

Day, Brett; Bateman, Ian; Lake, Iain

**Working Paper**

## Estimating the demand for peace and quiet using property market data

CSERGE Working Paper EDM, No. 06-03

**Provided in Cooperation with:**

The Centre for Social and Economic Research on the Global Environment (CSERGE), University of East Anglia

*Suggested Citation:* Day, Brett; Bateman, Ian; Lake, Iain (2006) : Estimating the demand for peace and quiet using property market data, CSERGE Working Paper EDM, No. 06-03, University of East Anglia, The Centre for Social and Economic Research on the Global Environment (CSERGE), Norwich

This Version is available at:

<https://hdl.handle.net/10419/80288>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

**Estimating the demand for peace and quiet  
using property market data**

**by**

**Brett Day, Ian Bateman and Iain Lake**

**CSERGE Working Paper EDM 06-03**

**Estimating the demand for peace and quiet using property market data**

by

**Brett Day<sup>1</sup>, Ian Bateman<sup>1</sup> and Iain Lake<sup>2</sup>**

<sup>1</sup> Centre for Social and Economic Research on the Global Environment (CSERGE),  
School of Environmental Sciences, University of East Anglia, NR4 7TJ, UK

<sup>2</sup> Centre for Environmental Risk, School of Environmental Sciences,  
University of East Anglia, NR4 7TJ, United Kingdom.

**Author contact details**

email: [brett.day@uea.ac.uk](mailto:brett.day@uea.ac.uk)

tel: +44 (0)1603 591413

**Acknowledgements**

The support of the Economic and Social Research Council (ESRC) is gratefully acknowledged. This work was part of the interdisciplinary research programme of the ESRC Research Centre for Social and Economic Research on the Global Environment (CSERGE).

Funding for this research was provided by the Department for Environment Trade and the Regions (now Department for Transport). The quality of the output was greatly enhanced through the kind cooperation of the Land Registry and the Valuation Office Agency.

**ISSN 0967-8875**

**Abstract**

For many industrialised nations environmental noise is emerging as a local pollutant of major concern. Incorporating such concerns into policy guidance tools such as social cost benefit analysis requires estimates of the monetised benefits of noise reduction. Using a two-stage hedonic pricing methodology we estimate a system of structural demand equations for different sources of transport-related noise. Our application applies state-of-the-art econometric techniques. In the first stage, we identify market segments using model-based clustering techniques and estimate separate hedonic price functions for each segment. We spatially smooth the data to account for omitted spatial covariates and employ a semiparametric estimator to allow flexibility in functional form. In the second stage, we control for nonlinearity of the budget constraint and identify demand relationships using techniques that account for problems of endogeneity and censoring of the dependent variable. We report welfare estimates for peace and quiet that we believe to be the first derived from property market data in a theoretically consistent manner.

**Keywords:** Noise valuation, hedonic pricing, model-based clustering, partial linear model, spatial autocorrelation, demand system, simultaneous-equation Tobit.

**JEL Classifications:** Q51, C14, C21, C24

## 1. INTRODUCTION

For many industrialised nations environmental noise is emerging as a local pollutant of major concern. In the European Union (EU), for example, it is claimed that around 20% of the population are exposed to environmental noise levels that scientists and health experts consider unacceptable (EC, 1996). Indeed, the Directive on the Assessment and Management of Environmental Noise requires member states of the EU to draw up “action plans” designed to reduce noise where necessary and to maintain environmental noise quality where it is good (EC, 2002).

Of course, noise reduction measures are costly and establishing whether such measures are worthwhile requires comparison of these costs with the social benefits of a quieter environment. One framework within which this comparison can be made is social cost-benefit analysis (SCBA). Attempts to integrate noise considerations into SCBA have been hampered, however, by the absence of monetary valuations of the welfare benefits of noise reductions. For example, the UK in common with many other EU member states and other industrialised countries, undertakes SCBA of transport projects but fails to include within its analysis monetarised estimates of the noise (dis)benefits of projects (Odgaard *et al*, 2005; Grant-Muller *et al*, 2001; Hayashi and Morisugi, 2000). Whilst there is an aspiration to include noise (dis)benefits in this analysis (HM Treasury, 2003, Annex 2; DfT, 2004a), no reliable figures are available for that purpose.

This paper reports on research undertaken in the UK that attempts to estimate welfare values for the avoidance of transport-related noise in urban areas. Our research applies the method of hedonic pricing to property market data collected from the city of Birmingham in the UK. As is well known, hedonic pricing uses techniques of multiple regression to isolate implicit prices (i.e differences in the market price of property attributable to marginal differences in property characteristics). There already exists a large literature reporting on studies undertaken in various urban property markets that have identified implicit prices for noise (reviewed in Bateman *et al.*, 2001).

Like any market price, implicit prices for noise provide a monetary estimate of the welfare benefits of marginal changes in noise within that property market. However, like any market price, the implicit prices that arise in a property market reflect only the particular conditions of supply and demand that exist in that market. Accordingly, an implicit price for noise estimated in one urban area offers little indication of the benefits of changes in noise that are experienced in different urban areas. Far better would be to use the information provided by observations of households’ choices of noise exposure when faced by different implicit prices to identify the demand relationship for peace and quiet. Areas under such a demand curve provide valid approximations to the welfare benefits of non-marginal changes in noise exposure (Bartik, 1988) and, because they are based on underlying preferences and not market prices, indicate values that might be used in benefits transfer exercises.

Unfortunately, estimating demand relationships from hedonic pricing data is a theoretically and analytically challenging task.<sup>1</sup> Indeed, this study is one of only a handful that have attempted to estimate demand relationships using the hedonic pricing method in a theoretically consistent manner (other examples include, Bajic, 1993, Cheshire and Sheppard, 1998, Palmquist and Isangkura, 1999, Boyle *et al.*, 1999 and Zabel and Kiel, 2000). Moreover, as far as the authors are aware it is the only study to estimate demand relationships for avoidance of transport-related noise pollution using property price data.

---

<sup>1</sup> Haab and McConnell (2002) state that “With a few exceptions, researchers have abandoned any attempts to recover preferences, and work instead with the hedonic price function” (p. 251).

Our study, is unique in other ways. As briefly described in Section 3, our data is one of the largest and richest property market data sets yet compiled. We make use of numerous data sources to assemble details of the structural characteristics of properties and take advantage of the powers of geographical information systems (GIS) to construct variables that capture a diversity of features of properties' locations. In particular, our data set contains measures of noise exposure from road, rail and aircraft traffic, such that our study is the first, of which we are aware, to investigate these three sources of noise pollution simultaneously.

In addition, we investigate our data for evidence of market segmentation. To this end we apply techniques of model-based clustering, a statistically rich set of analytical tools that use the data itself to inform on the pattern of segmentation in the property market (Fraley and Raftery, 1998). These techniques have never previously been applied in this context and, as far as the authors can ascertain, neither have they been applied in any other field of applied econometrics.

Furthermore, we contribute to the substantial literature on the specification of hedonic price functions. In particular, we combine the partial linear specification of Anglin and Gencay (1996) with the smooth spatial effects estimator of Gibbons and Machin (2003). Accordingly, we employ a specification in which variables that we believe *a priori* to be key determinants of market price enter the hedonic price function nonparametrically. Moreover, we account for the possibility of omitted spatial covariates by spatially smoothing the data. Finally, we correct the standard errors of our parameter estimates for spatial autocorrelation using a two stage generalised moments estimator suggested by Kelejian and Prucha (1999).

Our demand analysis is also unique in that we estimate a demand system for more than one source of noise pollution. In so doing we, we correct the data for problems brought about by non-linearity of the budget constraint and apply Amemiya's generalised least squares estimator (Newey, 1987) to account for the endogeneity of implicit prices and the censoring of the quantity variable at the background level of urban noise.

The paper is organised as follows. Section 2 discusses the hedonic pricing method. Section 3 provides a brief description of our data set. Section 4 describes our econometric approach to estimating hedonic price functions and deriving the implicit price of noise. Section 5 reports on the results of this approach when applied to our data. Section 6 discusses our econometric approach to estimating demand relationships for peace and quiet. Section 7 reports on the application of this estimating approach to our data. Section 8 describes the welfare estimates derived from our demand analysis and Section 9 provides some concluding remarks.

## 2. THE METHOD OF HEDONIC PRICING

The problem with estimating a demand function for peace and quiet is that noise avoidance is not traded in its own market. Accordingly, we are unable to directly observe either market prices or associated levels of demand. Nevertheless, as first observed by Ridker (1967) and Ridker and Henning (1967), households reveal their preferences for the environmental quality of their residential location when purchasing a residence in the property market. Of course, properties are differentiated by more than just their exposure to noise. Indeed, property can be thought of as a multi-attributed commodity providing households with a variety of different services. Rosen (1974) endowed economists with a substantive theoretical framework within which to study market equilibria for such commodities. These equilibria are characterised by a market clearing price function, the hedonic price function (HPF), that relates the various attributes of a commodity to the price that commodity can fetch in the market. In the property market, for example, prices are determined by the qualities or characteristics of a property's structure, environs and location.

In the spirit of Rosen (1974) our estimation strategy involves two stages. In the first stage we estimate the market-clearing HPF;

$$P = P(\mathbf{z}) \tag{1}$$

where  $\mathbf{z}$  is a vector (with elements  $z_k$ ,  $k = 1, 2, \dots, K$ ) of property attributes and  $P$  is the price that a property with those attributes commands in the market. For each household in a sample, therefore, the implicit price of a property attribute can be calculated from the HPF by evaluating  $\partial P(\mathbf{z})/\partial z_k \equiv p_{z_k}$ . As in the standard model of consumer choice, Rosen showed that utility maximising households will choose a residential location where the level of each property attribute is such that their marginal willingness to pay (MWTP) for each attribute just equals its implicit price. Accordingly, the second stage of the estimation procedure is to regress observed levels of residential noise exposure on calculated implicit prices in order to estimate a demand function for peace and quiet.

There has been considerable debate as to whether the demand function can be identified using this two stage procedure using data from a single market (see Ekeland et al. 2002 for a detailed review). Whilst the literature on this subject has been voluminous and occasionally misguided the problem is in essence quite simple; to estimate a household's demand curve one must have information on at least two points along its length, data from a single market provides just one.<sup>2</sup> As a consequence, researchers attempting to identify demand functions for attributes must follow one of two avenues; either (1) impose structure on the model of household preferences and use the non-linearity of the HPF to afford identification (as per Ekeland et al., 2004) or (2) introduce exogenous price variation into the data by collecting observations from a variety of property markets each characterised by unique and differing HPFs (as per Brown and Rosen, 1982; Murray, 1982; McConnell and Phipps, 1987). In our application we follow the latter strategy.

The process of recovering an estimate of the demand curve is further complicated by the fact that in hedonic markets, prices per unit of a characteristic are not necessarily constant. Under such circumstances, households effectively simultaneously choose the price and quantity of the property characteristic such that prices must be treated as endogenous in the estimation process. More fundamentally, when marginal prices are not constant the standard duality results of consumer theory no longer hold. In particular, Marshallian demand

---

<sup>2</sup> Data from a single market provides no exogenous variation in price. Differences in observed choices must result solely from differences in demand shifters such that the price-quantity relationship cannot be identified.

functions do not necessarily manifest themselves as downward sloping curves relating prices to quantities. In our application, we follow the procedure proposed by Murray (1983) and Palmquist (1988) in which adjustments are made to the income argument in the demand relationship so as to estimate 'pseudo'-Marshallian demand functions whose properties resemble those of standard Marshallian demand functions in a world of parametric prices. We discuss this estimation strategy in Section 6 and its implications for calculating welfare measures in Section 9.



### 3. THE DATA

The research results presented in this paper are based on property purchases in the City of Birmingham. Birmingham is the UK's second largest city with a large and diverse housing stock and an equally diverse population. The property market is highly developed and sufficiently active to ensure that competition amongst vendors, brokers and purchasers enforces a market rigour that obliges properties to sell at a firmly established market price.

Details of the selling price, sale date and full address for each residential property transaction during 1997 were obtained from the databases of the UK Land Registry. Details of the structural characteristics of each property were obtained from the Valuation Office Agency (VOA). Amongst other details, the VOA provided data on the number of bedrooms and bathrooms in each property, total floor area, the property's age, whether the property was a bungalow or house (flats are not included in the analysis), whether the property was detached, semi-detached, in a terrace or at the end of a terrace and whether the property had central heating and access to off-road parking. Furthermore, the VOA classifies properties according to age and style of construction into one of around 30 property types called Beacon Groups that provide an additional indication of property quality.

