT	.82383	.21695	.04336
LOTHER	.47296	.32869	.66765

TABLE 2.2

Neighborhood Variable Factor Loadings

	County Crime		
	C1COUNT	C2COUNT	C3COUNT
MURDER	.06813	-.00052	.96827
RAPE	.55439	-.18065	-.20387
ROBTOT	.32175	.56032	.14427
AGGASS	.81212	.12873	.16987
BRGTOT	.84438	.06556	.07755
LAACC	.72212	.45213	.03945
LFAUTO	.65886	.43326	.03298
LSHOP	.02252	.89895	-.03334
LPURSE	-.03738	.81785	-.01740
LBIKE	.60597	.15135	.04941
ATHEFT	.63471	.58028	.02262
LOTHER	.50711	.65724	.03000

TABLE 2.3

Neighborhood Variable Factor Loadings

	City Education					
	E1CITY	E2CITY	E3CITY	E4CITY	E5CITY	E6CITY
ENROLMNT	-.05678	-.12660	-.55815	-.09771	-.02111	.65788
PSRATIO	-.00671	.04550	.25620	.03248	.18804	.84561
ATTEND	.12995	-.02332	.71647	.25411	.20825	.02914
TEACHEXP	-.03984	.07276	.13289	.14111	.76220	.05412
ADMINEXP	.10810	.09245	.71474	-.22664	-.09662	.10992
PSMASTERS	.14375	-.08660	-.08416	-.35723	.62402	.13713
SAS3	.81105	.07766	.09496	-.41941	-.09579	.04813
VOCGE3	.81196	-.20805	-.04249	.00945	.19088	-.09262
READGE3	.91010	-.27326	.01098	.04917	.07745	-.05054
LANGGE3	.92735	-.20955	.12503	-.09878	.03503	.04646
MATHGE3	.93185	-.19471	.07397	-.06712	-.07471	.01194
SAS5	.85485	.17330	.10974	.20618	-.12826	.05367
VOCGE5	.75675	.43683	-.00249	-.05166	.07219	-.06648
READGE5	.85904	.39722	.05171	.00298	-.03533	.04853
LANGGE5	.85713	.37620	.14904	.10240	.14230	-.01710
MATHGE5	.85864	.41122	.05447	.03954	.05066	-.06600
VGEDIFF	-.00184	.81352	.04795	-.07796	-.13669	.02580
RGEDIFF	.01472	.84203	.05359	-.05532	-.13884	.12351
LGEDIFF	.06512	.82186	.05795	.27741	.16384	-.08487
MGEDIFF	.08885	.84156	-.01288	.14126	.16691	-.11100
SASDIFF	.01029	.12959	.01478	.90592	-.03962	.00476

TABLE 2.4

Neighborhood Variable Factor Loadings

	County Education				
	E1COUNT	E2COUNT	E3COUNT	E4COUNT	E5COUNT
ENROLMNT	-.02328	-.02480	.45517	.71747	-.12677
PSRATIO	.01334	-.02590	-.01172	.88760	.09159
ATTEND	.40526	-.01658	.35677	-.09899	.64216
TEACHEXP	.37225	.05556	.36615	-.45160	.07276
ADMINEXP	-.09980	.00114	.79915	-.03633	.02033
PSMASTERS	.24326	.07728	-.57971	-.21086	-.02201
SAS3	.85870	.11978	-.01919	-.10699	-.28597
VOCGE3	.85159	-.11911	-.19776	-.08085	.16394
READGE3	.97241	-.01271	-.04691	-.04597	.00837
LANGGE3	.94636	-.06250	.03216	-.00686	.02660
MATHGE3	.93227	.00837	-.08473	-.00410	.01599
SAS5	.71328	.50154	-.15358	-.04458	.28392
VOCGE5	.81399	.48338	-.02363	-.08655	.02461
READGE5	.83931	.48755	-.02491	-.03670	.13317
LANGGE5	.85417	.42631	-.04259	-.04086	.06050
MATHGE5	.84984	.44991	-.10520	-.00393	.04770
VGEDIFF	.26524	.76526	.16061	-.03763	-.12642
RGEDIFF	.02045	.88832	.02678	.00433	.22469
LGEDIFF	-.02836	.85096	-.12694	-.06077	.06338
MGEDIFF	.30816	.74308	-.07481	-.00168	.06073
SASDIFF	-.16955	.50754	-.17806	.07964	.74544

The school data set consists of school attributes and student test scores for each elementary school in the study area. Thus, there are no comparability problems across the City/County boundary. The principal component analysis was therefore performed three ways: once for the entire area; once for the City; and once for the County schools. The separate results for the City and for the County are presented in Tables 2.3 and 2.4.

