Environ Resource Econ (2007) 37:211–232
DOI 10.1007/s10640-007-9121-8

Beyond implicit prices: recovering theoretically consistent and transferable values for noise avoidance from a hedonic property price model

Brett Day · Ian Bateman · Iain Lake

Published online: 16 May 2007

Abstract Using a two-stage hedonic pricing methodology we estimate a system of structural demand equations for different sources of transport-related noise. In the first stage, we identify market segments using model-based clustering techniques and estimate separate hedonic price functions (HPFs) for each segment. In so doing, we show how a semiparametric spatial smoothing estimator outperforms other standard specifications of the HPF. In the second stage, we control for non-linearity of the budget constraint and identify demand relationships using techniques that account for problems of endogeneity and censoring of the dependent variable. Our estimated demand functions provide 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 · Non-market valuation · Hedonic pricing · Model-based clustering · Partial linear model · Spatial smoothing · 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 health experts consider unacceptable (EC 1996). Indeed, the *Directive on the Assessment and Management of Environmental*

B. Day (✉) · I. Bateman
Centre for Social and Economic Research on the Global Environment,
School of Environmental Sciences, University of East Anglia, Norwich NR4 7TJ, UK
e-mail: brett.day@uea.ac.uk

I. Lake
Centre for Environmental Risk, School of Environmental Sciences,
University of East Anglia, Norwich NR4 7TJ, UK

Noise requires member states of the EU to draw up "action plans" to combat noise pollution (EC 2002).

Of course, noise reduction measures are costly and establishing whether such measures are worthwhile requires comparison of these costs with the benefits of a quieter environment. One framework within which this comparison can be made is cost–benefit analysis (CBA).[1] Attempts to integrate noise considerations into CBA have been hampered, however, by the absence of monetary valuations of the welfare benefits of noise reductions. For example, until very recently the UK, like many other nations, undertook CBA of transport projects but failed to incorporate monetarised estimates of noise (dis)benefits in the analysis (Odgaard et al. 2005).

For David Pearce, of course, this sort of omission was like a red rag to a bull. He argued vociferously for policy appraisal that provided a "level playing field" (a favoured phrase) upon which environmental costs and benefits were afforded the same consideration as other costs and benefits. Indeed, in recent years, David dedicated much of his research effort to the field of non-market valuation and encouraged others to do so also. Moreover, he used his considerable influence to promote the outputs of such research and to persuade and cajole decision-makers into using non-market valuation in policy appraisal. The research reported in this paper is a testament to his achievements in both those pursuits. The work results from the last PhD thesis that David supervised before retiring from academic life (Day 2005). Moreover, the findings of the research have fed directly into the policy-making process, being the basis for the Department for Transport's guidance on the appraisal of transport related noise impacts of multi-modal transport plans (DfT 2006; Nellthorp et al. forthcoming).

In this research the method of hedonic pricing is applied to property market data collected from the City of Birmingham. As is well known, hedonic pricing uses techniques of multiple regression to isolate implicit prices; that is, differences in property prices attributable to marginal differences in property characteristics.[2] Numerous previous studies have used this technique to report implicit prices for noise from a variety of urban property markets (for a review see Bateman et al. 2001). According to the theory, an implicit price for noise provides a monetary estimate of the welfare benefit of a marginal change in noise. However, like any market price, implicit prices are market specific, reflecting only the particular conditions of supply and demand that exist in that property market. Accordingly, an implicit price for noise estimated in one urban area offers little indication of the benefits of changes in noise experienced elsewhere. 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.[3] 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; Boyle et al. 1999; Cheshire and Sheppard 1998; Palmquist and Isangkura 1999; and Zabel and Kiel 2000). Moreover,

[1] See the discussions of CBA in general given by Turner (2007) and the early use of hedonic pricing within CBA given by Barde (2007) in this collection.

[2] The most frequently applied alternatives to hedonic pricing are stated preference methods as discussed by Carson and Groves (2007) in this collection.

[3] Haab and McConnell (2003) state that "With a few exceptions, researchers have abandoned any attempts to recover preferences, and work instead with the hedonic price function" (p. 251).

as far as the authors are aware it is the only study to estimate such demand relationships for avoidance of transport-related noise using property price data.

Our study is unique in other ways. As briefly described in Sect. 3, our data is one of the largest and richest property market datasets yet compiled. 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 other fields 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 non-parametrically whilst simultaneously accounting 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 correct the data for problems arising from the 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.

2 The method of hedonic pricing

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

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

where z is a vector (with elements z_k; $k = 1, 2, \ldots, K$) of property attributes (including noise exposure) 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 $\partial P(z)/\partial z_k \equiv p_{z_k}$.

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 given data from just one market (Ekeland et al. 2002). 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. 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, where individual attributes are bundled into a composite good commanding a single market price, it is not necessarily the case that the implied price per unit of an attribute is a constant. Under such circumstances the standard duality results of consumer theory no longer hold. In particular, Marshallian demand functions do not necessarily manifest themselves as downward sloping curves in price-quantity space. In our application, we follow the procedure proposed by Murray (1982) 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.

3 The data

Birmingham is the UKs second largest city with a diverse housing stock and an diverse population. Details of the selling price, sale date and full address for each residential property transaction in the city and surrounds during 1997 were obtained from 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 (apartments were excluded), whether the property was detached, semi-detached, in or at the end of a terrace and whether the property had central heating and access to off-road parking. Further, the VOA classifies properties by age and style of construction into one of around 30 property types that provide an additional indication of property quality.

Addresses were geolocated using a GIS. Subsequently digital datasets were used to estimate each property's garden area and to calculate car travel times and walking distances from each property to (dis)amenities including the city centre, airport, primary schools, parks, 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. A similar index was constructed to describe proximity to local commercial centres based on the quantity and proximity of shops.

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.

Digital noise maps provided estimates of each property's exposure to daytime road, rail and air traffic noise measured in decibels LEQ (DEFRA 2000). LEQ summarises a location's noise exposure by determining the continuous noise level that has the equivalent energy content as the varying noise signal experienced at that location.