Addresses were geolocated using a GIS. Subsequently GIS datasets were used to provide details of the garden area and to calculate straight line distances, car travel times and walking distances from each property to (dis)amenities including the central business district, primary schools, shops, railway stations, industrial sites and landfills. To further refine the accessibility measures, an index was constructed that combined the walking-distance proximity of local primary schools with measures of school quality. Figure 1 illustrates values for this index for properties in the dataset for one location in the study area. A similar index was constructed to describe proximity to local commercial centres based on the quantity and proximity of shops.

**Figure 1: Primary School accessibility scores for a selection of properties**

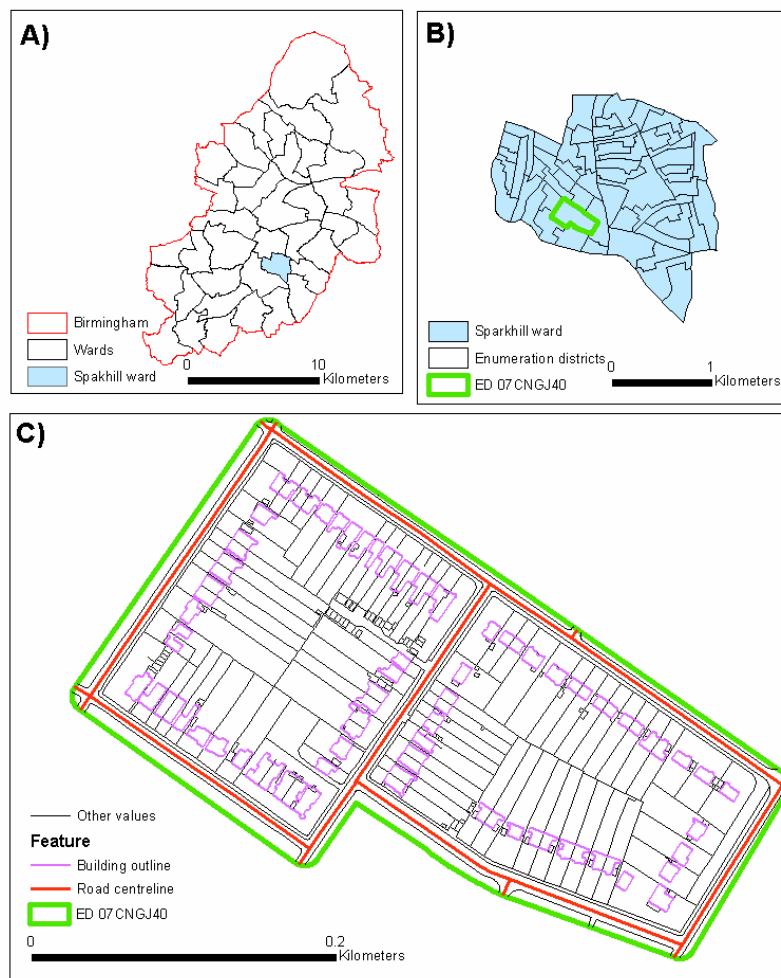


Using a GIS, data on land uses and the location and orientation of each property was combined with information on the landscape topology and building heights to calculate indices describing the views available from each property. View indices were constructed for recreational parkland and water surfaces.

Data pertaining to each property's exposure to daytime road, rail and air traffic noise measured in decibels LEQ were provided by the Department of Environment Food and Rural Affairs.

Data on the socioeconomic composition of property neighbourhoods were drawn from the 1991 UK census provided by the Office for National Statistics (ONS). The smallest area over which census data is provided by the ONS is an enumeration district (ED). Birmingham is divided into 1,940 EDs, with each ED containing an average of 191 households. EDs are gathered into larger scale political units known as wards of which Birmingham contains 39. The hierarchy of these administrative areas is illustrated in Figure 2.

**Figure 2: Hierarchy of administrative areas in Birmingham**



Census data provides hundreds of variables describing the socioeconomic characteristics of the households inhabiting each ED. We condense this excess of neighbourhood attributes into a more manageable set of indices using factor analysis. Following standard procedures we retain four factors that are interpreted as measuring the increasing; (1) wealth of households (2) presence of inhabitants from ethnic minorities (3) age of inhabitants, (4) presence of households with children.

Complete data records were successfully compiled for some 10,848 residential property transactions in Birmingham in 1997. A summary of the variables used in the analysis is provided in Appendix A.

#### 4. FIRST STAGE: ECONOMETRIC MODEL

It is almost universally acknowledged that urban property markets are not homogenous entities. Rather they are composed of a collection of distinct but related market segments. Watkins (2001) argues that segmentation of the market arises from the fact that households tend to fall into distinct “consumer groups” whilst the housing stock comprises a variety of distinct “product groups”. The matching of consumer groups to product groups gives rise to segmentation of the market whilst the differing conditions of supply and demand within each segment mean each is characterised by a unique HPF. Following Bajic (1993), we introduce price variation into the second stage estimation procedures by identifying market segments and fitting separate HPFs to each segment.

Our intention in this study is to identify segments as groups of properties with similar structural attributes inhabited by households with similar socioeconomic characteristics. One data-driven process by which such groups might be identified is known as *cluster analysis* (see Jain et al., 1999 for a recent review of this set of techniques). Indeed techniques of cluster analysis have previously been applied to the classification of properties into market segments; notably by Abraham et al. (1994), Goetzmann and Wachter (1995), Hoesli et al. (1997) and Bourassa et al. (1999). However, these studies use relatively simple clustering algorithms that provide no independent statistical indication of the nature or number of clusters to be found in the data. In this research we take advantage of recent advances in clustering techniques to answer both these important questions. In particular, we employ *model-based cluster analysis* (Banfield and Raftery, 1993; Fraley and Raftery, 1998).

The starting point for any cluster analysis is to define a set of  $P$  variables by which the data is to be clustered. Each observation can then be represented by a point in the  $P$ -space defined by those variables. The objective of the analysis is to identify the various groups of observations that form distinct clusters in this  $P$ -space. The fundamental assumption of model-based clustering is that each data point is drawn from a population of such points constituting all the members of the cluster. Moreover, the location, size and shape of this underlying population can be approximated by a probability distribution.<sup>3</sup> The data observed by the researcher is the composite of data points drawn from a finite number of such clusters.

To formalise, each  $P$ -dimensional data point  $\mathbf{x}$  arises from a super population comprising a mixture of  $M$  populations,  $C_1, C_2, \dots, C_M$ , in some proportions  $\pi_1, \pi_2, \dots, \pi_M$  respectively, where;

$$\sum_{j=1}^M \pi_j = 1 \quad \text{and} \quad \pi_j \geq 0 \quad (j = 1, 2, \dots, M) \quad (2)$$

If we assume that each population,  $C_j$ , can be modelled as a  $p$ -dimensional Gaussian distribution with mean vector  $\boldsymbol{\mu}_j$  and covariance matrix  $\boldsymbol{\Sigma}_j$ , then the probability density function (pdf) of an observation  $\mathbf{x}$  is of the form;

---

<sup>3</sup> Assuming a Gaussian distribution, for example, would imply that clusters are ellipsoidal. It would also assume that the likelihood of observing data points belonging to a particular cluster is greater near the mean location of that cluster than at its periphery.

$$\begin{aligned}
f(\mathbf{x}|\boldsymbol{\theta}, \boldsymbol{\pi}) &= \sum_{j=1}^M \pi_j f_j(\mathbf{x}|\boldsymbol{\theta}_j) = \sum_{j=1}^M \pi_j \phi_j^p(\mathbf{x}|\boldsymbol{\mu}_j, \boldsymbol{\Sigma}_j) \\
&= \sum_{j=1}^M \pi_j \left(2\pi^p |\boldsymbol{\Sigma}_j|\right)^{-\frac{1}{2}} \exp\left(-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_j)' \boldsymbol{\Sigma}_j^{-1}(\mathbf{x}-\boldsymbol{\mu}_j)\right)
\end{aligned} \tag{3}$$

where  $\boldsymbol{\theta} = (\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \dots, \boldsymbol{\theta}_M)$  is the vector of parameters associated with the assumed distributions of the  $M$  clusters,  $\boldsymbol{\pi} = (\pi_1, \pi_2, \dots, \pi_M)$  is the vector of mixing proportions,  $f_j(\mathbf{x}|\boldsymbol{\theta}_j)$  is the pdf of cluster  $C_j$  which is given the specific  $p$ -dimensional Gaussian form denoted  $\phi_j^p(\mathbf{x}|\boldsymbol{\mu}_j, \boldsymbol{\Sigma}_j)$ . Thus, in the Gaussian case  $\boldsymbol{\theta}_j$  comprises the elements of the vector  $\boldsymbol{\mu}_j$ , which determine the mean location of each cluster in  $p$ -space, and the distinct elements of the covariance matrices  $\boldsymbol{\Sigma}_j$ , which determine the geometric proportions of each cluster.

As described by Fraley and Raftery (1998), given some initial assumption concerning the number of clusters, equation (3) allows one to formulate a likelihood function for the clustering problem. Moreover, the location and shape parameters describing each cluster that maximise this likelihood can be estimated using a simple EM algorithm. In addition, the number of clusters best characterising the data can be identified by comparing models assuming different numbers of clusters using Bayes Information Criteria (BIC). Since we suspect that geographical location will play an important role in determining market segment membership, we make use of a procedure suggested by Posse (2001) and post-process the clustering classification to take account of the spatial information.

We estimate separate HPFs for each cluster identified in the data. Following standard practice (e.g. Allen et al., 1990) we assume that the properties in a cluster form a separate market segment if the HPF for that cluster differs significantly from those of the other clusters.

Unfortunately, economic theory provides little guidance as to how to specify the functional form of HPFs. Indeed, recent research simulating market-clearing HPFs using hypothetical data reveals that under all but the most contrived of assumptions, the equilibrium HPF will tend to be highly non-linear (Heckman et al., 2003). Nesheim (2002) even reports that for certain parameter values, a kinked price function is required in order to attain equilibrium. Moreover, it is widely recognised that location, or at least attributes of location, plays a major role in determining property prices. Even with the advances in data collation provided by developments in digital data manipulated in GIS, it seems unlikely that any data set will be sufficiently comprehensive to capture every aspect of location that might induce correlation in prices over space. Accordingly, our regression function must allow for non-linearities in known covariates and for the possibility of omitted spatial covariates.

Anglin and Gencay (1996) were the first to apply Robinson's (1988) *partial linear model* (PLM) in order to introduce flexibility into the functional form of hedonic equations. In the PLM part of the model is specified parametrically whilst the rest is estimated using non-parametric techniques. We allow for non-linearities by including in the nonparametric part of the model those variables that we believe *a priori* to be key determinants of market price.

Gibbons and Machin (2002) also exploit the PLM, but to account for omitted locational covariates. They include the coordinates of each observation in the nonparametric part of the model in what they label their *smooth spatial effects* estimator.

Here we combine the specifications of Anglin and Gencay (1996) and Gibbons and Machin (2002) by including both location coordinates and property characteristics in the nonparametric part of the PLM. Our specification is as follows;

$$\ln P_i = \mathbf{z}_i \boldsymbol{\beta} + q(\mathbf{x}_i, \mathbf{c}_i) + \varepsilon_i \quad (4)$$

where  $\mathbf{z}_i$  is a  $k$ -vector of regressors whose influence on property price is determined by the parameters  $\boldsymbol{\beta}$ , whilst  $\mathbf{x}_i$  is a  $p$ -vector of regressors and  $\mathbf{c}_i$  is the two-dimensional vector of coordinates whose joint influence on prices is determined by the unknown function  $q(\cdot)$ .

Robinson shows that the model in (4) can be rewritten as;

$$\ln P_i - E[\ln P | \mathbf{x}_i, \mathbf{c}_i] = (\mathbf{z}_i - E[\mathbf{z} | \mathbf{x}_i, \mathbf{c}_i]) \boldsymbol{\beta} + \varepsilon_i \quad (5)$$

suggesting that  $\boldsymbol{\beta}$  can be estimated in a two-step procedure;

- First, the unknown conditional means  $E[\ln P | \mathbf{x}_i, \mathbf{c}_i]$  and  $E[\mathbf{z} | \mathbf{x}_i, \mathbf{c}_i]$  are estimated using a non-parametric estimation technique.
- Second, the estimates are substituted in place of the unknown functions in (5) and ordinary regression techniques employed to estimate  $\boldsymbol{\beta}$ .

Indeed, Robinson shows that the resulting parameter estimates are asymptotically equivalent to those that would be derived if the true functional form of  $q(\cdot)$  were known and could be used in the estimation.