Analysis for the entire study area yielded five components, accounting for 79% of the variance. Percent masters and all of the achievement test scores load heavily on Factor 1. Attendance also has a high loading on this factor, which can be plausibly interpreted as representing *student performance*. This aspect of student performance is, however, likely to be heavily influenced by student quality. With the exception of the SAS test, all of the test difference scores load heavily onto Factor 2. This might represent a second aspect of student performance, one that reflects *school quality* rather than student quality. School enrollment and pupil/staff ratio load heavily onto Factor 3. This factor may be interpreted as representing *school size*. The *SAS difference* variable is the only one to load on Factor 4. Administrator experience and teacher experience both load on Factor 5, causing it to be interpreted as *staff experience*.

The factors change only slightly when the City and County schools are analyzed separately. In the City (Table 2.3), six factors were extracted, accounting for 79% of the variance. E1CITY may be interpreted as the student quality aspect of student performance. E2CITY may be interpreted as the school quality aspect of student performance, and E3CITY as administrator experience. The only variable loading on E4CITY is SASDIFF. E5CITY is teacher experience and E6CITY is school size.

For the County (Table 2.4), five factors were extracted, accounting for 79% of the variance. Once again, the first two factors (E1COUNT and E2COUNT) refer to student and school quality. E3COUNT is interpreted as staff experience, E4COUNT as school size and E5COUNT as SASDIFF.

With the exception of school size, all of the factors measure unambiguously positive attributes of school quality, and thus are expected to have positive coefficients in the hedonic regressions. There is some evidence in the public expenditure literature (see Hirsch, 1973, for example) that up to a point increased school size is beneficial; past that point it is disadvantageous. As a result, the sign for school size may be either positive or negative.

HEDONIC PRICE REGRESSIONS

Three hedonic price regressions are presented and interpreted in this section. Their fit is generally good, explaining more than 60% of the regression variance. Most variables have the correct signs and the interpretation of coefficients is straightforward.

Examination of the correlation matrix (Table 3) of neighborhood characteristics shows these variables to be largely uncorrelated with each other. Whereas researchers fear possible multicollinearity with census tract level variables, only one pair in the City has a correlation greater than .50 (PBLACK and C2CITY), and only three others have correlations greater than .30. No pair of variables in the County has a correlation greater than .25. Although "folklore" might imply that increased crime, deteriorating education and racial change progress together at the neighborhood level, the data provide little evidence.

Regression CITY1 is displayed in Table 4, columns A and B. The preferred functional form uses the logarithm of house price as the dependent variable (λ equals 0). Column A gives the unstandardized coefficient of the regression; Column B provides the hedonic price of the variable, $\beta_x P^{1-\lambda}$, where x is the variable, and P = bundle price. All hedonic prices are evaluated at the mean house price of $34,650.

All else equal, a detached house sells for approximately 21.6% more than a group house in Baltimore City. Additional rooms are worth approximately $1,400 more, and additional stories approximately $2,000 more. Garages and bathrooms have the expected positive values, while each house loses about $94 in value with each additional year of age. This may be an underestimate, since many houses in the City have been renovated partially or entirely on the inside, and are thus much "younger" than their stated ages.

The crime coefficients are substantial and significant. A one unit change in C1CITY, representing non-violent property loss, implies a $795 fall in house price. C2CITY, the violent crime component, implies a loss of $3,143, and C3CITY, or "shopping center crime" implies a price decrease of $3,721. Kain and Quigley (1970), for example, find little impact of crime on individual house prices. They attribute their insignificant coefficients and measurement problems to lack of variation from the limitation of their sample to the central city of St. Louis. In this case, either sufficient variance does exist, or, in fact, the variables available do a better job in discriminating the various levels of crime.