Data on the socioeconomic composition of property neighbourhoods were drawn from the 1991 UK census. For administrative purposes Birmingham is divided into 39 political units known as wards (see panel A of Fig. 1). Each ward is further divided into a number of enumeration districts (EDs) which form the smallest area over which census data is provided

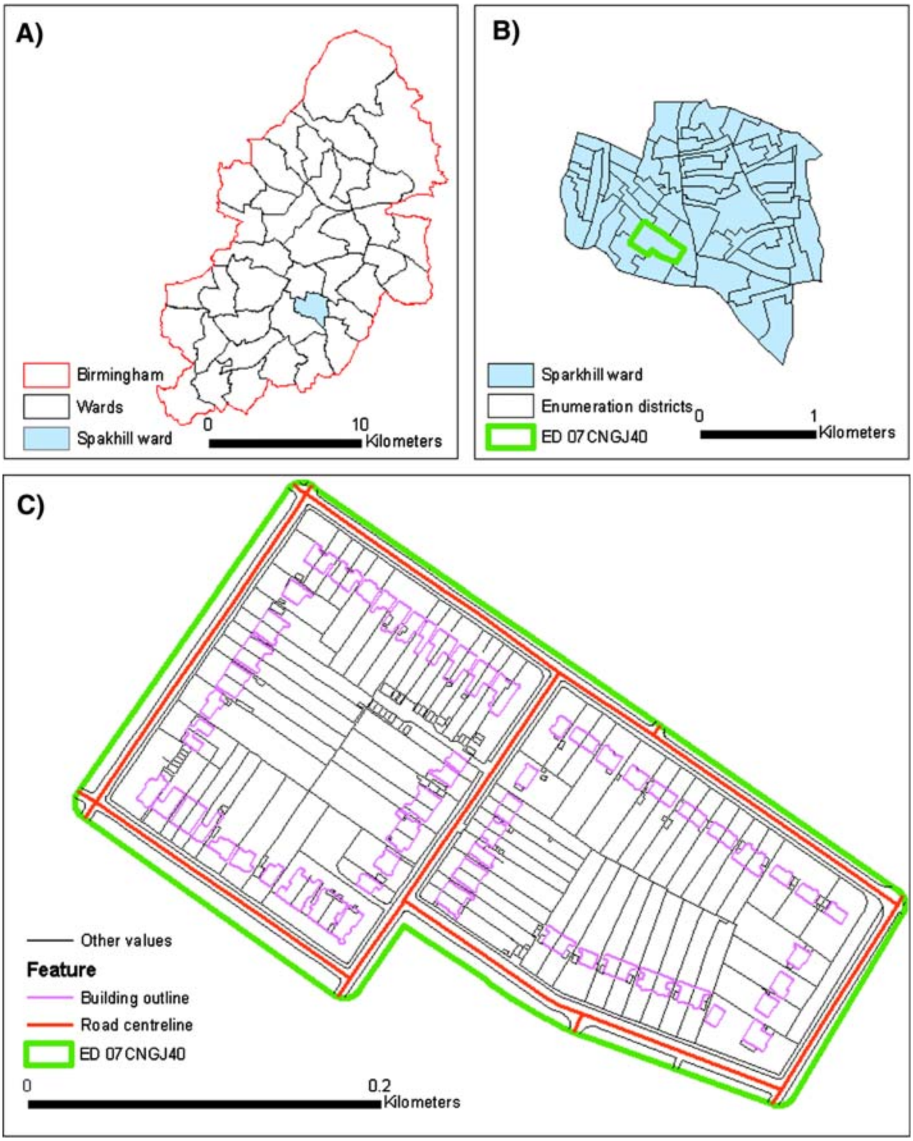

Fig. 1 Hierarchy of administrative areas in Birmingham

(see panel B of Fig. 1). Birmingham is divided into 1,940 EDs, with each ED containing an average of 191 households (see panel C of Fig. 1).

Census data provides hundreds of variables describing the socioeconomic characteristics of households residing in each ED. Factor analysis was employed to condense this excess of neighbourhood attributes into a more manageable set of indices. 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.

Applying these various techniques, data records were successfully compiled for some 10,848 residential property transactions in Birmingham in 1997. A detailed description of the data can be found in Bateman et al. (2004).

4 First stage: econometric model

It is generally 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 the property market comprises households that form "consumer groups" and a housing stock that comprises distinct "product groups". The matching of consumer groups with product groups gives rise to segmentation of the market whilst the differing conditions of supply and demand within each segment imply that 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 is to identify segments as groups of properties with similar structural attributes inhabited by households with similar socioeconomic characteristics. Cluster analysis has been used for this purpose by previous researchers (e.g. Abraham et al. 1994; Bourassa et al. 1999; and Goetzmann and Wachter 1995). However, such studies have used 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 employ *model-based cluster analysis* (Fraley and Raftery 1998) a recently developed clustering technique that answers both these important questions. As far as the authors are aware, this is the first economic application of this technique.

The starting point for any cluster analysis is to define a set of M variables by which the observations are to be clustered. The objective of the analysis is to identify the various groups of observations that form distinct clusters in the M-space defined by those variables. The fundamental assumption of model-based clustering is that the members of each group form a cluster whose location and shape can be approximated by a P-dimensional probability distribution. The data observed by the researcher is the composite of data points drawn at random from a finite number of such clusters.

Fraley and Raftery (1998) examine the case in which the members of each cluster are assumed to be Gaussian distributed. They show that given an initial assumption concerning the number of clusters, the likelihood function for the clustering problem can be maximised using a simple EM algorithm. Moreover, the number of clusters best characterising the data can be identified by comparing models assuming different numbers of clusters according to Bayes Information Criteria. 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. Following standard practice (e.g. Allen et al. 1995) 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 the functional form of HPFs. Indeed, recent research simulating markets for differentiated goods reveals that the equilibrium HPF will tend to be highly non-linear or even kinked (Heckman et al. 2003; Nesheim 2002). Moreover, it is widely recognised that location plays a major role in determining property prices. Even with the advances in data collation provided by 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) introduced flexibility into the functional form of the HPF by using a *partially linear model* (PLM). The PLM specification allows prices to depend on some covariates linearly and others non-parametrically. We allow for non-linearities by including in the non-parameteric part of the specification those variables that we believe *a priori* to be key determinants of market price.

Gibbons and Machin (2003) also exploit the PLM specification in order to account for omitted locational covariates. They spatially smooth their data by including the location of each observation in the non-parametric part of the model.

Here we combine the specifications of Anglin and Gencay (1996) and Gibbons and Machin (2003) by including both spatial coordinates and property characteristics in the non-parametric part of the PLM. As such our specification allows for non-linearity with respect to certain attributes whilst controlling for spatial variation in prices that is not captured by the regressors in our data set. We refer to this model as the PLM with spatial smoothing (PLM + SS). Our specification is as follows;

$$\ln P_i = z_i \boldsymbol{\beta} + q\left(\boldsymbol{x}_i, \boldsymbol{c}_i\right) + \varepsilon_i \tag{2}$$

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

Robinson (1988) shows that the model in (2) can be rewritten as;

$$\ln P_i - E\left[\ln P|\boldsymbol{x}_i, \boldsymbol{c}_i\right] = \left(z_i - E\left[z|\boldsymbol{x}_i, \boldsymbol{c}_i\right]\right)\boldsymbol{\beta} + \varepsilon_i \tag{3}$$

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

- Estimate the unknown conditional means $E\left[\ln P|\boldsymbol{x}_i, \boldsymbol{c}_i\right]$ and $E\left[z|\boldsymbol{x}_i, \boldsymbol{c}_i\right]$ using non-parametric regression.
- Substitute these estimates for the unknown functions in (3) and apply standard regression techniques to recover $\boldsymbol{\beta}$.