We estimate the quantities  $E[\ln P | \mathbf{x}_i, \mathbf{c}_i]$  and  $E[\mathbf{z} | \mathbf{x}_i, \mathbf{c}_i]$  using non-parametric kernel regression with a multivariate Gaussian kernel. We fix the bandwidth for the elements of  $\mathbf{c}_i$  to a prespecified value,  $b$ , such that each observation is spatially smoothed over the same circular area. We adopt a bandwidth matrix for  $\mathbf{x}_i$  that is proportional to the sample covariance matrix of  $\mathbf{x}$  and estimate this proportion,  $h$ , using cross-validation. Finally, we use a two-step adaptive kernel estimation procedure that, in the second step, employs an observation-specific proportion,  $h_i$ , calculated to allow for differences in the density of the data in the region of each observation.

The data can be expressed as differences from the spatially smoothed conditional expectations; that is  $\tilde{y}_i = \ln \tilde{P}_i = \ln P_i - E[\ln P | \mathbf{x}_i, \mathbf{c}_i]$  and  $\tilde{z}_i = \mathbf{z}_i - E[\mathbf{z} | \mathbf{x}_i, \mathbf{c}_i]$ . Stacking the  $\tilde{y}_i$  data to form the  $N \times 1$  matrix  $\tilde{Y}$  and the  $\tilde{z}_i$  vectors to form the  $N \times k$  matrix  $\tilde{Z}$  allows the parameters to be estimated according to;

$$\tilde{\boldsymbol{\beta}} = (\tilde{Z}' \tilde{Z})^{-1} \tilde{Z}' \tilde{Y} \quad (6)$$

Despite the use of a spatial smoothing estimator, the possibility remains that a variety of features of each property's location are not captured by the model. Since these features are held in common by properties in close proximity, it is possible that the regression residuals will exhibit spatial correlation. If this is so then (6) will return inefficient estimates of the model parameters and biased estimates of the parameters' standard errors.

In our application we find clear evidence of spatial autocorrelation in our regression residuals. For simplicity, we assume that this spatial error dependence takes the form;

$$\boldsymbol{\varepsilon} = \rho \mathbf{W} \boldsymbol{\varepsilon} + \mathbf{u} \quad (7)$$

where  $\varepsilon$  is the  $[N \times 1]$  vector of random error terms with mean zero,  $\mathbf{W}$  is an  $N \times N$  weighting matrix,  $\rho$  is the error dependence parameter to be estimated and  $\mathbf{u}$  is the usual  $N \times 1$  vector of random error terms with expected value zero and variance-covariance matrix  $\sigma^2 \mathbf{I}$ . Notice that  $\rho = 0$  implies  $\varepsilon = \mathbf{u}$  and there is no spatial dependence in the data. The association between one property and another is captured by the weighting matrix,  $\mathbf{W}$ . The diagonal elements of  $\mathbf{W}$  are zero, whilst the off-diagonal elements that represent the potential spatial dependence between observations, are non-zero only if properties are closer than some predetermined distance,  $d$ . In our application we adopt a binary weights matrix in which the  $w_{ij}^{\text{th}}$  element of  $\mathbf{W}$  is initially set to one if the  $i^{\text{th}}$  and  $j^{\text{th}}$  property are located within  $d$  metres of each other, otherwise that element is set to zero.

Following Bell and Bockstael (2000) we employ a generalised moments (GM) estimator developed by Kelejian and Prucha (1999) to estimate the model with spatial error dependence. This is a two step estimator. The first step recovers a consistent estimate of the spatial correlation parameter,  $\hat{\rho}$ , whilst the second step estimates the parameters of the HPF by reformulating (6) as the feasible generalised least squares (FGLS) estimator;

$$\hat{\beta} = \left( \tilde{\mathbf{Z}}' \left( (\mathbf{I} - \hat{\rho} \mathbf{W})' (\mathbf{I} - \hat{\rho} \mathbf{W}) \right)^{-1} \tilde{\mathbf{Z}} \right)^{-1} \tilde{\mathbf{Z}}' \left( (\mathbf{I} - \hat{\rho} \mathbf{W})' (\mathbf{I} - \hat{\rho} \mathbf{W}) \right)^{-1} \tilde{\mathbf{Y}} \quad (8)$$

Since we include the noise variables in  $\mathbf{z}$ , the implicit price of noise for each observation can be calculated as;

$$\begin{aligned} p_{z_k} &= \exp \left( E[\ln P | \mathbf{x}_i, \mathbf{c}_i] + (z_i - E[z | \mathbf{x}_i, \mathbf{c}_i]) \hat{\beta} \right) \hat{\beta}_{z_k} \\ &= \exp \left( \ln \hat{P}_i \right) \hat{\beta}_{z_k} \end{aligned} \quad (9)$$

where  $\hat{\beta}_{z_k}$  is the parameter estimated on the noise variable  $z_k$ .

## 5. ESTIMATION OF IMPLICIT PRICES FOR NOISE

We apply techniques of model-based clustering so as to identify groups (clusters) of properties in the multidimensional space defined by each property's floor area, garden area as well as the four factors indicating the relative wealth, ethnicity, adult age composition and family composition of the property's neighbourhood. Our analysis provides evidence for the existence of eight clusters in the data. Inspecting the property and neighbourhood characteristics of the observations in each cluster reveals that each delineates a distinct and readily interpretable market segment. A summary of these market segments is provided in Table 1.

**Table 1: Characteristics of Birmingham market segments**

Market Segment	Property Size	Income	Ethnicity	Children
<b>1. Low-Income, White</b>	Small	Low	White	Many
<b>2. Low-Income, Ethnic</b>	Small	Low	Ethnic	Many
<b>3. Middle-Income, Young, No Children</b>	Small	Medium	Mixed	Few
<b>4. Middle-Income, White</b>	Small	Medium	White	Some
<b>5. Middle-Income, Ethnic</b>	Medium	Medium	Ethnic	Some
<b>6. High-Income, Large Properties</b>	Large	High	Mixed	Few
<b>7. High-Income, Small Properties</b>	Small	High	White	Some
<b>8. High-Income, Medium Properties</b>	Medium	High	White	Many

We partition the data according to the clustering classification and seek to estimate separate HPFs for each partition. We include in the regression analysis all of the variables described in the appendix. In particular, our measures of noise exposure from road, rail and air traffic are taken as dB in excess of 55dB LEQ, a threshold commonly regarded as a measure of background noise in urban areas. Along with spatial coordinates, we include in the nonparametric part of the specification variables describing each property's floor area, garden area, construction age and sale date. The spatial smoothing bandwidth,  $b$ , was set to a distance approximating the radius of a ward.

We examined the residuals from regressions of the smoothed data (Equation 6) for outliers and found concentrations of observations in the extreme left hand tail of the error distributions. Many of these 'under-priced' properties have apparently sold outside the market and therefore off no information on the HPF for that market.<sup>4</sup> After individual investigation we trimmed 1.9% of the observations from the data and re-estimated (6). We examined the residuals from this regression for evidence of spatial correlation over an area roughly equating to that of an ED. In all eight cases Moran's I statistic (Cliff and Ord, 1972) is large enough to reject the null hypothesis of no spatial dependence between residuals with greater than 95% confidence. A robust test proposed by Kelejian and Robinson (1992) supports this conclusion. Accordingly we apply the GM estimator (8) to estimate the parameters of the HPFs.

Subsequently, we test to see whether the HPFs from the different partitions of the data can be regarded as statistically different functions using a series of pairwise Wald tests. In each case, the test statistics are significant at a greater than 95% level of confidence. The results imply that different price structures characterise the property markets of the different data

---

<sup>4</sup> For example, a whole row of properties were sold within a few months of each other at prices well below the apparent market rate. Examination of recent aerial photographs provided an explanation; the houses had since been demolished to make way for a road widening scheme

partitions and we take this as evidence to support our contention that we have identified separate market segments through the technique of cluster analysis (as per Allen et al. 1995) .

Table 2 provides some summary statistics and a selection of the parameter estimates from the HPFs for the different market segments.<sup>5</sup> The final three rows of the table list the number of observations in each market segment,  $N$ , the number of parameters in each market segment regression,  $K$ , and the unadjusted  $R^2$  statistic for each regression (calculated according to the procedure suggested by Anglin and Gencay, 1996). Based on a comparison of this statistic, it appears that the models perform reasonably well for all six middle and high income market segments (segments 3 to 8) but rather less well in the two low income market segments (segments 1 and 2).

Parameter estimates for the first two variables are taken from the nonparametric part of the regression models. Technically, the values reported are not coefficients but averages of the slope of the hedonic function estimated nonparametrically for each observation. Since both floor area and garden area enter the regression function as logarithms, the quantities listed in the table for these variables can be interpreted as elasticities. Not surprisingly, property prices are increasing in floor area and garden area in each market segment.

---

<sup>5</sup> Full results can be found in Appendix B.



**Table 2: Selected parameter estimates from the first stage Regressions**

Variable	Market Segment							
	1	2	3	4	5	6	7	8
In (Floor Area)	0.287 <sup>†</sup>	0.392 <sup>†</sup>	0.304 <sup>†</sup>	0.196 <sup>†</sup>	0.347 <sup>†</sup>	0.293 <sup>†</sup>	0.386 <sup>†</sup>	0.454 <sup>†</sup>
In (Garden Area)	0.059 <sup>†</sup>	0.077 <sup>†</sup>	0.070 <sup>†</sup>	0.062 <sup>†</sup>	0.162 <sup>†</sup>	0.172 <sup>†</sup>	0.089 <sup>†</sup>	0.123 <sup>†</sup>
Bedrooms 2	0.000	0.007	-0.028**	-0.028	-0.035	-0.002	-0.030***	-0.045**
WCs 2	0.034*	0.036*	0.034**	0.023	0.012	0.013	-0.011	-0.003
Garage	0.043***	-0.027	0.064***	0.068***	0.071***	0.031	0.026***	0.056***
Detached House	-0.046	-0.083	0.051	0.154***	0.087***	0.161***	0.124***	0.113***
Wealth	0.103***	-0.001	0.056***	0.129***	0.114***	0.161***	0.065***	0.084***
Ethnicity	-0.101***	0.048**	-0.043**	-0.111***	-0.045***	-0.147***	-0.085***	-0.100***
Primary School	0.113***	0.125**	0.289***	0.207***	0.107*	0.076	0.074**	0.091*
Landfill	5E-05***	-6E-05**	-1E-05	4E-05**	-4E-05**	1E-05	4E-05***	3E-05*
Road Noise (over 55dB)	.0018	.0035*	-.0053***	-.0028**	-.0055***	-.0021	-.0018**	-.0025**
Rail Noise (over 55dB)	-.0084*	-.0068	-.0063*	-.0135***	-0.0050	-.0049	.0001	-.0085**
Aircraft Noise (over 55dB)	-.0160***		-.0154	.0032	.0339	-.0230	-.0064	.0033
<b>N</b>	<b>1484</b>	<b>1016</b>	<b>1523</b>	<b>1362</b>	<b>1058</b>	<b>424</b>	<b>2341</b>	<b>1432</b>
<b>K</b>	<b>41</b>	<b>37</b>	<b>46</b>	<b>44</b>	<b>47</b>	<b>57</b>	<b>44</b>	<b>52</b>
<b>R<sup>2</sup></b>	<b>.552</b>	<b>.618</b>	<b>.751</b>	<b>.691</b>	<b>.778</b>	<b>.886</b>	<b>.679</b>	<b>.830</b>

† Average nonparametric derivative. No significance estimated.

\* Significant at 10% level of confidence

\*\* Significant at 5% level of confidence

\*\*\* Significant at 1% level of confidence

The next four rows of Table 2 report parameter estimates for dummy variables describing characteristics of properties. Roughly speaking the parameter estimates on these constants relate the percentage difference in a price of a property with this characteristic compared to the base case (see Halvorsen and Palmquist, 1980, for a more correct interpretation). For example, the variable “Bedrooms 2” compares the price of properties with two bedrooms to the base case; properties with three bedrooms. These estimates tend to behave as expected. The majority of coefficients are negatively signed, though only in market segments 3, 7 and 8 is this statistically significant. Popular perception might suggest that this variable should be more important in determining property prices. Of course, in the models presented here, the overall size of the property is controlled for by including ‘floor area’ as a regressor in the nonparametric part of the model. Accordingly a variable indicating less bedrooms should really be interpreted as ‘less but bigger bedrooms’ and the relative unimportance of this variable seems more acceptable.

The “WCs 2” variable compares property prices of ‘two-toileted’ properties with those in the base case with one toilet. Market segments 1, 2 and 3 each have significant positive coefficients though, once again, the general lack of significance of this variable may be ascribed to the high correlation between the number of toilets and overall size of a property, a feature for which we have already controlled.