Four of the six education quality measures have the correct sign. A one unit increase of school quality (E1CITY) implies a $2,253 increase in house price. E2CITY, E4CITY and E6CITY are also positive, although only E6CITY is significantly so (implying a $1,217 increase in house price).

Regressions for Baltimore County (COUNTY1 and COUNTY2) are displayed in columns C through F. COUNTY1 is estimated in log-linear form for comparability to CITY1, although COUNTY2, with the dependent variable in "square-root inverse" form is, strictly speaking, the better estimator, and is discussed below. By inspection, the structural variables do about as well, the educational variables a little better, and the crime variables not quite as well as the City regression with respect to posited signs and magnitudes. Hedonic prices are calculated using mean bundle price of $52,680.

Garage spaces, number of bathrooms and status as a detached house provide about the same percentage premiums in the County as they do in the City, although the hedonic prices are higher since houses in the County are generally worth half again as much as those in the City. The variable NSTORIES, although signif-

TABLE 3

Correlations of Neighborhood Variables

	City										
	1	2	3	4	5	6	7	8	9	10	11
1. E1CITY	1.00	--	--	--	--	--	--	--	--	--	--
2. E2CITY	.02	1.00	--	--	--	--	--	--	--	--	--
3. E3CITY	-.20	-.09	1.00	--	--	--	--	--	--	--	--
4. E4CITY	.05	.03	.10	1.00	--	--	--	--	--	--	--
5. E5CITY	.11	.06	.09	.08	1.00	--	--	--	--	--	--
6. E6CITY	-.14	.10	.25	.19	.06	1.00	--	--	--	--	--
7. C1CITY	.03	.02	.04	.15	.06	.11	1.00	--	--	--	--
8. C2CITY	-.44	-.05	-.04	-.08	.04	-.01	-.19	1.00	--	--	--
9. C3CITY	-.26	-.10	.05	-.04	-.08	.00	-.16	.39	1.00	--	--
10. PBLACK	-.48	-.12	.10	-.03	.28	.20	.01	.58	.19	1.00	--
11. PCHANGE	-.03	.05	.23	.10	.04	.40	.11	-.02	.21	.18	1.00

	County									
	1	2	3	4	5	6	7	8	9	10
1. E1COUNT	1.00	--	--	--	--	--	--	--	--	--
2. E2COUNT	-.04	1.00	--	--	--	--	--	--	--	--
3. E3COUNT	-.07	.02	1.00	--	--	--	--	--	--	--
4. E4COUNT	.03	-.02	-.11	1.00	--	--	--	--	--	--
5. E5COUNT	.16	-.04	.04	-.02	1.00	--	--	--	--	--
6. C1COUNT	-.23	-.17	.03	.02	-.16	1.00	--	--	--	--
7. C2COUNT	-.01	.07	.02	.02	-.04	.23	1.00	--	--	--
8. C3COUNT	-.03	.04	.03	.03	-.04	-.08	.16	1.00	--	--
9. PBLACK	-.10	-.07	-.08	.01	.02	.00	.08	-.04	1.00	--
10. PCHANGE	-.17	-.03	.01	-.15	-.03	.04	-.05	.03	.74	1.00