We use a Nadaraya–Watson estimator with a Gaussian kernel to estimate the quantities $E\left[\ln P|\boldsymbol{x}_i, \boldsymbol{c}_i\right]$ and $E\left[z|\boldsymbol{x}_i, \boldsymbol{c}_i\right]$. We fix the bandwidth for the elements of $\boldsymbol{c}_i$ to a pre-specified value, b, such that each observation is spatially smoothed over the same circular area. We adopt a bandwidth matrix for $\boldsymbol{x}_i$ that is proportional to the sample covariance matrix of $\boldsymbol{x}$ and estimate this proportion, h, using cross-validation. Finally, to allow for differences in the density of the data, we apply an adaptive-kernel procedure to provide an observation-specific bandwidth, h_i.

The data can be expressed as differences from the conditional expectations; that is $\tilde{y}_i = \ln \tilde{P}_i = \ln P_i - E\left[\ln P|\boldsymbol{x}_i, \boldsymbol{c}_i\right]$ and $\tilde{z}_i = z_i - E\left[z|\boldsymbol{x}_i, \boldsymbol{c}_i\right]$. If the $\tilde{y}_i$ are stacked to form the $N \times 1$ vector $\tilde{\boldsymbol{Y}}$ and the $\tilde{z}_i$ vectors stacked to form the $N \times k$ matrix $\tilde{\boldsymbol{Z}}$, then the parameters can be estimated according to;

$$\tilde{\boldsymbol{\beta}} = \left(\tilde{\boldsymbol{Z}}'\tilde{\boldsymbol{Z}}\right)^{-1}\tilde{\boldsymbol{Z}}'\tilde{\boldsymbol{Y}} \tag{4}$$

Despite spatially smoothing the data, some locational features may not be captured by the model. Since these features are held in common by properties in close proximity, the regression residuals may exhibit spatial correlation. In that case, (4) will return inefficient estimates of the parameters and biased estimates of their standard errors.

We assume that spatial error dependence will take the form;

$$\boldsymbol{\varepsilon} = \rho \boldsymbol{W}\boldsymbol{\varepsilon} + \boldsymbol{u} \tag{5}$$

where $\boldsymbol{\varepsilon}$ is the $[N \times 1]$ vector of regression residuals, ρ is the error dependence parameter to be estimated and $\boldsymbol{u}$ is the usual $N \times 1$ vector of random error terms with expected value zero and covariance matrix $\sigma^2\boldsymbol{I}$. Notice that $\rho = 0$ implies $\boldsymbol{\varepsilon} = \boldsymbol{u}$ and there is no spatial dependence in the data. The spatial association between properties is captured by the $N \times N$ weighting matrix, $\boldsymbol{W}$. The diagonal elements of $\boldsymbol{W}$ are zero, whilst the w_{ij}th off-diagonal element is initially set to one if the ith and jth 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 $\hat{\rho}$, whilst the second step estimates the parameters of the HPF using the feasible generalised least squares estimator;

$$\hat{\boldsymbol{\beta}} = \left(\tilde{\boldsymbol{Z}}' \left((\boldsymbol{I} - \hat{\rho}\boldsymbol{W})'(\boldsymbol{I} - \hat{\rho}\boldsymbol{W})\right)^{-1} \tilde{\boldsymbol{Z}}\right)^{-1} \tilde{\boldsymbol{Z}}' \left((\boldsymbol{I} - \hat{\rho}\boldsymbol{W})'(\boldsymbol{I} - \hat{\rho}\boldsymbol{W})\right)^{-1} \tilde{\boldsymbol{Y}} \tag{6}$$

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

$$\begin{aligned} p_{z_k} &= \exp\left(E\left[\ln P|\boldsymbol{x}_i, \boldsymbol{c}_i\right] + \left(\boldsymbol{z}_i - E\left[\boldsymbol{z}|\boldsymbol{x}_i, \boldsymbol{c}_i\right]\right)\hat{\boldsymbol{\beta}}\right)\hat{\beta}_{z_k} \\ &= \exp\left(\ln \hat{P}_i\right)\hat{\beta}_{z_k} \end{aligned} \tag{7}$$

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. Having post-processed the cluster definitions to incorporate geographic information into the classification of observations to clusters, our analysis provides evidence for the existence of eight clusters. Table 1 summarises the characteristics of the properties and inhabitants in each cluster.

We partition the data according to the clustering classification and estimate separate HPFs for each partition using the PLM+SS model. We include in the regression analysis all of the variables described Sect. 3. In particular, our measures of noise exposure from road, rail and air traffic are taken as dB in excess of 55 dB LEQ, a threshold commonly regarded as a measure of background noise in urban areas. Along with spatial coordinates, we include in the non-parametric part of the specification variables describing each property's floor area, garden area, construction age and sale date. The spatial smoothing bandwidth, b, was fixed at a distance approximating the radius of a ward.

We examined the residuals from the PLM+SS regressions 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 offer

Table 1 Characteristics of Birmingham market segments

Market segment	Property size	Income	Ethnicity	Children
1. Low-income, White	Small	Low	White	Many
2. Low-income, Ethnic	Small	Low	Ethnic	Many
3. Middle-income, Young, No children	Small	Medium	Mixed	Few
4. Middle-income, White	Small	Medium	White	Some
5. Middle-income, Mixed	Medium	Medium	Mixed	Some
6. High-income, Large properties	Large	High	Mixed	Few
7. High-income, Small properties	Small	High	White	Some
8. High-income, Medium properties	Medium	High	White	Many

Table 2 Unadjusted R^2 statistics for different specifications of market segment HPFs

Market segment	N	LM	PLM	PLM + SS
1. Low-income, White	1,484	0.425	0.454	0.559
2. Low-income, Ethnic	1,016	0.446	0.510	0.618
3. Middle-income, Young, No children	1,523	0.647	0.721	0.764
4. Middle-income, White	1,362	0.657	0.700	0.709
5. Middle-income, Mixed	1,058	0.649	0.722	0.778
6. High-income, Large properties	424	0.810	0.875	0.890
7. High-income, Small properties	2,341	0.587	0.640	0.695
8. High-income, Medium properties	1,432	0.751	0.820	0.839

no information on the HPF for that market.[4] After individual investigation we trimmed 1.9% of the observations from the data and re-estimated (4). We examined the residuals from this regression for evidence of spatial correlation over an area approximating 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) further supports this conclusion. As such we apply the GM estimator (6) to estimate the parameters of the HPFs.

To assess the overall fit of the estimated HPFs and the performance of the PLM + SS model, observe Table 2. The second column of Table 2 indicates the number of observations in each segment. The remaining columns list unadjusted R^2 statistics (that is, measures of the proportion of the total variation in the dependent variable explained by the model) for each HPF estimated using three nested model specifications; a standard linear specification (LM), the partially linear specification without spatial smoothing (PLM) and the partially linear model with spatial smoothing (PLM + SS).