In contrast, the garage constant provides consistently positive and in the main highly significant parameters. Notably the only market segment in which the parameter estimate is negative, though not significant, is the inner-city market segment 2. This finding may reflect the fact that, in contrast to suburban locations, few properties in the inner city are provided with a garage (mostly having been built in the Victorian era) and that in this market segment less people have access to or requirement of their own vehicle.

Parameter estimates for the ‘detached house’ constant show a similar pattern. The negative and insignificant parameters for market segments 1 and 2 reflect the fact that less than 1% of properties in these market segments are detached. Amongst the more affluent market segments being detached adds considerably to the selling price of a property.

The variables describing the socioeconomic characteristics of neighbourhoods tend to be important in explaining variation in property prices. In Table 2 we present parameter estimates for the Wealth and Ethnicity factors. The impact of neighbourhood wealth is unequivocal. In all but one case the increasing wealth of the inhabitants of an area manifests itself in higher property prices. A more interesting pattern is revealed from the ethnicity variable. In seven of the eight market segments, the increasing presence of residents from ethnic minorities tends to decrease property prices. In contrast, the parameter on the ethnicity variable in market segment 2 (Black and Asian inner city), is positive and significant. Evidently in market segment 2 the increasing presence of residents from ethnic minorities impacts property prices favourably. Combining these two observations suggests that in the City of Birmingham the market rewards ethnic homogeneity.

The variable for primary schools combines distance and school quality into a single index. High scores indicate increasing quality and/or ease of access. The results here corroborate anecdotal evidence and that of recent studies (Gibbons and Machin, 2003) suggesting that increasing primary school quality and proximity inflates property prices. Positive parameter estimates are returned for all of the market segments though one of these, market segment 6, is not significant. Since market segment 6 is the high-income large property market segment this finding may simply reflect the relative lack of households with young children and/or the availability of alternative educational opportunities that reduce the perceived importance of state funded educational institutions.

Results for the landfill site variable are more equivocal. In five of the market segments the

coefficients are, as expected, positive and four of these are statistically significant. However, two of the remaining market segments return negative and significant parameter coefficients. As we cannot support the implication that households in these market segments actually prefer to be near to a landfill site we have to conclude that other locational factors, not adequately controlled for by the models, are driving this result.

Finally, let us consider the focus variables for this research project, those on transport-related noise. Notice first that there are many more observations of properties suffering from road noise than there are rail noise. Indeed, the relative paucity of properties exposed to rail traffic noise suggests that it will be relatively more difficult to find a statistically significant relationship between property prices and rail traffic noise pollution. The same could be said of aircraft noise which is concentrated, to a large extent, in market segments 1, 4 and 7 whose members include properties located in the East of the city near Birmingham International Airport. Indeed, in market segment 2 there are no properties exposed to aircraft noise.

In accordance with prior expectations the majority of parameter estimates on road and rail noise pollution are negative. In contrast the estimates on the air noise variable are far more erratic. In particular, of the seven market segments in which implicit prices for air traffic noise have been estimated, three have coefficients that imply a positive price for increasing noise. Of course, the coefficients estimated for some of the market segments are based on a very small sample of observations suffering from aircraft noise pollution. Amongst the three market segments which contain the majority of the properties exposed to aircraft noise (market segments 1, 4 and 7) our models the parameters have the expected sign in market segments 1 and 7 but not in market segment 4.

One reason for the relatively poor performance of the aircraft noise variable may be our choice of econometric model. Unfortunately, air traffic noise is considerably less localised than that arising from either road or rail traffic. Indeed properties over a large area will experience very similar levels of air traffic noise. A short-coming of the modelling approach adopted in this research is that much of the influence of these wide-area spatial effects is diluted by spatial smoothing of the data. Indeed, econometric specifications not reported show that when spatial effects are not modelled, parameter estimates on air traffic noise fall into line with prior expectations. Nevertheless, the authors believe that spatially smoothing the data is a sacrifice worth making. In particular, considerable improvement is realised in the ability of the models to identify local-area spatial effects such as those resulting from road and rail traffic noise.

Turning to the road noise variables, first observe that coefficients estimated for market segments 1 and 2 are positive. Whilst we are happy to accept that a price premium for peaceful environments may not exist in these two low-income market segments, it is implausible to conclude that households are actually willing to pay more for noisier properties. Clearly, some important aspect of the local environments in these two market segments is not captured by our specification of the HPF. This tallies with our observation that the econometric models of the HPF in these two market segments show the poorest fit to the data.

The remaining market segments return negative road noise coefficients ranging from a value of  $-.0021$  to a value of  $-.0053$ . In other words, our models indicate that a one decibel increase in road traffic noise can reduce the selling price of a property between 0.21% and 0.53%. Encouragingly, five out of the remaining six market segments have coefficients that are significant at the 95% level of confidence.

With regard to the rail noise coefficients, observe that on the whole the estimates are relatively larger in magnitude to those found to characterise road noise. On average the models indicate an NSDI of 0.67%. In all but one case, market segment 7, the rail noise

coefficients are negative.

The results presented in Table 2 are generally pleasing. In the main, the coefficient estimates are of the correct sign and mostly have plausible magnitudes. The models reveal that a 1dB increase in road noise reduces property prices by between 0.21% and 0.53%, depending on market segment. The rail noise estimates are relatively larger in magnitude, indicating that on average a 1dB increase in rail noise will reduce property prices by 0.67%. We find little evidence of a relationship between air traffic noise and property prices. This finding is almost certainly a result of our econometric approach which tends to subsume wide-area spatial effects into other parameters of the model.

For each household, we estimate the implicit price of each type of noise using Equation (9).

## 6. THE ECONOMETRIC MODEL: SECOND STAGE

In the second stage we seek to estimate a demand system for 'peace and quiet'. Notice that in this stage we invert the measure of noise so as to frame the analysis in terms of the good 'peace and quiet' rather than the bad 'noise pollution'. Moreover, the demand for 'peace and quiet' is assumed to be censored from above at the background urban noise level of 55dB. In addition, given the insignificant findings with regards to the implicit price of aircraft noise, we restrict our investigation to the estimation of demand equations for the avoidance of road and rail traffic noise.

For the purposes of a demand analysis we would prefer our implicit prices to be expressed as annual rents. According to theory, if one assumes that housing rental and purchase markets operate perfectly, then a property's purchase price will be equal to the discounted sum of the stream of future annual rents realisable from that property over its expected lifetime. It is mathematically convenient to assume that the expected lifetime of a property is infinite such that given some discount rate,  $\tau$ , implicit prices for property characteristics can be converted to implicit rents according to;

$$r_{z_i} = \tau \cdot p_{z_i} \quad (10)$$

The key issue, to which we shall return in the next section, is determining  $\tau$ ; the rate of discount used by households when comparing property prices with property rents.

Having invoked a two-stage budgeting hypothesis, the relevant argument in the demand relationship is not household income, but the household's total expenditure in purchasing a property. Of course, nonlinearity in the HPF implies non-constant implicit prices such that this total expenditure does not trace out the usual linear budget constraint in property attribute space.

Non-linearity of the budget constraint is a considerable problem for the estimation of Marshallian demand functions. First, since prices are not parametric to the consumer's problem, households choice of residential location implies a simultaneous choice of both the price and quantity of property attributes. Accordingly, our demand function specification must account for the endogeneity of prices. Even more troublesome is the fact that when implicit prices are not constant a household's choice of attribute levels will depend not only on marginal prices but on the parameters of the entire hedonic price function. Under these circumstances, even if one makes the standard assumptions concerning preferences, Marshallian demand functions do not necessarily manifest themselves as downward sloping curves in price quantity space (Palmquist, 1988).

One well-established procedure for dealing with this problem is that proposed by Murray (1983) and Palmquist (1988) in which the set of implicit prices at the chosen bundle are used to linearise the budget constraint. In effect, the expenditure variable is adjusted so as to indicate the 'virtual' expenditure that would lead a household to the same choice of property if the prices of characteristics were constant and equal to the implicit prices at their chosen bundle. We include linearised rather than actual expenditure in the specification of the demand system and, as a result, the functions we estimate are 'pseudo'-Marshallian demand functions. This estimation strategy has been applied previously in the context of hedonic models by Bajic (1993) and Boyle et al. (1999) amongst others. Moreover, we convert our linearised expenditure,  $\tilde{y}$ , argument into an annual equivalent,  $\tilde{m}$ , according to;

$$\tilde{m} = \tau \cdot \tilde{y} \quad (11)$$

In theory, the implicit prices of all property attributes should enter the demand relationship for 'peace and quiet'. In practice this would require the inclusion of an impractically large

number of covariates. Consequently, we ignore cross-price effects for all attributes other than those relating to the noise environment.

Our econometric analysis seeks to estimate two separate demand equations; one for peace and quiet from road traffic noise a second for peace and quiet from rail traffic noise. In general form, we express these two relationships as;

$$\begin{aligned} q_1^* &= \tilde{q}_1(r_1, r_2, \tilde{m}, \mathbf{s}) \\ q_2^* &= \tilde{q}_2(r_1, r_2, \tilde{m}, \mathbf{s}) \end{aligned} \quad (12)$$

We denote equations and variables pertaining to road noise by the subscript 1 and those pertaining to rail noise by the subscript 2. Furthermore,  $q_k^*$  ( $k = 1, 2$ ) is the quantity of peace and quiet;  $\tilde{q}_k(\cdot)$  ( $k = 1, 2$ ) is the pseudo demand relationship,  $r_k$  ( $k = 1, 2$ ) are the annualised implicit prices for peace and quiet,  $\tilde{m}$  is linearised annual expenditure on property services and  $\mathbf{s}$  is a vector of socioeconomic demand shifters.

To maintain simplicity we estimate (12) as a linear demand system. That is;

$$q_{ki}^* = \mathbf{Y}_i \boldsymbol{\beta}_k + \mathbf{X}_{1i} \boldsymbol{\gamma}_k + u_{ki} = \mathbf{Z}_i \boldsymbol{\delta}_k + u_{ki} \quad (k = 1, 2; i = 1, \dots, N) \quad (13)$$

For reasons that shall become clear shortly, we have gathered together the implicit prices,  $r_1$  and  $r_2$ , and the natural logarithm of expenditure,  $\ln(\tilde{m})$ , for household  $i$  into the  $1 \times 3$  vector  $\mathbf{Y}_i$ . Moreover, we have redesignated the vector of socioeconomic demand shifters ( $\mathbf{s}$ ) for household  $i$  as the vector  $\mathbf{X}_{1i}$ . In addition, the explanatory variables of the demand function for household  $i$  can be gathered together into the vector  $\mathbf{Z}_i = [\mathbf{Y}_i, \mathbf{X}_{1i}]$  whilst the parameters of the demand functions can be gathered together in the vector  $\boldsymbol{\delta}_k = [\boldsymbol{\beta}_k, \boldsymbol{\gamma}_k]$  ( $k = 1, 2$ ). Finally,  $u_{ki}$  is an unobservable household specific random term capturing elements of the demand relationships not accounted for in the model specification.

Whilst the linear specification in (13) contravenes restrictions implied by economic theory we regard the linear demand function as a reasonable approximation.<sup>6</sup> Our choice of functional form is also driven by pragmatic considerations. In particular, our data suffer from both censoring and endogeneity and we note that techniques for dealing with these difficulties are well developed for linear regression.

One such difficulty is that our data is heavily censored from above with many households choosing properties at the corner solution imposed by the existence of an urban background level of noise.<sup>7</sup> Our dependent variable, therefore, is of the form;

$$q_{ki} = \begin{cases} q_{ki}^* & \text{if } q_{ki}^* < 0 \\ 0 & \text{otherwise} \end{cases} \quad (k = 1, 2; i = 1, 2, \dots, N) \quad (14)$$

Unfortunately, in the presence of censoring, standard approaches to the estimation of a

---

<sup>6</sup> Several econometric specifications have been developed that seek to explicitly impose the restrictions implied by economic theory. One example of such a specification is the almost ideal demand system (AIDS) of Deaton and Muellbauer (1980), a specification which was employed by Sheppard and Cheshire (1998) and Palmquist and Isangkura (1999) in their demand analyses based on hedonic data.