TABLE 4

Hedonic Price Regressions--City and County

	a	b	c	d	f	e
	CITY1	Hed Price	COUNTY1	Hed Price	COUNTY2	Hed Price
DWELCODE	.1958	7.494	.1931	11.221	.02929*	13.276
NROOMS	.04030*	1.396	.06895*	3.632	.009909	3.788
POOL	-.003060	- .106	.1099*	6.120	.006197	2.452
PATIO	.1226*	4.519	-.01588	- .830	- .0005066	- .193
AIRCOND	.2262*	8.795	.02699	1.441	.003794	1.481
NSTORIES	.05704*	1.976	-.05192*	2.735	- .005532*	-2.115
GARAGE	.1341*	4.647	.1422*	7.491	.01463*	5.594
NBATHROOM	.1755*	6.081	.1692*	8.913	.01958*	7.487
DIST	.02204*	.210	-.007308	.104	- .0008309	.098
$DIST^2$	-.0003678	--	.0001161	--	.00001360	--
AGE	-.002709*	- .094	-.003562*	- .188	- .0005480*	- .210
E1CITY	.06501*	2.253	--	--	--	--
E2CITY	.02232	.773	--	--	--	--
E3CITY	-.02507*	- .860	--	--	--	--
E4CITY	.01336	.463	--	--	--	--
E5CITY	-.04499*	-1.558	--	--	--	--
E6CITY	.03513*	1.217	--	--	--	--
E1COUNT	--	--	.09180*	4.836	.01250*	4.779
E2COUNT	--	--	.02501*	1.318	.003244*	1.240
E3COUNT	--	--	.001707	.089	.0003342	.128
E4COUNT	--	--	-.02321*	-1.223	- .002255*	- .862
E5COUNT	--	--	.000803	.042	.0001951	.074
PBLACK	-.002859*	- .100	-.000408	- .021	- .00002951	- .011
PCHANGE	.006112*	.211	-.001962*	- .103	- .0002920*	- .111

TABLE 4 (Continued)

	a	b	c	d	f	e
	CITY1	Hed Price	COUNTY1	Hed Price	COUNTY2	Hed Price
C1CITY	-.02293*	- .795	--	--	--	--
C2CITY	-.09071*	-3.143	--	--	--	--
C3CITY	-.1074*	-3.721	--	--	--	--
C1COUNT	--	--	-.005167	- .272	- .001714**	- .655
C2COUNT	--	--	-.007184	- .038	.001517	.580
C3COUNT	--	--	-.02128*	-1.121	- .003098*	-1.184
ZONE1CIT	.3676*	15.394	--	--	--	--
ZONE2CIT	.3792*	15.978	--	--	--	--
ZONE3CIT	.2920*	11.750	--	--	--	--
ZONE1COU	--	--	.08202*	4.501	.01747*	7.373
ZONE2COU	--	--	.04899	2.645	.01017*	4.115
ZONE3COU	--	--	.1163*	6.497	.02008*	8.607
ZONE4COU	--	--	.01193	.632	.006341	2.511
CONSTANT	2.0820	--	3.2832	--	1.6200	--
λ	0.0***	--	0.0	--	-0.5***	--
$\bar{R}^2$	.637	--	.655	--	.683	--
S.E.E.	.3663	--	.2189	--	.02732	--

*Significant at 5% level ($t > 1.96$)

**$t = 1.81$

***Significant at 5% level ($\frac{1}{2}\chi^2(1) > 1.92$)

icantly positive in the City, is significantly negative in the County. In the City, 75% of the sample houses are attached; an additional floor probably implies more room in this context. In the County, however, only 41% of the sample houses are attached; NSTORIES, here, could imply the disadvantage of having to use stairs, rather than having the entire house on one floor.

Two of the three crime components for the County have the correct signs but only C3COUNT (murder) is significant, yielding a decline in house price of $1,184. C1COUNT (with a t-statistic of 1.81) implies a discount of $655. In general, the magnitudes of these coefficients are not comparable to City coefficients. Perhaps purchasers in the County do not see crime differences as major problems. More probably, however, there is a lack of significant variation necessary to estimate the hedonic prices properly. Baltimore County is basically a large suburban area, and it is possible that the crime variable does not vary enough to allow meaningful hedonic price estimation.

Valuation of educational quality in the County is substantial.

A one unit increase in E1COUNT implies close to a $4,800 increase in house value; a one unit increase in E2COUNT brings house prices up by $1,318. All variables have correct signs and three of the five are significant.