Comparison of the LM and PLM reveals that the latter semiparametric estimator affords significant gains in explanatory power; on average across the segments a 6% increase in the R^2 score. Clearly, the linear specification provides a relatively poor approximation to the functional relationship between property prices and the variables contained in the non-parametric part of the specification. Moreover, moving from the PLM to the PLM + SS model, the R^2 scores increase by a further 5% on average. Notice that in the two low-income market segments (segments 1 and 2) where the fit of the LM is particularly poor, the PLM + SS provides the relatively greatest increase in model fit (13% and 17%, respectively). Evidently,

[4] 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.

despite the regressors in our model containing numerous descriptions of locational quality (e.g. neighbourhood socioeconomics as well as proximity to the city centre, airport, schools, parks, shops, railway stations, industrial and landfill sites), certain features of the spatial environment that have a substantive impact on property prices are not captured by the PLM model. The PLM + SS model is clearly superior to the other specifications in its ability to describe variation in property prices.

Of course, our particular concern is not goodness of fit but the impact of model specification on estimates of the coefficients of the HPF, particularly those relating to transport noise. Tables 3–5 document the coefficients for road, rail and air traffic noise, respectively. The second column in each of these tables records the number of properties in each market segment exposed to levels of noise from that source above the urban background level (55 dB). Notice that our data contains many more properties exposed to above-threshold levels of road noise than rail noise. Accordingly, we would expect our estimates of the road noise parameters to be estimated with relatively greater precision than those for rail noise. Notice further that the considerable geographical separation in market segments means that aircraft noise is concentrated largely in the three segments (1, 4 and 7) that are closest to Birmingham International Airport.

The subsequent columns of Tables 3–5 record coefficient estimates for the LM, PLM and PLM + SS specifications. Moreover, the final column records coefficients from a Bayesian update of the PLM + SS model that simultaneously constrains the three noise coefficients to the negative orthant (for computational details see Geweke 1986). The standard errors reported in the tables are corrected for spatial error dependence using the GM estimator described previously. Since we assume that exposure to noise, *ceteris paribus*, can only reduce the selling price of a property, significance is judged through a one-tailed t-test of the

Table 3 Road noise parameters of the HPFs (standard errors in brackets)

Market segment	Obs. above 55 dB threshold	LM	PLM	PLM + SS	PLM + SS constrained
1. Low-income, White	307	0.0008 (0.0014)	0.0008 (0.0014)	0.0018 (0.0013)	−0.0006† (0.0005)
2. Low-income, Ethnic	208	0.0008 (0.0019)	0.0013 (0.0019)	0.0035 (0.0019)	−0.0008† (0.0007)
3. Middle-income, Young, No children	521	−0.0047*** (0.0011)	−0.0050*** (0.0011)	−0.0053*** (0.0010)	−0.0053† (0.0010)
4. Middle-income, White	299	−0.0029** (0.0014)	−0.0030** (0.0014)	−0.0028** (0.0014)	−0.0028† (0.0013)
5. Middle-income, Mixed	271	−0.0072*** (0.0018)	−0.0067*** (0.0017)	−0.0055*** (0.0017)	−0.0057† (0.0017)
6. High-income, Large properties	167	−0.0016 (0.0027)	−0.0018 (0.0027)	−0.0021 (0.0027)	−0.0032† (0.0020)
7. High-income, Small properties	567	−0.0025 (0.0009)	−0.0026*** (0.0009)	−0.0018** (0.0009)	−0.0019† (0.0009)
8. High-income, Medium properties	383	−0.0035*** (0.0012)	−0.0033*** (0.0012)	−0.0025** (0.0012)	−0.0025† (0.0011)

† Significance test not meaningful
* Significant at 10% level of confidence
** Significant at 5% level of confidence
*** Significant at 1% level of confidence

Table 4 Road noise parameters of the HPFs (standard errors in brackets)

Market segment	Obs. above 55 dB threshold	LM	PLM	PLM + SS	PLM + SS constrained
1. Low-income, White	27	−0.0079* (0.0050)	−0.0081* (0.0050)	−0.0084** (0.0048)	−0.0091† (0.0044)
2. Low-income, Ethnic	40	−0.0052 (0.0050)	−0.0061 (0.0050)	−0.0068* (0.0049)	−0.0073† (0.0042)
3. Middle-income, Young, No children	92	−0.0077** (0.0036)	−0.0086*** (0.0035)	−0.0063** (0.0034)	−0.0065† (0.0031)
4. Middle-income, White	43	−0.0132*** (0.0040)	−0.0141*** (0.0039)	−0.0135*** (0.0039)	−0.0135† (0.0039)
5. Middle-income, Mixed	53	−0.0040 (0.0050)	−0.0084** (0.0051)	−0.0050 (0.0050)	−0.0066† (0.0040)
6. High-income, Large properties	15	−0.0026 (0.0119)	−0.0146 (0.119)	−0.0050 (0.0141)	−0.0136† (0.0096)
7. High-income, Small properties	61	0.0019 (0.0032)	0.0020 (0.0033)	0.0001 (0.0032)	−0.0025† (0.0019)
8. High-income, Medium properties	48	−0.0085** (0.0039)	−0.0078** (0.0039)	−0.0085** (0.0041)	−0.0087† (0.0039)

† Significance test not meaningful
* Significant at 10% level of confidence
** Significant at 5% level of confidence
*** Significant at 1% level of confidence

Table 5 Aircraft noise parameters of the HPFs (standard errors in brackets)

Market segment	Obs. above 55 dB threshold	LM	PLM	PLM + SS	PLM + SS constrained
1. Low-income, White	199	−0.0177*** (0.0041)	−0.0178*** (0.0040)	−0.0160*** (0.0049)	−0.0153† (0.0049)
2. Low-income, Ethnic	0				
3. Middle-income, Young, No children	17	−0.0031 (0.0099)	−0.0071 (0.0095)	−0.0154 (0.0142)	−0.0191† (0.0115)
4. Middle-income, White	195	−0.0007 (0.0040)	0.0016 (0.0040)	0.0032 (0.0044)	−0.0025† (0.0021)
5. Middle-income, Mixed	10	0.0170 (0.0187)	0.0121 (0.0188)	0.0339 (0.220)	−0.0096† (0.0085)
6. High-income, Large properties	3	0.0260 (0.0359)	0.0280 (0.0396)	−0.0226 (0.1884)	−0.1585† (0.1179)
7. High-income, Small properties	190	−0.0084** (0.0040)	−0.0089** (0.0038)	−0.0063* (0.0046)	−0.0072† (0.0040)
8. High-income, Medium properties	30	−0.0004 (0.0140)	−0.0053 (0.0130)	0.0033 (0.142)	−0.0102† (0.0080)

† Significance test not meaningful
* Significant at 10% level of confidence
** Significant at 5% level of confidence
*** Significant at 1% level of confidence

null hypothesis that the coefficient is equal to zero.[5] Since the coefficients in the final column are constrained to be non-positive we do not report significance in this case.

Consider the road noise coefficients in Table 3. In accordance with prior expectations, the majority of these parameter estimates are negative. Notice, however, that across all three unconstrained specifications, those for market segments 1 and 2 remain 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; a supposition supported by the observation that our models show the poorest fit to the data in these two cases.

