department of economics
"S. Cognetti de Martiis"

Via Po, 53 – 10124 Torino (Italy)
Tel. (+39) 011 6704917 - Fax (+39) 011 6703895
URL: http//www.de.unito.it

WORKING PAPER SERIES

GIS and Geographically Weighted Regression in stated preferences analysis of the externalities produced by linear infrastructures

Sergio Giaccaria e Vito Frontuto

Dipartimento di Economia "S. Cognetti de Martiis"

Working paper No. 10/2007

Università di Torino

GIS and Geographically Weighted Regression in stated preferences analysis of the externalities produced by linear infrastructures

Sergio Giaccaria
Vito Frontuto

Dipartimento di Economia
Università di Torino
Via Po, 53
I-10124 Torino (Italy)
Corresponding author: <sergio.giaccaria@unito.it>

February 1, 2007

Abstract

The paper uses Contingent Valuation to investigate the externalities from linear infrastructures, with a particular concern for their dependence on characteristics of the local context within which they are perceived. We employ Geographical Information Systems and a spatial econometric technique, the Geographic Weighted Regression, integrated in a dichotomous choice CV in order to improve both the sampling design and the econometric analysis of a CV survey. These tools are helpful when local factors with an important spatial variability may have a crucial explanatory role in the structure of individual preferences. The Geographic Weighted Regression is introduced, beside GIS, as a way to enhance the flexibility of a stated preference analysis, by fitting local changes and highlighting spatial non-stationarity in the relationships between estimated WTP and explanatory variables. This local approach is compared with a standard double bounded contingent valuation through an empirical study about high voltage transmission lines. The GWR methodology has not been applied before in environmental economics. The paper shows its significance in testing the consistency of the standard approach by monitoring the spatial patterns in the distribution of the WTP and the spatial stability of the parameters estimated in order to compute the conditional WTPs.

Acknowledgements: We are grateful to Riccardo Scarpa, Edward Morey and Anna Alberini for their suggestions, discussions and contributions.

Keywords:

Stated preferences; Contingent valuation; GIS; Geographic Weighted Regression; externalities; linear infrastructures; spatial analysis.

JEL Classification: *C21, D62, H5, O13, O22, Q51.*

Introduction

In stated preference studies, a spatial criterion is generally used to delimit the scenario of environmental change. Public choices about land use (both in rural and urban contexts) are a typical framework for the analysis of stated preferences within which spatially defined attributes are used to improve the description of the hypothetical scenario (*inter alia* Riganti, Alberini, Longo *et al.* 2005; Johnston, Swallow, Bauer, 2002) . In the study of the environmental impact from linear infrastructure networks, such as railways, highways, and power lines, we suggest that the implementation of a spatial perspective may help improving stated preferences studies. A spatial perspective proves to be especially relevant when the analysis deals with preferences of residential households who are familiar with the infrastructure, whose location can determine their perception of the impacts.

We use well-experimented distance-based approaches (Bateman, 2002) to include among the explanatory variables, the proximity of respondents to the infrastructure being valued, the possible interactions with other local infrastructures, and quantitative indicators of environmental quality. The paper aims to verify whether perceived externalities are highly sensitive to these features, which we define *local context variables* (LCV). We present two alternatives for the econometric modelling of this contextual analysis. The first, which we call *global approach,* is the well-experimented double-bounded model for contingent experiments based on dichotomous choices (Hanemann, Loomis, Kanninen, 1991). It gives us results that are global in the sense that the relationships between WTP and explanatory variables are fixed for the whole population from which we draw inference. The second approach, which we call *local*, is based on a geostatistical technique – the Geographically Weighted Regression (GWR) – employed to verify the stationarity of relationships across space (Fotheringham *et al.*, 2002). GWR has been employed before in the study of real estate markets, in ecology to examine spatial stationarity between species richness and environmental drivers (Foody, 2004; Foody, 2005), in epidemiology (Nakaya *et al*, 2005), and other case studies developed mainly by the research group that first introduced this methodology (Fotheringham *et al*, 2002). To our knowledge, this is the first application of this approach in environmental economics.

In the first section we discuss the omitted variable bias which can be controlled for by considering local context variables. In the second section we illustrate the definition of local context of linear infrastructures in a geographical sense. In the rest of the paper we control for context dependence of WTPs in the case of the valuation of externalities caused by High Voltage Transmission Lines, as a relevant example of linear infrastructure. Section 3 illustrates the structure of the case study. Section 4 analyses the global case with a standard econometric approach. Section 5 applies the GWR approach to the same data, thus shifting the focus from summarizing parameters at a global level to modelling local heterogeneity geographically (Fotheringham *et al*; 2001). We compare the goodness-of-fit and the relative predictive performance of global and local models, using GIS to map the numerical outputs from GWR: WTPs, parameters and residuals. We also test for local variations of WTP and use GWR to identify sub-regional areas within which the impacts of the infrastructure can be considered as spatially stationary. Section 6 concludes.