<sup>7</sup> Specifically, that in only 2,723 households choose locations enduring road traffic noise above the urban background level (taken as 55dB), whilst only 379 choose locations enduring rail traffic noise above the urban background level.

linear function such as (13) will return biased and inconsistent estimates of the model parameters. To progress, we impose the assumption that the error term  $u_{ki}$  is distributed normally with mean zero and constant variance  $\sigma_k^2$ . Furthermore, we assume that the error terms are uncorrelated across households that is;

$$\text{cov}(u_{ki}, u_{kj}) = 0 \quad (i, j = 1, 2, \dots, N; i \neq j; k = 1, 2) \quad (15)$$

Under these assumptions, one might estimate the parameters of the demand relationships using Tobit regression. However, in our case, a further difficulty is introduced by the fact that implicit prices are endogenous. Moreover, the cross-price terms are also endogenous as is linearised expenditure ( $\tilde{m}$ ) since this is calculated using these same implicit prices. Accordingly, the own-price, cross-price and expenditure variables making up the vector  $\mathbf{Y}_i$  are endogenous to the choice problem whilst the socioeconomic demand shifters in the vector  $\mathbf{X}_{1i}$  are considered exogenous.

The most general technique for handling endogenous variables is the method of *instrumental variables* (IV). The fundamental ingredient of any IV procedure is a set of instrumental variables. Crucially, each instrumental variable must be correlated with the endogenous variables in the model, but must be independent of those elements of the joint decision process (i.e. the simultaneous choice of quantity and price) that are not captured by the model. Here we assume that these uncaptured elements reflect specific characteristics of the household and their attitude to noise. As such we follow the suggestion of Cheshire and Sheppard (1998) and select as instrumental variables the values of  $r_{jk}$  and  $\tilde{m}_j$  chosen by other nearby households. We define “nearby” as proximity in a multidimensional space defined by geographic location, socioeconomic characteristics (i.e. neighbourhood factors) and property characteristics (i.e. floor and garden area). Our matrix of instrumental variables consists of  $r_{jk}$  and  $\ln(\tilde{m}_j)$  for the three nearest neighbours to each observation and the squares of these values. We gather the instrumental variables for each household into the vector of exogenous regressors  $\mathbf{X}_{2i}$ .

Now the reduced-form equations for the endogenous explanatory variables can be written as the set of three linear equations;

$$\mathbf{Y}_i = \mathbf{X}_i \mathbf{\Pi} + \mathbf{v}_i = \mathbf{X}_{1i} \mathbf{\Pi}_1 + \mathbf{X}_{2i} \mathbf{\Pi}_2 + \mathbf{v}_i \quad (i = 1, \dots, N) \quad (16)$$

Here,  $\mathbf{Y}_i$  is each household’s observed choice of road price, rail price and expenditure,  $\mathbf{X}_i$  is the matrix containing all the exogenous variables (demand shifters,  $\mathbf{X}_{1i}$ , and instruments,  $\mathbf{X}_{2i}$ ) and  $\mathbf{\Pi}_1$  and  $\mathbf{\Pi}_2$  are reduced form parameters which can be gathered together into the matrix  $\mathbf{\Pi}$ . Finally,  $\mathbf{v}_i$  is a vector of disturbances.

To construct reduced form demand equations we may substitute (16) into (13) giving;

$$\begin{aligned} q_{ki}^* &= (\mathbf{X}_i \mathbf{\Pi} + \mathbf{v}_i) \boldsymbol{\beta}_k + \mathbf{X}_{1i} \boldsymbol{\gamma}_k + u_{ki} \\ &= \mathbf{X}_{1i} (\mathbf{\Pi}_1 \boldsymbol{\beta}_k + \boldsymbol{\gamma}_k) + \mathbf{X}_{2i} (\mathbf{\Pi}_2 \boldsymbol{\beta}_k) + \mathbf{v}_i \boldsymbol{\beta}_k + u_{ki} \\ &= \mathbf{X}_{1i} \boldsymbol{\alpha}_{1k} + \mathbf{X}_{2i} \boldsymbol{\alpha}_{2k} + \mathbf{v}_i \boldsymbol{\beta}_k + u_{ki} \\ &= \mathbf{X}_i \boldsymbol{\alpha}_k + \mathbf{v}_i \boldsymbol{\beta}_k + u_{ki} \end{aligned} \quad (k = 1, 2; i = 1, \dots, N) \quad (17)$$

where  $\boldsymbol{\alpha}_{1k} = \mathbf{\Pi}_1 \boldsymbol{\beta}_k + \boldsymbol{\gamma}_k$ ,  $\boldsymbol{\alpha}_{2k} = \mathbf{\Pi}_2 \boldsymbol{\beta}_k$  and  $\boldsymbol{\alpha}_k$  gathers together the vectors  $\boldsymbol{\alpha}_{1k}$  and  $\boldsymbol{\alpha}_{2k}$ .

Amemiya (1979) suggested that one could estimate  $\hat{\mathbf{\Pi}}$ , the reduced form parameters in (16), using OLS. Subsequently, it is possible to estimate the reduced form parameters of the

demand equations,  $\hat{\alpha}_k$ , by inserting the estimated residuals,  $\hat{v}_i$  into (17) and applying Tobit regression techniques.

Of course, our interest is not in the parameters of the reduced form equations,  $\alpha_k$ , but in the parameters of the structural demand equations  $\delta_k$  ( $k=1,2$ ). However, Amemiya (1979) observed that  $\alpha_k$  and  $\delta_k$  are related through the equation;

$$\alpha_k = D(\Pi)\delta_k \quad (k = 1, 2) \quad (18)$$

where  $D(\Pi) \equiv [\Pi, I_1]$  is the selection matrix made up of ones and zeros such that  $X_{1i} = X_i I_1$ .

If we knew the true values of the reduced form parameters  $\alpha_k$  and  $\Pi$  then it would be a simple task to retrieve  $\delta_k$  using (18). Unfortunately, we only have estimates of the reduced form parameters,  $\hat{\Pi}$  and  $\hat{\alpha}_k$ . Since these estimates will not be exactly equal to the true values, Equation (18) can be reformulated as per Amemiya (1979) to give;

$$\begin{aligned} \hat{\alpha}_k &= D(\hat{\Pi})\delta_k + (\hat{\alpha}_k - \alpha_k) + \beta(\hat{\Pi} - \Pi) \\ &= D(\hat{\Pi})\delta_k + \eta_k \end{aligned} \quad (k = 1, 2) \quad (19)$$

where  $\eta_k = (\hat{\alpha}_k - \alpha_k) + \beta(\hat{\Pi} - \Pi)$ . Amemiya (1979) shows how  $\delta_k$  can be estimated by applying generalised least squares to (19) in which the weighting matrix is constructed from a consistent estimator of the asymptotic covariance matrix of  $\sqrt{N}(\hat{\alpha} - D(\hat{\Pi}))$ . To construct such a weighting matrix requires assumptions to be made concerning the disturbance terms of the reduced form equations for the endogenous explanatory variables (16) and those in the demand equations of (13). Amemiya assumes that conditional on the exogenous variables  $X_i$ , these disturbance terms are multivariate normal;

$$(u_{ki}, v_{ki}) \sim N(\mathbf{0}, \Sigma_k) \quad (i = 1, 2, \dots, N; k = 1, 2) \quad (20)$$

The actual form of the weighting matrix is beyond the scope of this discussion. The formulas can be found in Newey (1987). This procedure for estimating  $\delta_k$  has become known as Amemiya Generalised Least Squares (AGLS). Amemiya (1979) and Newey (1987) show that these estimates of the structural parameters are more efficient than other possible estimators.



## 7. ESTIMATION OF DEMAND FUNCTIONS

In estimating the demand relationships, there are a number of data issues with which we have had to contend. First, notice that the first stage hedonic regression returns results that allow for negative prices for peace and quiet (most notably in market segments one and two). Clearly, negative prices defy both theory and common sense; are we to believe that households are prepared to pay more for properties in noisier locations? As a result we treat these negative prices as the result of sampling error where the 'true' implicit prices for peace and quiet are in the region of zero. Accordingly we set these prices to a value of zero in the regression analysis.

Second, we must decide upon a discount rate with which to annualise prices and expenditure. To this end we examined numerous Birmingham estate agent and letting agent listings in order to find pairs of similar properties being offered for rent and sale respectively. Properties were matched on a variety of features, namely location (to first 4 digits of the postcode), number of bedrooms, property type and construction age. Comparison of the sale price and rental price of matched properties provides an estimate of the required discount rate. This data from 2003 suggest a discount rate that equates to a value halfway between the mortgage rate and the base rate. Applying this simple rule of thumb to the Birmingham data from 1997 provides a discount rate of 6.78%.<sup>8</sup>

Finally, our data lacks information on the specific characteristics of the households making the property purchases. Accordingly we use the factor scores describing the socioeconomic characteristics of all households in the property's local neighbourhood as a proxy for each household's actual characteristics.

Applying the AGLS estimator to our data we estimate the parameters of the pseudo demand functions for rail and road noise and report those in Table 3.

In accordance with expectations, the own-price terms in both pseudo demand equations are negative and highly significant. Interestingly, the two own-price coefficients are relatively close in value suggesting the slopes of the two pseudo demand curves are relatively similar.

We expect the cross-price terms to be negative, reflecting the complementarity relationship that exists between 'peace and quiet' from different sources. That is to say, to attain a quiet environment, the household must purchase 'peace and quiet' from different sources of noise pollution in roughly similar quantities. Increases in the price of 'peace and quiet' from one particular source will not only reduce demand for 'peace and quiet' from that source but will also tend to reduce demand for 'peace and quiet' from other sources. The regression results, provide weak evidence supporting this hypothesis. In the rail demand equation the parameter on the road noise price is negative and significant at the 95% level of confidence. However, contrary to expectations, the parameter for rail noise price in the road demand equation is positive though not statistically significant.

Demand for 'peace and quiet' is greater for households allocating greater virtual expenditure to the purchase of their property. As Blomquist (1989) observes, it does not necessarily follow that demand is increasing in actual expenditure though in numerous special cases this will also be true. In contrast to the similarity in own-price parameter estimates, the parameter estimate for expenditure is a great deal larger in the rail traffic noise equation. This indicates that the income elasticity of demand for 'peace and quiet' from rail traffic noise is greater than that for road traffic noise. It seems that wealthier households are increasingly more likely to choose environments free from rail traffic noise than from road traffic noise.

---

<sup>8</sup> In employing this rule of thumb, we must assume a certain degree of myopia on behalf of house buyers. That is, households are assumed to calculate  $d$  based purely on current interest rates and not on the expected future values of these rates as would be more theoretically correct. In reality, this assumption is probably not too far-fetched.

**Table 3: Parameters of the pseudo demand functions for road and rail noise**

	Road		Rail	
	Coefficient (Std. Err.)	t-score	Coefficient (Std. Err.)	t-score
Road Noise Price	-0.419*** (0.100)	-4.19	-0.399** (0.197)	-2.03
Rail Noise Price	0.104 (0.071)	1.47	-0.446*** (0.140)	-3.18
Log Pseudo Expenditure	7.649*** (2.614)	2.93	20.297*** (5.210)	3.90
Ethnicity Factor	1.438*** (0.247)	5.81	0.729 (0.501)	1.46
Age Factor	-1.506*** (0.430)	-3.50	2.978*** (0.840)	3.55
Family Factor	1.084*** (0.253)	4.28	0.535 (0.497)	1.08
Constant	-52.735*** (19.431)	-2.71	-124.726*** (38.432)	-3.25
Uncensored Obs:	2,723		379	
Censored Obs:	7,918		10,262	