Comparing across specifications, we observe that the PLM returns estimates of the road noise coefficients that are almost indistinguishable from the LM. It appears that the non-linearities captured by the PLM, whilst greatly improving the overall fit of the model, are uncorrelated with road noise. In contrast, spatially smoothing the data has a more significant impact on the road noise parameter estimates. Where the PLM + SS estimates differ markedly from the PLM (segments, 3, 5 and 7), the coefficients are seen to reduce in absolute value. In these cases, spatial smoothing appears to be capturing the impact of some price-depressing spatial process, not captured by the variables in our model, that is correlated with areas of high road noise.[6] We conclude that the PLM + SS specification, in controlling for more general spatial influences on house prices, reduces omitted variable bias and provides more accurate estimates of the road noise parameters.

The PLM + SS model returns negative road noise coefficients for market segments 3–8 that range in value from -0.0018 to -0.0055. In other words, our preferred specification indicates that a 1 dB increase in road traffic noise reduces the selling price of a property by between 0.18% and 0.55%; a measure termed the noise depreciation sensitivity index (NSDI). Encouragingly, five of these six estimates are significant at greater than a 95% level of confidence.

Now let us consider the rail traffic noise coefficients listed in Table 4. Again the estimates are generally of the expected sign except that for segment 7. Moreover, the PLM + SS estimates are significant with greater than 90% confidence in five market segments. Given the relatively small number of properties exposed to rail noise above the background level, this level of significance suggests the effects are reasonably strong. Indeed, the rail noise coefficients are generally larger in magnitude than those observed on road noise. On average the models indicate an NSDI of 0.67%.

Comparison of the PLM and PLM + SS specifications shows that where there are large differences in the coefficients, the parameters tend to fall in absolute value. Again we conclude that spatial smoothing has the effect of removing from the rail noise coefficient the extraneous influence of omitted spatial processes that correlate with locations of high rail noise.

The aircraft noise coefficients shown in Table 5 are somewhat more erratic than those for either road or rail noise. Of course, given the small number of properties exposed to aircraft noise in most segments, it is not surprising that we are unable to detect significant effects. In the three market segments which contain the majority of properties exposed to aircraft noise,

[5] Our thanks to an anonymous reviewer who indicated that a one-tailed, not two-tailed, test was appropriate in this case.

[6] One can only conjecture what these processes might be, but in addition to noise, road traffic is associated with other negative externalities including air pollution, congestion and a lack of on-street parking, and the community severance resulting from the barrier effect of major roads.

the parameters have the expected sign in market segments 1 and 7, being significant at a 99% and 90% level of confidence respectively, but not in market segment 4.

The PLM + SS specification shows clear advantages in terms of goodness of fit suggesting that there are spatial processes influencing property prices that are omitted in the less comprehensive specifications. Moreover, we have shown that accounting for these influences by spatially smoothing the data impacts upon the noise coefficients; usually to reduce their absolute value. Accordingly, we take the PLM + SS as our preferred specification. In the interests of brevity, we do not discuss the coefficient estimates on the numerous other regressors. A comprehensive discussion can be found in Bateman et al. (2004).

Using a series of pairwise Wald tests, we find that we can reject, with greater than 95% confidence, the hypothesis that the same price function characterises the data in any two of the partitions identified by the cluster analysis. We take this as evidence to support our contention that the cluster analysis has identified separate market segments (as per Allen et al. 1995).

For each household, we estimate the implicit price of each type of noise using Eq. 7. Under the reasonable assumption that noise is considered a bad, the second stage estimation requires that these implicit prices are negative. The existence of a scattering of positive (though insignificant) noise coefficient estimates, therefore, is problematic. As such, we also estimate implicit prices using the constrained coefficient estimates listed in the final columns of Tables 3–5.

It is worth making a few preliminary comments on the research findings at this juncture. First, Tables 3–5 tell a fairly consistent story. If we ignore insignificant results, our data suggest that the implicit price of aircraft noise tends to be relatively larger than that for rail noise which in turn is relatively larger than that for road noise. Of course, as previously discussed, implicit prices are determined by the complex interaction of supply and demand in the market. Accordingly, whilst we could speculate that this observed pattern is a result of a relatively greater demand for environments free from rail noise than from road noise and a greater demand still for environments free from aircraft noise, such a conclusion would only be justified if the supply conditions of those different environments were identical. This leads on to a second comment. Notice that across the various market segments, there is considerable variation in the coefficients (and, as such, implicit prices) for each particular type of noise. Again these differences might be explained as the outcome of differing supply and demand conditions in the various segments. Of course, this fact introduces a problem for policy-makers; implicit prices, by their very nature, are market specific and are not quantities that can be transferred across markets. It is this observation that motivates our desire to move beyond implicit prices and carry out a second-stage hedonic analysis in order to recover underlying preference relationships. In particular, transferring welfare estimates derived from preference relationships across markets is theoretically defensible, transferring implicit prices across markets is not.

6 The econometric model: second stage

In the second stage we pool the data from all market segments and seek to estimate a demand system for peace and quiet. Notice that we invert the measure of noise so as to frame the analysis in terms of the good ‘peace and quiet’. Moreover, the demand for peace and quiet is assumed to be censored from above at the background urban noise level of 55 dB. We set the scale for our peace and quiet measure such that 55 dB is zero and each extra dB of noise exposure is one less dB of peace and quiet. In addition, given the concentration of aircraft noise in just a few market segments, we restrict our investigation to the estimation of demand equations for the avoidance of road and rail traffic noise where we observe variation in implicit prices across the range of market segments.

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 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. We assume the same relationship holds between the implicit prices and implicit rents of property characteristics. It is mathematically convenient to suppose that the expected lifetime of a property is infinite such that for some discount rate, τ, implicit prices for noise are converted to implicit rents according to;

$$r_{z_k} = -\tau \cdot p_{z_k} \tag{8}$$

where the negation inverts rents for noise into rents for peace and quiet.

Following Palmquist and Isangkura (1999), we invoke a two-stage budgeting hypothesis such that the relevant argument in the demand relationship is not household income, but the household's total expenditure in purchasing a property.

Of course, non-linearity 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 econometric model 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 HPF. 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 procedure for dealing with this problem is that proposed by Murray (1982) and Palmquist (1988). Here a new variable is calculated for each household indicating the 'virtual' expenditure that would be necessary to purchase their chosen property if prices were in fact constant at the implicit prices characterising their chosen bundle. We include virtual 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. We convert virtual expenditure, $\tilde{y}$, into an annual equivalent, $\tilde{m}$, according to;

$$\tilde{m} = \tau \cdot \tilde{y} \tag{9}$$

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. A similar assumption is made by Palmquist and Isangkura (1999).