1. Endogeneity and context dependence

Environmental externalities from linear infrastructures impact on wide areas generally characterized by a significant degree of spatial heterogeneity. Most methodologies that analyze individual preferences in order to infer economic values of these externalities are set in the theoretical framework of methodological individualism and inferential methods. From an operative point of view, these valuations are potentially meaningful for planners and decision-makers as a support for choosing the optimal location of the infrastructures. The results of these valuations, however, are sensitive to the scale employed in the design of the study. In the framework of inferential quantitative analysis, the trade-off between global and local approaches needs to receive careful consideration, as the problem of shifting from global to local involves important methodological implications. For example, valuing the externalities of an infrastructure network within a regional area, in to account for patterns of spatial heterogeneity we could design a geographically stratified sample so as to draw inference from an individual level to an aggregated level. Obtaining aggregated estimates starting from individual preferences is the standard case of *statistical inference* (achieved for instance with hedonic pricing or contingent valuation methods). On the contrary, *ecological inference* is defined as the prediction of individual behaviours from grouped data (for example using benefit transfer).

In the case of linear infrastructures we suggest that the externalities are perceived from individuals in a way that is dependent on local context features. Spatial patterns have been pointed out as relevant explanatory contextual factors, in rural land use choice (Bockstael, 1996) and in the valuation of rural amenities (Johnston, Swallow and Bauer, 2002). It appears reasonable that the same can occur for disamenities such as linear infrastructures, often located in rural and marginalized areas. To account for the context-dependence in a spatial perspective, Anselin proposes the concept of "spatial externality" (Anselin, 2002). Both in spatial econometrics and quantitative geography, considerable empirical and methodological efforts are concerned with the exploration of possible sources of bias in statistical or ecological inference.

Two crucial topics, in this field, are *endogeneity* and *spatial correlation*. The term endogeneity, here, refers mainly to biases from omitted variables (Guevara, Ben-Akiva, 2005). Factors that have an explanatory role in a regression model may be omitted (as it could be in the case of local context variables). In common regression models, these (omitted) independent variables are considered as part of an unobservable component, modelled as residuals. The omitted variables are correlated with the dependent variable (for example the probability of saying "no" to a bid of a certain amount in a CV survey), and consequently also with other explanatory variables. This correlation generates some degree of association between explanatory variables and residuals, and can involve a bias, as it violates the assumption of independence of residuals of most common econometric regression models. Even if the econometric model does not suffer from this restriction, omitting local context variables could anyway result in models that are not well-specified.

In order to check whether contextual factors can be considered as a source of omitted variables endogeneity in a contingent experiment, we compare different models (*global* and *local*), testing the explanatory performance of local context variables (LCV). Spatial definition of the context, in the study, is achieved with variables obtained through GIS techniques. A detailed review of the literature about the role of GIS in applied environmental economic studies is offered by Bateman (2002), who deals both with theory, econometrics and implementation. Note that in this first phase of collecting data, GIS is mainly a technical support employed to manage different sources of alphanumerical data and maps, rather than a methodology for the econometric analysis.

Aggregating and processing data for global and local analysis
The scheme in Figure 1 draws the operative structure that we followed. Arrows are flows of data merged by GIS or treated through statistical inference. From the upper right part of the scheme, we see regional GIS data as a source of maps and geo-referenced information. The sample design (on the left of the scheme) aims to define how many observations are to be collected in each geographical strata. Identifying respondents leads to the phase of telephone interviewing, while locations of respondents are specified with GIS. Geo-referenced households data have inherited LCV data, their location ruled as a link to merge LCV to the survey data, and as a way to calculate distances between the household, the HVTL route, other linear infrastructures. The complete dataset is analyzed through the global (left side of the scheme) and the local GWR approach (on the left side).