As noted above, the "chain rule" method can be used to examine the impacts of individual measures summarized in the principal components vectors. Using the notation described in equation (4), the incremental value of measure y_1, summarized in component x_1 is

$$\frac{\delta P}{\delta y_i} = P^{1-\lambda}(\Sigma_k \beta_k \gamma_{ki}) \qquad (5)$$

where γ_{ki} refers to the factor score of the ith measure in the kth factor. A one standard deviation increase in third-grade mathematics score (MATHGE3), for example, is associated with a $537 price increase in the County and a $594 price increase in the City

The racial variables in the two regressions give provocative results. In CITY1, an additional percentage point black population implies a $100 fall in house price; in the County (COUNTY2), the fall is $11 and the coefficient is not significant. Although the analysis attempts to control for house and neighborhood quality, it does not include neighborhood socioeconomic variables such as income, education or employment status. To the extent that lower values of these variables (reflecting neighborhood disamenities) are correlated with percentage black, its coefficient will be biased downward. This bias may be larger in the City because of a larger variation in socioeconomic variables than the more suburban County. This "omitted variables" problem confounds any serious attempt to test hypotheses about racial discrimination or prejudice; they await the availability of more detailed Census data.

The variable PCHANGE is designed to measure neighborhood racial stability. In the City, all else equal, a house in a neighborhood that has gone from 10% to 20% black over the past decade is worth about $2,100 more than a house in a neighborhood that has remained 20% black. In the County, the same circumstances lead to a decline in price by $1,100.

This may result from the submarket segmentation discussed above. The black population in Baltimore City grew by approximately 150,000 between 1970 and 1980. The subsequent search for additional black housing may have led to increased demand in white neighborhoods, both in the City and in the County. Although some small amount of renovation, rehabilitation and new construction did occur in the City, it was probably not sufficient to meet the increased demand. Market segmentation due to work-place constraints, search costs or discrimination may have perpetuated the higher prices. On the other hand, considerable new construction occurred in the more suburban County at the same time as this increased demand; enough, perhaps, to keep prices stable or even to lower them.

CONCLUSION

This article has demonstrated how economic theory and econometric techniques can be used to impute a valuation to a good that does not explicitly enter the market, that is, neighborhood. This is not novel, as many other studies have examined some number of neighborhood characteristics in this framework. More important here is the specification of several variables to describe education (six in the City and five in the County), and three variables to describe crime. Most have the correct sign and plausible magnitudes. The education, crime and racial variables exhibit only modest multicollinearity. This allows a reasonably good measure of individual effects; it implies, however, that proxy variables for many neighborhood effects may be difficult to find.

Economic theory suggests that the hedonic prices facing homeowners can be used to reflect, in a rough way, the benefits of improvements in education or decreases in crime. An increase of one unit of E1CITY implies an increase of $2,253 in house value. If there are 1,000 houses in the improved area this implies an increase in property values of approximately $2.25 million, reflecting the long-term stream of benefits. If the long-term stream of costs is less than this, the one unit improvement, by benefit-cost criteria will be worthwhile. The increased home values may also lead to increased property tax collections.

This analysis must be modified severely to consider large area improvements, and the property value increments must be analyzed particularly carefully. (An improvement in education over a wide area may so increase the availability of the characteristics that its market valuation, and the values of houses that have it, may even fall. Lind, 1973, provides the necessary theoretical analysis and Freeman presents examples.) In addition, it is unclear which policy imputs can be used to control the outputs that enter into the education and crime characteristics. Nonetheless, this method may provide a valuable framework for approximating the results of programs with neighborhood impacts and for increasing the societal benefits achievable with limited resources.

REFERENCES

Box, G.E.P., & Cox, D.R. An analysis of transformations. *Journal of the Royal Statistical Society*, 1964, *B26*, 211-43.

Freeman, A.M. The hedonic price approach to measuring demand for neighborhood characteristics. In D. Segal (Ed.), *The economics of neighborhood.* New York: Harcourt, Brace, Jovanovich, 1979.

Frisbie, D.W., Fishbine, G., Hintz, R., Joelson, M., Nutter, J.B., & Gray, C.M. Crime in Minneapolis. Minnesota Crime Prevention Center, Minneapolis, 1978.

Goodman, A.C. A comparison of block group and census tract data in a hedonic housing price model. *Land Economics*, 1977, *5*, 483-87.

Goodman, A.C. Hedonic prices, price indices and housing markets. *Journal of Urban Economics*, 1978, *5*, 471-84.