Our econometric analysis seeks to estimate two separate demand equations;

$$\begin{aligned} q_1^* &= \tilde{q}_1\left(r_1, r_2, \tilde{m}, \mathbf{s}\right) \\ q_2^* &= \tilde{q}_2\left(r_1, r_2, \tilde{m}, \mathbf{s}\right) \end{aligned} \tag{10}$$

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

To maintain simplicity we estimate (10) as the linear demand system;

$$q_{ki}^* = \beta_0^k + \beta_m^k \ln(\tilde{m}) + \beta_1^k r_{1i} + \beta_2^k r_{2i} + \mathbf{s}_i \boldsymbol{\gamma}_k + u_{ki} \qquad (k = 1, 2; i = 1, \ldots, N) \tag{11}$$

where β_0^k is the demand curve intercept, β_m^k is the parameter on virtual expenditure, β_1^k and β_2^k, are respectively, the parameters on the prices of avoiding road and rail noise, whilst $\boldsymbol{\gamma}_k$ is the vector of parameters on the socioeconomic demand shifters and u_{ki} is an error term.

While the linear specification contravenes restrictions implied by economic theory,[7] our choice of functional form is driven by pragmatic considerations. In particular, we need to contend with numerous econometric problems that are relatively easier to handle in a linear model. For example, numerous households choose properties at the corner solution imposed by the existence of an urban background level of noise. Accordingly our dependent variable is censored from above taking the form;

$$q_{ki} = \begin{cases} q_{ki}^* & if \ \ q_{ki}^* < 0 \\ 0 & \text{otherwise} \end{cases} \qquad (k = 1, 2; \quad i = 1, 2, \ldots, N) \tag{12}$$

One standard way of analysing censored data of this type is to apply the Tobit model. Unfortunately, we have an additional problem with which we have to contend that may compromise the Tobit model; both implicit prices (r_k) and the virtual expenditure ($\tilde{m}$) that is calculated from them, are endogenous in the demand relationship.

The most general technique for handling endogenous variables is the method of instrumental variables (IV). This method requires the identification of a set of instrumental variables each of which must be both correlated with the endogenous variables as well as 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_{kj} 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). That is to say, we select the values chosen by households living in similar properties in similar neighbourhoods that are close by. Notice that we do not allow for households living in the exact same neighbourhood; that is, in the same (ED).[8]

[7] Several econometric specifications have been developed that seek to explicitly impose the restrictions implied by economic theory. Indeed Cheshire and Sheppard (1998) and Palmquist and Isangkura (1999) employ one such specification, the almost ideal demand system, in their hedonic pricing studies.

[8] An anonymous reviewer has raised the issue of whether the values of nearby households qualify as valid IVs. In particular, they note that in a world of spatial correlation, individuals with similar (and unobserved) characteristics may group themselves in the same area. In that case, our IVs may exhibit correlation with the error term in the same way as the original variables. In our defence, the households we chose as being "nearby" were not allowed to live in the same ED. Since road and rail noise in the urban environment are highly localised phenomena, it is unlikely that the noise environment in one ED will reflect that in a neighbouring or nearby ED. As such, if the omitted characteristics that are the source of our problem, are those relating to attitudes to the noise environment, then we have no reason to believe that these characteristics will be correlated across nearby EDs.

Amemiya (1979) describes an IV procedure for censored data that generalises the Tobit model and returns consistent estimates of the parameters of the linear demand system. The procedure is known as Amemiya Generalised Least Squares (AGLS). Newey (1987) shows that AGLS returns parameter estimates that are more efficient than other possible estimators.

7 Estimation of demand functions for peace and quiet

In estimating the demand relationships, there are a number of data issues with which we have had to contend. As noted previously, the PLM+SS specification of the HPF indicates that for a handful of segments, the price of peace and quiet is negative. 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? One response to this problem is to treat the negative prices from the PLM + SS model as resulting from sampling error where the 'true' implicit prices for peace and quiet are in the region of zero. The solution then, is to set aberrant prices to a value of zero in the regression analysis. A second response is to impose negativity on the first stage regressions using the constrained PLM + SS specification discussed previously. In what follows we compare the outcomes of these two approaches.

Second, we must decide upon a discount rate, τ 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. Comparison of the sale price and rental price of matched properties provides an estimate of the required discount rate. This data from 2003 suggests 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%.[9]

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.

Tables 6 and 7 record estimates of the parameters of the pseudo-Marshallian demand functions for peace and quiet from road and rail traffic, respectively. The functions have been estimated using implicit prices calculated from both the unconstrained PLM + SS specification (with negative prices set to zero) and the constrained PLM + SS specification.[10] Also, for each set of price estimates, the demand function parameters have been estimated using both the Tobit model and the AGLS model which corrects for the endogeneity of prices.

Consider first the issue of price endogeneity. If endogeneity is not a problem with this data then, assuming we have defined valid instruments, we would expect the Tobit and AGLS models to return very similar parameter estimates. Observe from Tables 6 and 7, however, that for both sets of price data, the parameter estimates from these two models differ significantly in magnitude. We calculate a test statistic suggested by Smith and Blundell (1986) to assess

[9] 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.

[10] In estimating standard errors for the parameters of the second stage regressions we do not correct for the fact that the implicit prices are themselves estimated with some level of imprecision. Accordingly, the p-values reported in Tables 6 and 7 overstate the significance of the parameter estimates and the width of the confidence intervals for the welfare values that follow in Tables 8 and 9 will be biased downwards. Notice, however, that not accounting for this added imprecision does not bias either the parameter estimates or the mean welfare values derived from them.

Table 6 Parameters of the pseudo demand functions for road noise (p-values in brackets)

	PLM + SS		PLM + SS constrained	
	Tobit	AGLS	Tobit	AGLS
Road noise price	−0.179 (0.000)	−0.419 (0.000)	−0.192 (0.000)	−0.420 (0.000)
Rail noise price	−0.007 (0.492)	0.104 (0.110)	−0.024 (0.008)	−0.012 (0.862)
Log pseudo expenditure	3.318 (0.000)	7.649 (0.006)	4.982 (0.000)	11.485 (0.000)
Ethnicity factor	0.611 (0.001)	1.438 (0.000)	0.820 (0.000)	1.652 (0.000)
Age factor	−0.131 (0.538)	−1.506 (0.000)	−0.076 (0.718)	−0.910 (0.057)
Family factor	1.272 (0.000)	1.084 (0.000)	1.273 (0.000)	1.160 (0.000)
Constant	−16.709 (0.000)	−52.735 (0.011)	−19.443 (0.000)	−80.644 (0.011)
Smith–Blundell test of exogeneity:	12.357 (0.000)		9.481 (0.000)	
Uncensored obs:	2,723			
Censored obs:	7,918			

Table 7 Parameters of the pseudo demand functions for road noise (p-values in brackets)