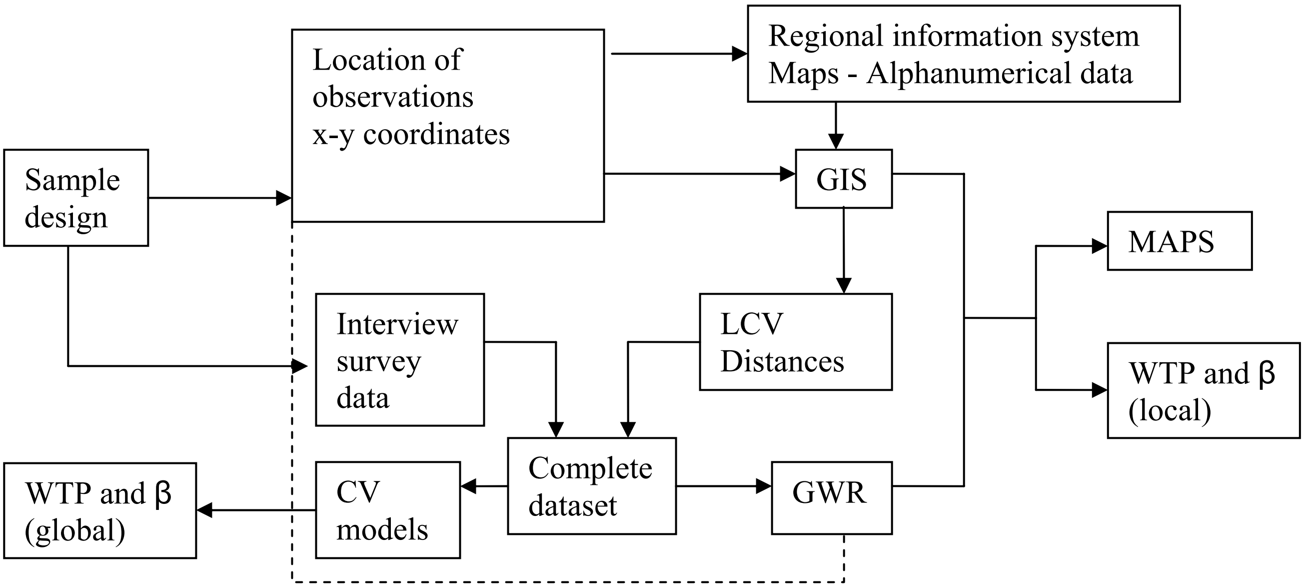


Figure 1 – Scheme of the structure of the study

2. Spatial boundaries for linear infrastructure impacts: defining the geographical context

The spatial context within which the externalities of linear infrastructures are perceived is conventionally identified by means of "corridors". Socioeconomic and ecosystemic impacts that can be intended as "external effects" are more intense in these areas. The literature offers examples of methodologies for the assessment of ecosystem impacts of linear infrastructures based on conceptual framework of corridors. Geneletti (2002, 2004), for instance, employs GIS tools in a analysis of the level of spatial fragmentation of habitats and biodiversity loss produced by roads. An automated estimation, performed with GIS, of the extension and of the perimeter of areas that represent typological units of ecosystems supplies numerical values for biophysical indicators, useful for instance in Environmental Impact Assessment. GIS is also widely applied in the field of Strategic Environmental Assessment (Directive 2001/42/CE). In this case, the aim of GIS is the visualization of environmental impacts through maps, where several typologies of impact are present and visual supports make it possible to identify the areas with low impacts.

3. An application to linear infrastructures: High Voltage Transmission Lines

Our application is focused on a contingent valuation survey implemented on a geographically stratified sample of households, living in the High Voltage Transmission Lines (HVTLs) corridors.[1] The valuation methods most frequently employed in the monetary estimate of the externalities produced by the presence of HVTLs are those based on the real estate market. In particular, the studies based on the hedonic price method observe the impact of the proximity to the HVTL on the value of properties. It is frequent to observe the use of local context information as explanatory variables in econometric models (*inter alia* DeRosiers, 2002; Sims and Dent, 2005). Other applications are based on a Stated Preferences approach aimed at providing a holistic valuation of total economic value (TEV). Contingent Valuation (CV) studies on HVTLs (Atkinson *et al.*, 2004; Tempesta and Marazzi, 2005, Rosato *et al.*, 2004) tend to be more inclined towards the estimation of cost and benefits related to visual encumbrance and landscape quality and emphasize non-use values of landscape.

In this paper, CV is used to estimate the damage impinging on all of the different components of TEV. To do so, it is crucial to identify the relevant externalities concerned with the presence of HVTLs. The main externalities deriving from HVTLs are referable to the perception of risk for human health, visual encumbrance, and other external effects for landowners.