\*\*\* Significant at 99% level of confidence

\*\* Significant at 95% level of confidence

\* Significant at 90% level of confidence

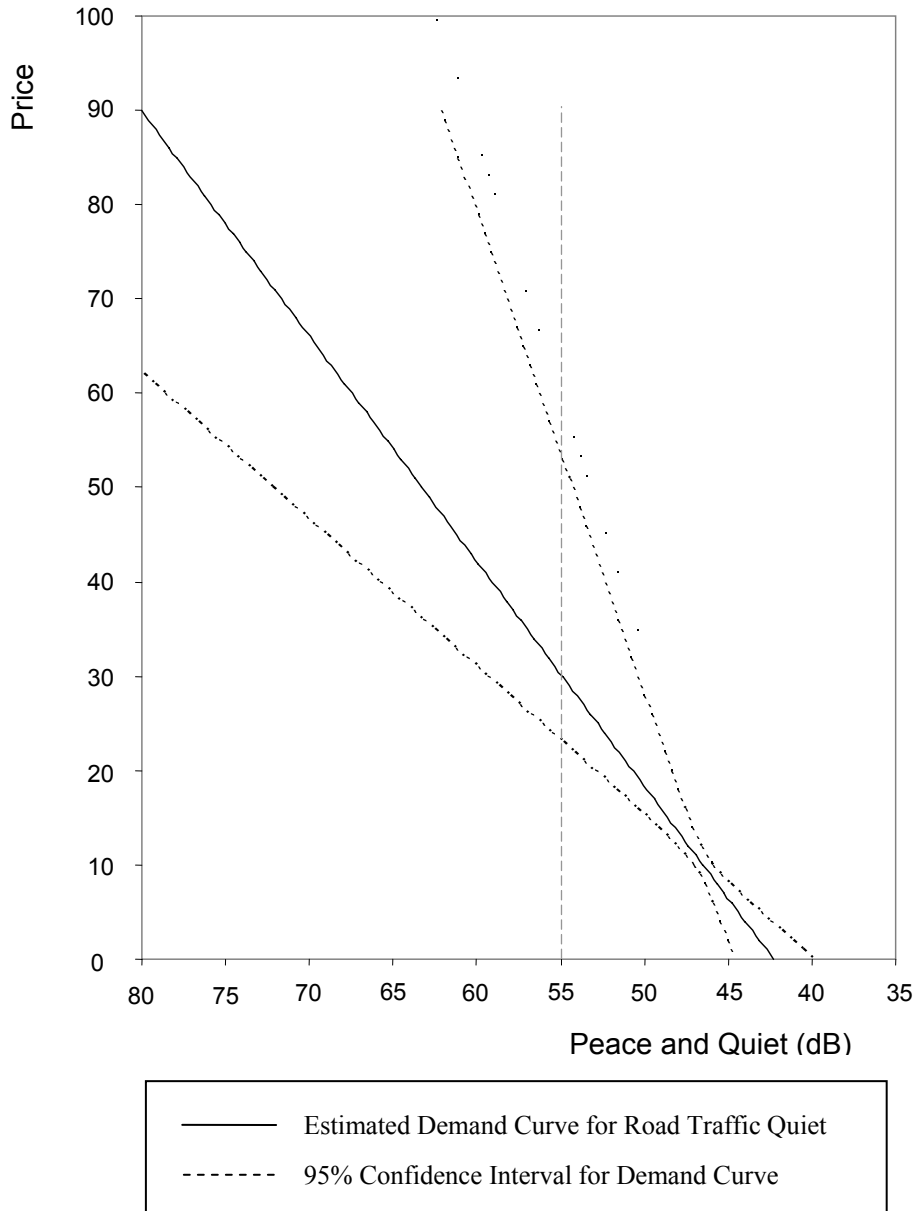
We are cautious in interpreting the parameters of the socioeconomic variables since these are only proxies for each household's actual characteristics. Reassuringly, the parameter estimates on the family composition factor are of the same sign and almost identical magnitude. These indicate that demand for noise avoidance is greater amongst households with young families. The ethnicity factor coefficient also takes the same sign in both equations indicating that households from the ethnic minorities demand relatively greater levels of 'peace and quiet', all else equal. However, for rail traffic noise at least, the parameter estimate is small in magnitude and statistically insignificant. The coefficient estimates for the adult age composition are the least easy to interpret. The estimates suggest that older households demand relatively more 'peace and quiet' from rail traffic noise but, conversely, relatively less 'peace and quiet' from road traffic noise. Since we have no particular theory to support these conflicting results we suggest that this is an appropriate juncture at which to execute "cautious scepticism".

We perform a number of specification tests on the models. Using both a familiar Hausman test and a test suggested by Smith and Blundell (1986), we find that for both demand functions the hypothesis that prices and expenditure are exogenous to the choice problem is clearly rejected. Moreover, we adopt the testing framework suggested by Weesie (2000) in order to test whether the two demand relationships can be considered as separate estimates of the same demand function for 'peace and quiet' from any source. We find that we can reject this hypothesis with a high level of confidence. and conclude that the demand function for 'peace and quiet' differs depending on whether the source of noise pollution is road traffic or rail traffic.

Figure 3 depicts the estimated demand function for road traffic noise at the means of the data. Projecting the demand curve down below 55dB reveals that marginal WTP falls to zero at a noise level of 42.3dB with a 95% confidence interval ranging from 39.9dB to 44.7dB. At first sight it is somewhat surprising that the confidence interval for the estimated demand curve narrows in the range below 55dB. As a matter of fact this is quite intuitive. The bulk of observations are for households facing relatively low prices and choosing properties below

55dB. Our confidence concerning the path of the demand curve is greatest where we have most data and becomes progressively less precise as we move away from the centre of the data.

**Figure 3: Road noise demand curve and its 95% confidence intervals**



## 8. WELFARE ESTIMATES

We estimate the benefits of an external change in the provision of 'peace and quiet' as areas under the pseudo demand curve and interpret these as approximating the exact welfare measure that would be given by areas under the compensated demand curve (Bartik, 1988). Table 4 reports welfare values for 1dB changes in noise pollution levels from road traffic and rail traffic (columns 3 and 5 respectively) calculated at the mean values of the covariate data. To calculate the welfare benefits from larger changes simply requires the relevant range of 1dB changes to be summed. This is illustrated in columns 4 and 6 of Table 4 where the benefits of particular 5dB changes are listed. The welfare estimates are values per annum reported in 1997 prices.

The mean values for road noise range from £31.49 per annum for a 1dB reduction from a 56dB baseline to £88.76 per annum for the same change from a 80dB baseline. The equivalent values for rail noise are higher, ranging from £83.61 to £137.41 per annum. One possible explanation for this observation is that the common unit used to measure noise from these two sources ( $dB_{LEQ}$ ) conceals important differences in the characteristics of road and rail noise.

**Table 4: Welfare estimates for changes in noise exposure at means of covariate data (95% confidence interval range)**

Noise Change		Welfare Change ( £ per Annum in 1997 prices)			
High	Low	Road		Rail	
56	55	31.49 (24.84 to 52.52)		83.61 (43.21 to 461.80)	
57	56	33.88 (26.51 to 57.26)		85.85 (44.43 to 473.39)	
58	57	36.26 (28.18 to 61.97)		88.09 (45.65 to 484.98)	
59	58	38.65 (29.81 to 66.64)		90.33 (46.87 to 496.58)	
60	59	41.04 (31.48 to 71.31)		92.57 (48.08 to 508.17)	
61	60	43.42 (33.14 to 75.97)	181.32 (140.90 to 309.87)	94.82 (49.30 to 519.77)	440.45 (228.24 to 2,424.92)
62	61	45.81 (34.81 to 80.64)		97.06 (50.52 to 531.36)	
63	62	48.19 (36.48 to 85.31)		99.3 (51.73 to 542.96)	
64	63	50.58 (38.15 to 89.97)		101.54 (52.95 to 554.55)	
65	64	52.97 (39.82 to 94.77)		103.78 (54.17 to 566.14)	
66	65	55.35 (41.49 to 99.57)	240.97 (182.41 to 426.53)	106.02 (55.39 to 580.05)	496.49 (258.67 to 2,714.78)
67	66	57.74 (43.16 to 104.37)		108.26 (56.60 to 594.85)	
68	67	60.12 (44.82 to 109.17)		110.51 (57.82 to 609.65)	
69	68	62.51 (46.50 to 113.97)		112.75 (59.04 to 624.45)	
70	69	64.9 (48.16 to 118.77)		114.99 (60.26 to 639.25)	
			300.62 (224.12 to 545.85)		552.53 (289.10 to 3,048.25)

71	70	67.28 (49.83 to 123.57)	360.27 (265.84 to 665.88)	117.23 (61.47 to 654.04)	608.57 (319.53 to 3,418.20)
72	71	69.67 (51.50 to 128.37)		119.47 (62.69 to 668.84)	
73	72	72.05 (53.17 to 133.18)		121.71 (63.91 to 683.64)	
74	73	74.44 (54.84 to 137.98)		123.96 (65.12 to 698.44)	
75	74	76.83 (56.50 to 142.78)		126.2 (66.34 to 713.24)	
76	75	79.21 (58.17 to 147.58)	419.92 (307.55 to 785.90)	128.44 (67.47 to 728.03)	664.61 (347.83 to 3,788.16)
77	76	81.6 (59.84 to 152.38)		130.68 (68.52 to 742.83)	
78	77	83.98 (61.51 to 157.18)		132.92 (69.57 to 757.63)	
79	78	86.37 (63.18 to 161.98)		135.16 (70.62 to 772.43)	
80	79	88.76 (64.85 to 166.78)		137.41 (71.67 to 787.23)	

Furthermore, Table 4 reports 95% confidence intervals for each of the reported welfare estimates. These are calculated using a non-parametric bootstrap procedure. It is immediately apparent that the 95% confidence intervals for the rail welfare estimates are considerably wider than those for the road welfare estimates. Of course, this is only to be expected given the fact that our dataset contains many fewer observations of exposure to rail traffic noise above the urban background level than is the case for road traffic noise.

## 9. CONCLUDING REMARKS

The research summarised in this paper represents one of the most comprehensive hedonic pricing studies carried out to date. The majority of previous studies have attempted simply to identify the implicit price function for peace and quiet using property market data. Since that implicit price function is simply a reflection of the specific conditions of supply and demand that exist in a property market, these studies provide no basis for calculating transferable welfare values.

In this research, we have attempted to go one step further and estimate demand functions for 'peace and quiet'. Since these functions identify a household's underlying preferences, that is, how they are prepared to trade-off between money and 'peace and quiet', they can be used as objects for transferring benefits across locations. As far as the authors are aware, the values presented in Table 4 are the first estimates of welfare values for peace and quiet derived from hedonic property market data in a theoretically correct manner.

In proffering these values out to the policy-making community, however, we draw attention to a number of caveats. First, our data relate only to households that purchase properties. Accordingly, the demand relationships reflect only the preferences of households participating in the property-purchasing market.

Second, there is a degree of circularity in the estimation strategy of introducing price variation through the identification of market segments. In particular, we would like our data to provide evidence of how households with similar characteristics choose when faced by different price functions in separate markets. Unfortunately, the market segments are defined in part by the socioeconomic characteristics of households such that they comprise relatively distinct socioeconomic groupings. Accordingly, the degree of overlap in household types across market segments may be rather less than might be hoped. Rather better would be for the first stage analysis to be applied in a second urban area characterised by a similar socioeconomic mix of inhabitants as that in Birmingham. The combination of the output from that second study with the Birmingham study would bring together data from similar households choosing in distinct markets and would provide a far stronger basis for identification of the demand relationship.

Finally a critical question for welfare analysis is the extent to which pseudo demand curves approximate compensated demand curves. This crucially depends on the size of income effects. If there are no income effects then the relationship between prices and quantities is not influenced by adjustments to expenditure. In this case, the pseudo demand curve will exactly replicate the compensated demand curve. As income effects become more pronounced, the pseudo demand curve will begin to differ from the compensated demand curve in ways that mirror the divergence between the ordinary demand curve and the compensated demand curve in a world of constant marginal prices. Theoretical work by Edlefsen (1981) and Blomquist (1989) has sought to describe the relationships that exist between the pseudo and ordinary demand functions when marginal prices are not constant. Unfortunately, there appears to be no theoretical research comparing the pseudo and compensated demand curves in a world of non-constant marginal prices to parallel that carried out by Willig (1976), Randall and Stoll (1980) and Hanemann (1991) in comparing the ordinary and compensated demand functions in a world of constant marginal prices. Accordingly, the authors are not in a position to indicate how closely the welfare estimates presented in Table 4 approximate the theoretically consistent values that would be derived from a compensated demand curve.

That said, we have taken care in our application to address all the major theoretical and empirical difficulties that afflict estimation of demand relationships from property market data. Moreover, the results of our analysis are statistically significant and conform with theoretical

expectations. Finally, our welfare estimates are of a magnitude that might be regarded as “reasonable”. We believe that these features of our study should be taken as an indication of the validity and robustness of the procedures used in this analysis.