	PLM + SS		PLM + SS constrained	
	Tobit	AGLS	Tobit	AGLS
Rail noise price	−0.043 (0.047)	−0.446 (0.001)	−0.036 (0.042)	−0.269 (0.039)
Road noise price	−0.224 (0.000)	−0.399 (0.041)	−0.214 (0.001)	−0.385 (0.050)
Log pseudo expenditure	8.139 (0.000)	20.297 (0.000)	9.020 (0.000)	18.005 (0.001)
Ethnicity factor	0.719 (0.053)	0.729 (0.126)	0.941 (0.013)	1.186 (0.031)
Age factor	1.623 (0.000)	2.978 (0.000)	1.652 (0.000)	2.553 (0.005)
Family factor	0.663 (0.094)	0.535 (0.276)	0.682 (0.085)	0.167 (0.733)
Constant	−36.410 (0.000)	−124.726 (0.001)	−43.515 (0.000)	−108.878 (0.008)
Smith–Blundell test of exogeneity:	3.877 (0.009)		1.934 (0.122)	
Uncensored obs:	379			
Censored obs:	10,262			

Table 8 Welfare estimates for changes in road noise exposure at means of covariate data (95% confidence interval range in brackets)

Noise change		Welfare change (£ per annum in 1997 prices)			
High	Low	PLM + SS		PLM + SS constrained	
		Tobit	AGLS	Tobit	AGLS
56	55	61.19 (49.34–81.11)	31.49 (24.84–52.52)	57.77 (47.74–76.26)	32.32 (26.58–46.40)
61	60	89.19 (71.09–120.00)	43.42 (33.14–75.97)	83.75 (68.20–112.45)	44.22 (35.31–66.69)
66	65	117.18 (92.86–159.61)	55.35 (41.49–99.57)	109.72 (88.62–148.69)	56.12 (44.03–87.04)
71	70	145.17 (114.58–198.96)	67.28 (49.83–123.57)	135.69 (108.97–184.91)	68.02 (52.77–107.31)
76	75	173.17 (136.55–238.47)	79.21 (58.17–147.58)	161.67 (129.27–221.52)	79.93 (61.52–127.57)
81	80	201.16 (158.37–277.98)	91.15 (66.52–171.58)	187.64 (149.65–257.78)	91.83 (70.32–147.84)

Table 9 Welfare estimates for changes in rail noise exposure at means of covariate data (95% confidence interval range in brackets)

Noise change		Welfare change (£ per annum in 1997 prices)			
High	Low	PLM + SS		PLM + SS constrained	
		Tobit	AGLS	Tobit	AGLS
56	55	645.27 (317.52–2,435)	83.61 (43.21–461.80)	789.02 (419.94–3,231)	127.11 (77.61–469.36)
61	60	761.72 (372.86–2,866)	94.82 (49.30–519.77)	928.99 (490.16–3,831)	145.67 (86.74–557.37)
66	65	878.16 (426.51–3,309)	106.02 (55.39–580.05)	1,068.97 (560.84–4,431)	164.24 (95.72–645.39)
71	70	994.61 (480.16–3,752)	117.23 (61.47–654.04)	1,208.94 (632.99–5,032)	182.80 (104.88–733.41)
76	75	1,134.35 (533.81–4,196)	128.44 (67.47–728.03)	1,348.91 (705.16–5,632)	201.36 (114.16–821.42)
81	80	1,227.50 (588.40–4,639)	139.65 (72.72–802.03)	1,488.88 (777.18–6,232)	219.93 (123.44–909.44)

whether accounting for the endogeneity of prices has a statistically significant impact on the parameter estimates. The test statistic is highly significant for both sets of data in the case of the road noise demand functions. The statistic is only significant for the unconstrained price data in the case of the rail noise demand functions. Taken as a whole, however, it seems that we are well-advised to control for endogeneity in our estimation of the demand functions.

In contrast, when comparing the AGLS estimates for the unconstrained and constrained price data we see only relatively modest differences in the parameters. A standard Hausman test of differences in parameters confirms that differences are not statistically significant for either the road or rail noise demand functions. Accordingly we focus our discussion on the parameter estimates from the AGLS model using the unconstrained price data.

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

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. For rail traffic noise, however, the parameter estimate is 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. We have no particular theory to support these conflicting results.

Finally, we adopt the framework suggested by Weesie (1999) in order to test whether the two demand relationships can be considered as separate estimates of the same general demand function for peace and quiet independent of source. We find that we can reject this hypothesis with a high level of confidence; preferences for peace and quiet appear to differ according to the source of the noise.

8 Welfare estimates

Estimates of the benefits of an external change in the provision of peace and quiet can be calculated as areas under the pseudo-Marshallian demand curves. For road and rail traffic, respectively, Tables 8 and 9 report welfare values for selected 1 dB changes in noise levels calculated at the mean values of the covariate data for each of the models in Tables 6 and 7. The welfare estimates are values per annum reported in 1997 prices.

Notice that in all cases, the preferred AGLS model returns more conservative estimates of the welfare changes. Clearly, our observation that accounting for price endogeneity has significant impacts on the estimated parameters of the pseudo-Marshallian demand functions carries over to welfare estimates based on those functions. In contrast, the welfare estimates derived from the constrained and unconstrained price data are remarkably similar. Indeed, in all cases it is impossible to reject the hypothesis that these are estimates of the same value. Accordingly, we again focus our discussion on the welfare values derived from AGLS model applied to the unconstrained data.

The mean values for road noise range from £31.49 per annum for a 1 dB reduction from a 56 dB baseline to £91.15 per annum for the same change from an 81 dB baseline. The equivalent values for rail noise are higher, ranging from £83.61 to £139.65 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. For example, in summarising a location's noise exposure as the continuous noise level with equivalent energy content, one might conjecture that the dB_{LEQ} measure provides a relatively poor scale by which to approximate the differences in disbenefit resulting from the extreme spikes of rail noise as compared to the relatively more constant pattern of road noise.

Furthermore, Tables 8 and 9 report 95% confidence intervals for each of the reported welfare estimates. These are calculated using a non-parametric bootstrap procedure. It is apparent that the 95% confidence intervals for the rail welfare estimates are considerably wider than those for the road welfare estimates, reflecting the substantially fewer observations of properties exposed to rail noise above the urban background level upon which those estimates are based.

9 Concluding remarks

While several previous studies have attempted to identify implicit prices for peace and quiet, those estimates reflect only the specific conditions of supply and demand that exist in the particular property markets studied. Such estimates 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 households' underlying preferences, they might be preferred as objects for transferring benefits across locations. As far as the authors are aware, the values presented in Tables 8 and 9 are the first estimates of welfare values for peace and quiet derived from hedonic property market data in a theoretically consistent manner.

In proffering these values 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 as opposed to property-rental 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 perhaps 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-Marshallian 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-Marshallian 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 say how closely the welfare estimates presented in Table 4 parallel 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. Indeed, in tackling these various problems our research brings together a host of frontier techniques, some of which have not previously been applied in economics research. 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.

Acknowledgements Funding for this research was provided by the Department for Environment Trade and the Regions (now Department for Transport) and by the PEDM which is funded by the UK Economic and Social Research Council (ESRC) Award M535255117. The quality of the output was greatly enhanced through the kind cooperation of the Land Registry and the Valuation Office Agency.