Hellman, D.A., & Naroff, J.L. The impact of crime on urban residential property values. *Urban Studies*, 1979, *16*, 105-12.

Hirsch, W.Z. *Urban economic analysis.* New York: McGraw Hill, 1973.

Kain, J.F. & Quigley, J.M. Measuring of value of housing quality. *Journal of the American Statistical Association*, 1970, *65*, 532-48.

King, A.T. *Property taxes, amenities and residential land values.* Cambridge, MA: Ballinger, 1974.

Lind, R.C. Spatial equilibrium, the theory of rents and the measurement of benefits from public programs. *Quarterly Journal of Economics*, 1973, *87*, 188-207.

Little, J.T. Residential preferences, neighborhood filtering and neighborhood change. *Journal of Urban Economics*, 1976, *3*, 68-84.

Rosen, S. Hedonic prices and implicit markets: Product differentation in pure competition. *Journal of Political Economy*, 1984, *82*, 34-55.

Schelling, T. A process of residential segregation: Neighborhood tipping. In Anthony H. Pascal (Ed.), *Racial discrimination in economic life.* Lexington, MA: D.C. Heath, 1972.

Segal, D. *The economics of neighborhood.* New York: Harcourt, Brace, Jovanovich, 1979.

Straszheim, M. Hedonic estimation of housing market prices: A further comment. *Review of Economics Land Statistics*, 1974, *56*, 404-06.

APPENDIX

Explanation of Variables

Housing Structure Variables	
NROOMS	Number of Rooms
POOL	Pool
PATIO	Patio
AIRCOND	Air conditioning
NSTORIES	Number of stories
GARAGE	Size of garage: 0 = no garage; 1 = 1-car garage; 2 = 2-car garage; etc.
NBATHRMS	Number of bathrooms
DIST	Distance from CBD (1,000's of ft.)
AGE	Age of structure
Education	
ENROLMNT	Student enrollment
PSRATIO	Pupil to staff ratio
ATTEND	Percent average daily attendance
TEACHEXP	Average teacher experience
ADMINEXP	Average administrator experience
PMASTERS	Percent of staff with masters degree or above
SAS3	Average SAS score—third grade
VOCGE3	Vocabulary average grade equivalent score—third grade

READGE3	Reading comprehension average grade equivalent score—third grade
LANGGE3	Language total average grade equivalent score—third grade
MATHGE3	Mathematical total average grade equivalent score—third grade
SAS5	Average SAS score—fifth grade
VOCGE5	Vocabulary average grade equivalent score—fifth grade
READGE5	Reading comprehension average grade equivalent score—fifth grade
LANGGE5	Language total average grade equivalent score—fifth grade
MATHGE5	Mathematical total average grade equivalent score—fifth grade
SASDIFF	SAS5 - SAS3
VOLDIFF	VOCGE5 - VOCGE3
READDIFF	READGE5 - READGE3
LANGDIFF	LANGGE5 - LANGGE3
MATHDIFF	MATHGE5 - MATHGE3
EiCITY	ith education principal component in Baltimore City
EiCOUNT	ith education principal component in Baltimore County

Crime

MURDER	Criminal homicide	
RAPE	Rape	
ROBTOT	Total robberies	
AGGASS	Aggravated assault	
BRGTOT	Total burglaries	
LAACC	Larceny—auto accessories	
LFAUTO	Larceny—from auto	Total Larcenies
LSHOP	Larceny—shoplifting	
LPURSE	Larceny—purse snatching	
LBIKE	Larceny—bicycle	
LOTHER	Larceny—other	
ATHEFT	Auto theft	
CjCITY	jth crime principal component in Baltimore City	
CjCOUNT	jth crime principal component in Baltimore County	

Race

BLACK70 (BLACK80)	Black Population (census tract level) in 1970 (1980)
TOTAL70 (TOTAL80)	Total population (census tract level) in 1970 (1980)
PBLACK	BLACK80/TOTAL80
PCHANGE	(BLACK80 − BLACK70)/(TOTAL70 + TOTAL80)/2)