**APPENDIX A: Data Descriptions**

<b>Variable</b>	<b>Mean</b>	<b>Std. Dev.</b>	<b>Min</b>	<b>Max</b>
Sale Price (£)	58,986	36,099	11,000	645,003
Structural Characteristics				
Floor Area (m <sup>2</sup> )	102.6	32.7	42	645
Garden Area (m <sup>2</sup> )	226.1	208	0	5,164
Garage (proportion)	0.436	0.496	0	1
Central Heating (proportion)	0.728	0.268	0	1
Age (decades)	6.1	2.76	0	11
WCs (proportion)				
One	0.794	0.404	0	1
Two	0.196	0.397	0	1
Three	0.009	0.094	0	1
> Three	0.001	0.029	0	1
Bedrooms (proportion)				
One	0.005	0.069	0	1
Two	0.172	0.377	0	1
Three	0.716	0.451	0	1
Four	0.083	0.276	0	1
Five	0.016	0.127	0	1
> Five	0.007	0.084	0	1
Storeys (proportion)				
One	0.021	0.145	0	1
Two	0.954	0.209	0	1
Three	0.021	0.143	0	1
> Three	0.003	0.058	0	1
Construction Type (proportion)				
Detached Bungalow	0.013	0.111	0	1
Semi-Detached Bungalow	0.008	0.090	0	1
End Terrace Bungalow	0.000	0.022	0	1
Terrace Bungalow	0.000	0.017	0	1
Detached House	0.116	0.320	0	1
Semi-Detached House	0.396	0.489	0	1
End Terrace House	0.115	0.319	0	1
Terrace House	0.352	0.478	0	1
Beacon Group (proportion)				
1. Unrenovated cottage pre 1919	0.000	0.019	0	1
2. Renovated cottage pre 1919	0.001	0.027	0	1
3. Small "industrial" pre 1919	0.040	0.195	0	1
4. Medium "industrial" pre 1919	0.226	0.418	0	1
5. Large terrace pre 1919	0.006	0.078	0	1
8. Small "villa" pre 1919	0.020	0.138	0	1
9. Large "villas" pre 1919	0.009	0.093	0	1
10. Large detached pre 1919	0.003	0.058	0	1
19. Houses 1908 to 1930	0.011	0.103	0	1
20. Subsidy houses 1920s & 30s	0.140	0.347	0	1
21. Standard houses 1919-	0.257	0.437	0	1



<b>Variable</b>	<b>Mean</b>	<b>Std. Dev.</b>	<b>Min</b>	<b>Max</b>
45				
24. Large houses 1919-45	0.016	0.124	0	1
25. Individual houses 1919-45	0.000	0.022	0	1
30. Standard houses 1945-53	0.045	0.207	0	1
31. Standard houses post 1953	0.190	0.392	0	1
32. Large houses post 1953	0.032	0.177	0	1
35. Individual houses post 1945	0.001	0.038	0	1
36. "Town Houses" post 1950	0.004	0.062	0	1
<b>Sale Date (proportion)</b>				
1 <sup>st</sup> Quarter (Jan. to Mar.)	0.214	0.410	0	1
2 <sup>nd</sup> Quarter (Apr. to June)	0.247	0.431	0	1
3 <sup>rd</sup> Quarter (July to Sept.)	0.287	0.452	0	1
4 <sup>th</sup> Quarter (Oct. to Dec.)	0.252	0.434	0	1
<b>Neighbourhood Characteristics</b>				
Poverty Factor	-0.375	0.855	-1.934	2.363
Sills Factor	0.180	1.000	-1.398	4.198
Age Factor	0.055	0.807	-3.216	3.143
Family Factor	-0.029	0.842	-3.198	3.791
Asian Factor	-0.045	0.942	-1.131	5.152
Black Factor	-0.240	0.750	-2.016	8.214
<b>Locational Characteristics</b>				
Proximity to City Centre (mins)	1,313	478	208	3,187
Proximity and Quantity of Shops	2.276	1.273	0.07	9.56
Proximity and Quality of Primary Schools	0.602	0.177	0.15	0.97
Walking time to Rail Station (secs)	1,846	1,013	21.05	5,525
Walking time to a Park (secs)	900	558	3.17	3,425
Driving time to Airport (secs)	2,388.215	655.134	602.19	4,386
Proximity to A-Type Industrial Processes (m)	2,463.592	1,820.591	21.94	10,204
Proximity to B-Type Industrial Processes (m)	814.103	527.842	10	3,333
Proximity to Land Fill sites (m)	946.611	608.089	10	3,472
<b>Environmental Characteristics</b>				
Views of Water (weighted m <sup>2</sup> )	0.480	7.543	0	348
Views of Parkland (weighted m <sup>2</sup> )	6.290	36.831	0	664
Road Traffic Noise (dB)	49.8	9.4	31.6	75.8
Rail Traffic Noise (dB)	36.8	12.6	0	74.7
Aircraft Noise (dB)	4.8	16.0	0	69

**APPENDIX B: Parameter Estimates for Hedonic Price Functions**

Variable	Parameter Estimates by Submarket (p-values in brackets)							
	1	2	3	4	5	6	7	8
<i>Nonparametric Variables:</i>								
In(Floor Area)	0.287	0.3919	0.3045	0.1957	0.347	0.2932	0.3864	0.4537
In(Garden Area)	0.0594	0.077	0.0705	0.062	0.1623	0.1719	0.0887	0.1231
Age	-0.009	-0.0059	-0.0082	-0.0054	-0.0208	-0.0046	-0.0076	-0.0218
Sale Date	0.0001	0.0002	0.0001	0.0001	0	0	0.0001	0.0002
<i>Structural Variables:</i>								
Bedrooms 1	0.0473 (.777)	-0.0292 (.906)	0.0453 (.784)	-0.0281 (.692)	-0.0557 (.792)		-0.065 (.726)	-0.4505 (.000)
Bedrooms 2	0.0002 (.993)	0.0073 (.695)	-0.0283 (.025)	-0.0285 (.104)	-0.0356 (.225)	-0.0018 (.982)	-0.0297 (.006)	-0.0448 (.055)
Bedrooms 3	base case	base case	base case	base case	base case	base case	base case	base case
Bedrooms 4	-0.1018 (.196)	0.0405 (.408)	0.0306 (.399)	0.0861 (.001)	0.0628 (.032)	0.0304 (.447)	0.0215 (.534)	0.0509 (.000)
Bedrooms 5		0.0422 (.728)	0.1771 (.056)		0.0958 (.068)	0.013 (.783)	0.2714 (.010)	0.0682 (.014)
Bedrooms 6		-0.2501 (.324)	0.2627 (.236)		0.001 (.991)	0.0076 (.903)		0.172 (.086)
Bedrooms 7					-0.042 (.860)	-0.17 (.069)		0.2138 (.146)
Bedrooms 8						0.0568 (.582)		
Bedrooms 9						-0.262 (.307)		
Bedrooms 10						-0.4596 (.131)		
Bedrooms 11						-0.265 (.355)		
Bedrooms 12						0.4654 (.228)		
WCs 1	base case	base case	base case	base case	base case	base case	base case	base case
WCs 2	0.0336 (.084)	0.0356 (.085)	0.0339 (.017)	0.0234 (.113)	0.0117 (.518)	0.013 (.661)	-0.011 (.262)	-0.0026 (.810)
WCs 3	0.0367 (.828)		0.1113 (.311)	0.0559 (.369)	-0.0365 (.830)	0.1252 (.052)	0.2219 (.033)	0.0317 (.289)
WCs 4				-0.0375 (.841)		-0.1734 (.238)		
WCs 5						0.0119 (.962)		
Floors 2	base case	base case	base case	base case	base case	base case	base case	base case
Floors 3	-0.0317 (.371)	-0.0326 (.781)	-0.0388 (.336)	-0.0747 (.035)	-0.0397 (.431)	-0.1833 (.003)	-0.0377 (.171)	-0.2109 (.000)
Floors 4			-0.1598 (.192)		-0.1974 (.034)	-0.6347 (.000)		-0.4681 (.000)
Floors 5			0.0281 (.923)			-0.9616 (.000)		
Floors 6						-0.0817 (.748)		

Variable	Parameter Estimates by Submarket (p-values in brackets)							
	1	2	3	4	5	6	7	8
Floors 7						0.6845 (.755)		
Garage	0.0434 (.001)	-0.0272 (.479)	0.0636 (.001)	0.068 (.000)	0.071 (.000)	0.0312 (.434)	0.026 (.000)	0.0562 (.000)
Central Heating	0.0837 (.000)	0.0161 (.555)	0.1005 (.002)	0.0797 (.004)	0.0155 (.689)	0.063 (.586)	0.0728 (.002)	0.2123 (.000)
Detached Bungalow	0.4157 (.071)			0.1045 (.020)	0.1052 (.279)	0.112 (.240)	0.089 (.066)	0.1498 (.000)
Semi-Detached Bungalow	-0.1725 (.291)			0.0558 (.125)	0.0071 (.966)		0.1292 (.043)	0.033 (.457)
End Terrace Bungalow	-0.0366 (.840)		-0.3607 (.196)	-0.2221 (.273)			0.219 (.770)	
Terrace Bungalow							0.063 (.678)	0.1464 (.518)
Detached House	-0.0462 (.367)	-0.0832 (.623)	0.051 (.304)	0.1545 (.000)	0.0866 (.003)	0.1606 (.000)	0.124 (.000)	0.1126 (.000)
Semi-Detached House	base case	base case	base case	base case	base case	base case	base case	base case
End Terrace House	-0.0573 (.000)	-0.0841 (.156)	-0.0691 (.008)	-0.1054 (.000)	0.0001 (.997)	-0.2559 (.014)	-0.0766 (.000)	-0.0657 (.023)
Terrace House	-0.0874 (.000)	-0.1046 (.053)	-0.0799 (.001)	-0.0859 (.000)	-0.0056 (.818)	-0.0158 (.806)	-0.0829 (.000)	-0.044 (.104)
BG1			-0.0563 (.765)	0.0929 (.641)			0.0591 (.699)	-0.4009 (.030)
BG2				0.3694 (.052)		-0.0443 (.736)		0.4335 (.018)
BG3	-0.0614 (.269)	-0.0632 (.016)	-0.0375 (.128)	-0.0509 (.423)	0.0019 (.982)		0.0677 (.226)	-0.0262 (.880)
BG4	0.0391 (.289)	base case	base case	0.1163 (.005)	base case	0.089 (.271)	0.018 (.496)	0.1041 (.002)
BG5			-0.129 (.277)		0.0774 (.117)	0.047 (.631)		0.1492 (.054)
BG8	0.0978 (.289)	0.1497 (.077)	0.0148 (.551)	-0.0438 (.621)	0.0644 (.119)	-0.2127 (.118)	0.0746 (.278)	0.0948 (.031)
BG9		-0.1015 (.675)			0.1622 (.038)	-0.0189 (.783)		0.1846 (.000)
BG10						-0.0609 (.439)		0.2077 (.142)
BG19	-0.8608 (.000)		-0.1435 (.117)	0.2484 (.006)	0.1522 (.024)	-0.0598 (.501)	0.1369 (.030)	0.0857 (.003)
BG20	base case	0.0814 (.353)	0.0329 (.419)	-0.0969 (.004)	0.0549 (.170)	-0.2742 (.041)	-0.0688 (.000)	-0.0009 (.979)
BG21	0.0885 (.000)	0.1886 (.026)	0.0363 (.391)	0.0316 (.285)	0.1122 (.005)	-0.0954 (.099)	base case	base case
BG24			-0.1849 (.375)		0.2013 (.026)	base case	0.1914 (.008)	0.1513 (.000)
BG25						0.2308 (.080)		
BG30	0.0364 (.267)	0.0756 (.755)	-0.0452 (.776)	-0.106 (.000)	0.1328 (.041)	-0.269 (.049)	-0.0216 (.298)	-0.0568 (.106)