References

Abraham JM, Goetzmann WN, Wachter SM (1994) Homogenous groupings of metropolitan housing markets. J Housing Econ 3(3): 186–206
Allen MT, Springer TM, Waller NG (1995) Implicit pricing across residential rental markets. J Real Estate Finance Econ 11: 137–151
Amemiya T (1979) The estimation of a simultaneous-equation Tobit model. Int Econ Rev 20(1): 169–181
Anglin PM, Gencay R (1996) Semiparametric estimation of a hedonic price function. J Appl Economet 11: 633–648
Bajic V (1993) Automobiles and implicit markets: An estimate of a structural demand model for automobile characteristics. Appl Econ 25(4): 541–551
Barde J-P (2007) Harnessing the political economy of environmental policy: David Pearce's contribution to OECD, Environ Resour Econ 37(1): 33–42 (this issue)
Bartik TJ (1988) Measuring the benefits of amenity improvements. Land Econ 64(2): 172–183
Bateman IJ, Day BH, Lake I, Lovett AA (2001) The effect of road traffic on residential property value: a literature review and hedonic pricing study. Scottish Executive and The Stationary Office, Edinburgh
Bateman IJ, Day BH, Lake I (2004) The valuation of transport-related noise in Birmingham. Technical Report to the Department for Transport. University of East Anglia, Norwich
Bell K, Bockstael NE (2000) Applying the generalised method of moments approach to spatial problems involving micro-level data. Rev Econ Stat 82(1): 72–82
Blomquist NS (1989) Comparative statics for utility maximization models with nonlinear budget constraints. Int Econ Rev 30: 275–296
Bourassa SC, Hamelink F, Hoesli M, MacGregor BD (1999) Defining housing submarkets. J Housing Econ 8: 160–183
Boyle KJ, Poor PJ, Taylor LO (1999) Estimating the demand for protecting freshwater lakes from eutrophication. Am J Agric Econ 81(5): 1118–1122
Brown JN, Rosen HS (1982) On the estimation of structural hedonic price models. Econometrica 50(3): 765–768
Carson RT, Groves T (2007) Incentive and informational properties of preference questions, Environ Resour Econ 37(1):181–210 (this issue)
Cheshire P, Sheppard S (1998) Estimating the demand for housing, land and neighbourhood characteristics. Oxford Bull Econ Stat 60(3): 357–382
Cliff A, Ord JK (1972) Testing for spatial autocorrelation among regression residuals. Geogr Anal 4: 267–284
Day BH (2005) Hedonic analysis of property markets: theory and applications to UK cities. PhD Thesis, Department of Economics, University College London.
Department for Transport (DfT) (2006) Transport analysis guidance. TAG Unit 3.3.2: Noise, London: DfT; http://www.webtag.org.uk/webdocuments/3_Expert/3_Environment_Objective/3.3.2.htm
Department for Environment, Food & Rural Affairs (DEFRA) (2000) A report on the production of noise maps of the City of Birmingham. DEFRA, London
Edlefsen LE (1981) The comparative statics of hedonic price functions and other nonlinear constraints. Econometrica 49: 1501–1520
Ekeland I, Heckman JJ, Nesheim L (2002) Identifying hedonic models. Am Econ Rev 92(2): 304–309
Ekeland I, Heckman JJ, Nesheim L (2004) Identification and estimation of hedonic models. J Polit Econ 112(1): 60–109
European Commission (EC) (1996) Future noise policy – European Commission Green Paper. Report COM(96) 540 final. European Commission, Brussels
European Commission (EC) (2002) Directive 2002/49/EC for the assessment and management of environmental noise. Official Journal of the European Communities L189/12, European Commission, Brussels

Fraley C, Raftery AE (1998) How many clusters? Which clustering method? Answers via model-based cluster analysis. Computer J 41(8): 578–588

Geweke J (1986) Exact inference in the inequality constrained normal linear regression model. J Appl Econ 1: 127–141

Gibbons S, Machin S (2003) Valuing English primary schools. J Urban Econ 53(2): 197–219

Goetzmann WN, Wachter SM (1995) Clustering methods for real estate portfolios. Real Estate Econ 23(3): 271–310

Haab TC, McConnell KE (2003) Valuing environmental and natural resources: the econometrics of nonmarket valuation. Edward Elgar, Northampton, MA

Hanemann WM (1991) Willingness to pay and willingness to accept: how much can they differ?. Am Econ Rev 81(3): 635–647

Heckman J, Matzkin R, Nesheim L (2003) Simulation and estimation of hedonic models. CENMAP working paper, CWP10/03. Institute of Fiscal Studies, UCL, London

Kelejian HH, Prucha IR (1999) A generalised moments estimator for the autoregressive parameter in a spatial model. Int Econ Rev 40(2): 509–533

Kelejian HH, Robinson DP (1992) Spatial autocorrelation: a new computationally simple test with an application to per capita county police expenditures. Regional Sci Urban Econ 22: 317–331

McConnell KE, Phipps TT (1987) Identification of preference parameters in hedonic models: consumer demands with nonlinear budgets. J Urban Econ 22: 35–52

Murray MP (1982) Mythical demands and supplies for proper estimation of Rosen's hedonic price model. J Urban Econ 14: 327–337

Nellthorp J, Bristow AL, Day BH (forthcoming) Introducing willingness-to-pay for noise changes into transport appraisal – an application of benefit transfer. Transport Rev

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 WK (1987) Efficient estimation of limited dependent variable models with endogenous explanatory variables. J Econometr 36(3): 231–250

Odgaard T, Kelly C, 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 RB (1988) Welfare measurement for environmental improvements using the hedonic model: the case of nonparametric marginal prices. J Environ Econ Manage 15: 297–312

Palmquist RB, Isangkura A (1999) Valuing air quality with hedonic and discrete choice models. Am J Agric Econ 81(5): 1128–1133

Posse C (2001) Hierarchical model-based clustering for large datasets. J Comput Graphical Stat 10(3): 464–486

Randall A, Stoll JR (1980) Consumer's surplus in commodity space. Am Econ Rev 70(3): 449–457

Robinson PM (1988) Root-N-consistent semiparametric regression. Econometrica 56: 931–954

Rosen S (1974) Hedonic prices and implicit markets: production differentiation in pure competition. J Polit Econ 82: 34–55

Smith RJ, Blundell RW (1986) An exogeneity test for a simultaneous equation Tobit model with an application to labor supply. Econometrica 54(3): 679–685

Turner K (2007) Limits to CBA in UK and European environmental policy: retrospect & future prospects. Environ Resour Econ 37(1): 253–269 (this issue)

Watkins CA (2001) The definition and identification of housing submarkets. Environ Plan A 33: 2235–2253

Weesie J (1999) Seemingly unrelated estimation: an application of the cluster-adjusted sandwich estimator. Stata Technical Bull 52: 34–47

Willig RD (1976) Consumer's surplus without apology. Am Econ Rev 66(4): 589–597

Zabel JE, Kiel KA (2000) Estimating the demand for air quality in four U.S. cities. Land Econ 76(2): 174–194