Variable	Parameter Estimates by Submarket (p-values in brackets)							
	1	2	3	4	5	6	7	8
BG31	0.0701 (.112)	0.0826 (.973)	0.2336 (.004)	base case	0.0771 (.388)	-0.2271 (.039)	0.0406 (.271)	-0.1033 (.039)
BG32	0.0798 (.694)		0.4211 (.000)	0.1032 (.017)	0.2838 (.011)	-0.0036 (.964)	-0.0058 (.904)	-0.019 (.710)
BG35						0.0362 (.805)		0.0271 (.760)
BG36	0.0206 (.858)	0.0901 (.971)	0.0125 (.916)	-0.0825 (.290)	-0.3153 (.027)			-0.2008 (.289)
<i>Neighbourhood Variables:</i>								
Wealth	-0.1032 (.000)	0.001 (.976)	-0.0561 (.002)	-0.1291 (.000)	-0.1137 (.000)	-0.1614 (.000)	-0.0646 (.000)	-0.084 (.000)
Ethnicity	-0.1013 (.001)	0.0479 (.020)	-0.0432 (.018)	-0.1108 (.006)	-0.0452 (.002)	-0.1469 (.000)	-0.0851 (.001)	-0.1001 (.004)
Age	-0.0057 (.722)	0.0248 (.330)	0.0043 (.699)	0.0352 (.002)	0.0452 (.006)	0.0549 (.001)	0.0167 (.045)	0.0611 (.000)
Family	-0.0331 (.004)	-0.0267 (.116)	-0.0313 (.004)	-0.0408 (.001)	-0.0349 (.010)	-0.028 (.181)	-0.01 (.361)	-0.0162 (.258)
<i>Environmental Variables:</i>								
Road Noise	0.0018 (.165)	0.0035 (.070)	-0.0053 (.000)	-0.0028 (.046)	-0.0055 (.002)	-0.0021 (.445)	-0.0018 (.047)	-0.0025 (.038)
Rail Noise	-0.0084 (.081)	-0.0068 (.168)	-0.0063 (.063)	-0.0135 (.001)	-0.005 (.316)	-0.0049 (.729)	0.0001 (.974)	-0.0085 (.039)
Air Noise	-0.016 (.001)		-0.0154 (.279)	0.0032 (.467)	0.0339 (.125)	-0.023 (.904)	-0.0063 (.169)	0.0033 (.819)
Park Views	-1E-04 (.517)	0.0007 (.038)	-0.0001 (.495)	0.0002 (.283)	-6E-05 (.779)	-0.0003 (.231)	2E-05 (.878)	0.0003 (.015)
Water Views	0.0048 (.036)	-0.0188 (.053)	0.0003 (.426)	-0.0086 (.015)	0.0047 (.139)	0.0019 (.375)	-0.0027 (.304)	0.0001 (.805)
<i>Locational Variables:</i>								
CBD	6E-05 (.153)	0.0001 (.260)	-8E-05 (.167)	-8E-05 (.054)	-0.0002 (.000)	-3E-05 (.766)	1E-06 (.963)	-7E-05 (.063)
Airport	-7E-05 (.059)	-0.0001 (.110)	7E-05 (.222)	-7E-05 (.095)	0.0001 (.029)	-4E-05 (.711)	-8E-05 (.003)	-6E-05 (.075)
Landfill	5E-05 (.005)	-6E-05 (.023)	-1E-05 (.686)	4E-05 (.012)	-4E-05 (.026)	1E-05 (.717)	4E-05 (.001)	3E-05 (.058)
Industry A	-1E-05 (.418)	5E-05 (.028)	4E-05 (.031)	1E-05 (.270)	2E-05 (.257)	3E-05 (.308)	1E-05 (.079)	2E-05 (.027)
Industry B	-2E-05 (.321)	0.0001 (.011)	2E-05 (.565)	-4E-05 (.030)	7E-06 (.799)	-6E-05 (.231)	-2E-05 (.065)	-7E-05 (.000)
Park	1E-05 (.444)	-1E-05 (.721)	5E-05 (.025)	-3E-05 (.061)	2E-05 (.444)	0.0002 (.000)	-1E-05 (.348)	-1E-05 (.480)
Railway Station	-1E-05 (.228)	6E-06 (.775)	-4E-05 (.021)	-6E-06 (.545)	-1E-05 (.549)	4E-05 (.186)	-1E-05 (.065)	-2E-05 (.006)
Shops	0.0191 (.046)	-0.0145 (.103)	0.025 (.006)	-0.0053 (.612)	-0.0068 (.502)	-0.0033 (.849)	-0.0332 (.000)	-0.0467 (.000)
Primary Schools	0.1134 (.009)	0.1246 (.049)	0.2895 (.000)	0.2072 (.000)	0.1067 (.077)	0.0764 (.482)	0.0741 (.013)	0.0914 (.073)
<i>K</i>	41	37	46	44	47	57	44	52
<i>h</i>	.694	.551	.574	.918	.603	.752	.547	.538
<i>N</i>	1484	1016	1523	1362	1058	424	2341	1432

## References

- Abraham, J. M., W. N. Goetzmann and S. M. Wachter (1994). "Homogenous groupings of metropolitan housing markets", *Journal of Housing Economics*, 3 (3), pp. 186-206.
- Allen, M. T., T. M. Springer and N. G. Waller (1995). "Implicit pricing across residential rental markets", *Journal of Real Estate Finance and Economics*, 11, pp 137-151.
- Amemiya, T., (1979). "The estimation of a simultaneous-equation Tobit model", *International Economic Review*, 20(1), pp. 169-181.
- Anglin, P. M., and R. Gencay (1996). "Semiparametric estimation of a hedonic price function", *Journal of Applied Econometrics*, 11, pp 633-648.
- Bajic, V., (1993). "Automobiles and implicit markets: An estimate of a structural demand model for automobile characteristics", *Applied Economics*, 25(4), pp. 541-551.
- Banfield, J. D., and A. E. Raftery (1993). "Model-based Gaussian and non-Gaussian clustering", *Biometrics*, 49, pp 803-821.
- Bartik, T. J., (1988). "Measuring the benefits of amenity improvements", *Land Economics*, 64(2), pp 172-183.
- Bateman, I. J., B. H. Day, I. Lake and A. A. Lovett (2001). *The Effect of Road Traffic on Residential Property Value: A Literature Review and Hedonic Pricing Study*, Edinburgh: Scottish Executive and The Stationary Office.
- Bateman I.J., B. H. Day and I. Lake (2004). *The Valuation of Transport-Related Noise in Birmingham*, Technical Report to the Department for Transport. Norwich: University of East Anglia.
- Bell, K., and N. E. Bockstael (2000). "Applying the generalised method of moments approach to spatial problems involving micro-level data", *Review of Economics and Statistics*, 82(1), pp 72-82.
- Blomquist, N. S., (1989). "Comparative statics for utility maximization models with nonlinear budget constraints", *International Economic Review*, 30, pp.275-296.
- Bourassa, S.C., F. Hamelink, M. Hoesli and B. D. MacGregor (1999). "Defining housing submarkets", *Journal of Housing Economics*, 8, pp. 160-183.
- Boyle, K. J., P. J. Poor and L. O. Taylor (1999). "Estimating the demand for protecting freshwater lakes from eutrophication", *American Journal of Agricultural Economics*, 81(5), pp. 1118-22.
- Brown, J. N., and H. S. Rosen (1982). "On the estimation of structural hedonic price models", *Econometrica*, 50(3), pp 765-768.
- Cheshire, P. and S. Sheppard (1998). "Estimating the demand for housing, land and neighbourhood characteristics", *Oxford Bulletin of Economics and Statistics*, 60(3), pp357-382.
- Cliff, A., and J. K. Ord (1972). "Testing for spatial autocorrelation among regression residuals", *Geographical Analysis* 4, pp. 267-284.
- Edlefsen, L. E. (1981), "The comparative statics of hedonic price functions and other nonlinear constraints", *Econometrica*, 49, pp.1501-1520.
- Ekeland, I., J. J. Heckman and L. Nesheim (2002). "Identifying hedonic models", *American Economic Review*, 92(2), pp. 304-309.
- Ekeland, I., J. J. Heckman and L. Nesheim (2004). "Identification and estimation of hedonic models", *Journal of Political Economy*, 112(1), pp. 60-109.
- European Commission (EC) (1996). *Future Noise Policy – European Commission Green Paper*. Report COM(96) 540 final, Brussels: European Commission.
- European Commission (EC) (2002). *Directive 2002/49/EC for the assessment and management of environmental noise*, Official Journal of the European Communities L189/12, Brussels: European Commission.
- Fraley, C., and A. E. Raftery (1998). "How many clusters? Which clustering method? Answers via model-based cluster analysis", *The Computer Journal*, 41(8), pp. 578-588.
- Gibbons, S., and S. Machin (2003). "Valuing English primary schools", *Journal of Urban Economics*, 53(2), 15.

- Goetzmann, W. N., and S. M. Wachter (1995). "Clustering methods for real estate portfolios", *Real Estate Economics*, 23 (3), pp. 271-310.
- Grant-Muller SM, Mackie PJ, Nellthorp J and Pearman AD (2001), 'Economic Appraisal of European Transport Projects - The State of the Art Revisited' in *Transport Reviews* 21(2), pp. 237-261.
- Haab, T. C., and K. E. McConnell. (2003). *Valuing Environmental and Natural Resources: The Econometrics of NonMarket Valuation*. Northampton, MA: Edward Elgar.
- Halvorsen, R., and P. Palmquist (1980). "The interpretation of dummy variables in semilogarithmic equations", *American Economic Review*, 70, pp. 474-475.
- Hanemann, W. M., (1991). "Willingness to Pay and Willingness to Accept: How Much Can They Differ?" *American Economic Review*, 81(3), pp.635-47.
- Hayashi and Morisugi (2000), 'International comparison of background concept and methodology of transportation project appraisal', *Transport Policy*, 7, pp. 73-88.
- Heckman, J., R. Matzkin and L. Nesheim (2003). "Simulation and estimation of hedonic models", *CENMAP Working Paper*, CWP10/03, Institute of Fiscal Studies, UCL, London.
- Hoesli, M., and B. D. MacGregor (1995). "The classification of local property markets in the UK using cluster analysis", in *The Cutting Edge: Proceedings of the RICS Property Research Conference, 1995, Vol. 1*, London: Royal Institution of Chartered Surveyors.
- Jain, A. K., M. N. Murty and P. J. Flynn (1999). "Data Clustering: A Review", *ACM Computing Surveys*, 31(3), pp. 264-323.
- Kelejian, H. H., and I. R. Prucha (1999). "A generalised moments estimator for the autoregressive parameter in a spatial model", *International Economic Review*, 40(2), pp 509-533.
- Kelejian, H. H., and Robinson, D. P., (1992). "Spatial autocorrelation: A new computationally simple test with an application to per capita county police expenditures", *Regional Science and Urban Economics*, 22, pp 317-331.
- McConnell, K. E., and T. T. Phipps (1987). "Identification of preference parameters in hedonic models: Consumer demands with nonlinear budgets", *Journal of Urban Economics*, 22, pp 35-52.
- Murray, M. P., (1982). "Mythical demands and supplies for proper estimation of Rosen's hedonic price model", *Journal of Urban Economics*, 14, pp 327-337.
- Nesheim, L., (2002). "Equilibrium sorting of heterogeneous consumers across locations: theory and empirical implications", *CEMMAP Working Paper*, CWP08/02, Institute of Fiscal Studies, Department of Economics, University College London.
- Newey, W. K. (1987). "Efficient Estimation of Limited Dependent Variable Models with Endogenous Explanatory Variables", *Journal of Econometrics*, 36(3), pp. 231-50.
- Odgaard T., Kelly C. and Laird J. (2005) Current practice in project appraisal in Europe: Analysis of country reports. HEATCO Work Package 3 Deliverable 1. IER, University of Stuttgart.
- Palmquist, R. B., (1988). "Welfare measurement for environmental improvements using the hedonic model: The case of nonparametric marginal prices", *Journal of Environmental Economics and Management*, 15, pp 297-312.
- Palmquist, R. B, and A. Isangkura (1999). "Valuing air quality with hedonic and discrete choice models", *American Journal of Agricultural Economics*, 81(5), pp1128-33.
- Posse, C., (2001). "Hierarchical model-based clustering for large datasets", *Journal of Computational and Graphical Statistics*, 10 (3), pp. 464-486.
- Randall, A., and J. R. Stoll (1980). "Consumer's Surplus in Commodity Space". *American Economic Review*, 70(3), pp. 449-57.
- Ridker, R. G. (1967). *Economic Costs of Air Pollution: Studies in Measurement*. New York: Praeger.
- Ridker, R. G. and J. A. Henning (1967). "The Determinants of Residential Property Values with Special Reference to Air Pollution", *Review of Economics and Statistics*, 49: 246-57.
- Robinson, P. M., (1988). "Root-N-consistent semiparametric regression", *Econometrica*, 56, pp 931-954.

- Rosen, S. (1974). "Hedonic prices and implicit markets: Production differentiation in pure competition", *Journal of Political Economy*, 82, pp 34-55.
- Smith, R. J., and R. W. Blundell (1986). "An exogeneity test for a simultaneous equation Tobit model with an application to labor supply", *Econometrica*, 54(3), pp.679-85.
- Strand, J., and M. Vagnes (2001), "The relationship between property values and railroad proximity: a study based on hedonic prices and real estate brokers' appraisals", *Transportation*, 28, pp. 137-156.
- Watkins, C. A., (2001). "The definition and identification of housing submarkets", *Environment and Planning A*, 33, pp. 2235-2253.
- Weesie, J., (1999). "Seemingly unrelated estimation: An application of the cluster-adjusted sandwich estimator", *Stata Technical Bulletin*, 52, pp.34-47
- Willig, R. D., (1976). "Consumer's surplus without apology", *American Economic Review*, 66(4), pp 589-597.
- Zabel, J. E. and K. A. Kiel (2000). "Estimating the demand for air quality in four U.S. cities", *Land Economics*, 76(2), pp 174-194.