Human health risk. This is the externality characterised by a higher level of uncertainty, both in terms of scientific knowledge, and in terms of social perception. The perception of health risk and the estimation of health damages related with the proximity to power lines are complicated by the fact that electric magnetic fields are present virtually everywhere. The uncertainty arises from the diffusion and cumulative impact of electronic equipments that act as emission sources (grounding systems in homes, transformers, switching gear in offices and homes, computers and household appliances). The outcomes of scientific research that examines the impact from exposure to electro-magnetic emissions produced by power lines are not unanimous (Extremely Low Frequencies are those referred to HVTLs). Gregory and von Winterfeldt (1996) have collected the main studies that investigate the existence of a link between the exposure to ELF and the set in of degenerative pathologies. They point out the existence of studies that recognize “a significant association between indirect measures of exposure and cancer” and “other studies that found no statistical evidence of such effects” (p.201).

Visual encumbrance and landscape quality. Linear infrastructures have, by nature, a strong impact on landscape quality. The technical features of power lines, in particular the presence of towers, determine a strong impact in term of aesthetic quality of the lands crossed.

Land use: the presence of the towers may impose serious constraints on the use of the occupied area. In general, this kind of impact derives from a limitation of property rights.

[1] The survey is part of a research project on the valuation of externalities from different types of infrastructures conducted in the Department of Economics of the University of Turin and financed by the Piedmont Region (Nucleo di Valutazione degli Investimenti Pubblici), whose support is gratefully acknowledged. The research staff of the project includes Ugo Colombino, Silvana Dalmazzone, Vito Frontuto and Sergio Giaccaria

3.1 Experimental design of the CV study

The study area is the Piedmont Region in north-western Italy (Figure 2). Considering only the transmission lines (132V, 220V, 380V), the power lines network crosses 786 municipalities and involves 2.613.904 inhabitants living within the corridor of HVTLs influence.

The hypothetical scenario proposed in the survey uses a referendum format. We hypothesize a regional program for the modernization and rationalization of the power lines network that involves the removal of some portions of infrastructure. Carrying out this program requires a contribution by citizens, presented in the form of payment of a *una-tantum* tax. The hypothetical market assumes the form of a local political market:

"…Suppose that your municipality, in order to decide whether to demolish a 5 Km portion of HVTL, is asking the opinion of citizens through a referendum. If you vote NO, the portion of power line will not be removed. If you vote YES, the line will be removed, but all the citizens will have to contribute the payment of a una-tantum tax. If the amount of the tax were X Euro, would you vote Yes or No?"

The 5 Km value chosen as the length of line to be removed was selected because this is the median length of power lines in the Piedmont municipalities.

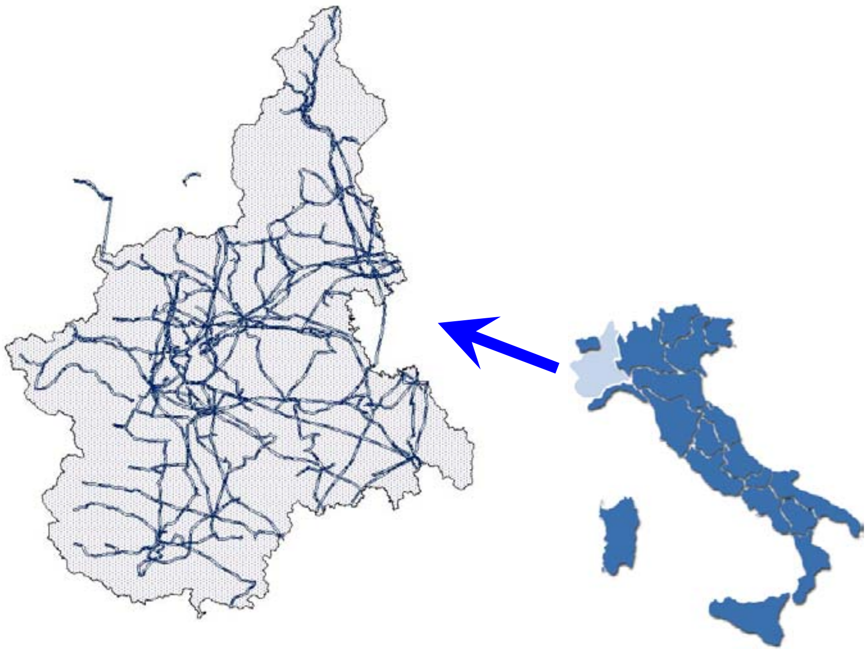


Figure 2. The HVTLs network in Piedmont

3.2 Sampling frame (geographical stratification)

The reference population includes the residents in proximity (i.e. within the corridor) of the actual lines of all of the regional network. The choice of the corridor width plays an important role in the experimental design. A review of empirical studies shows chosen widths of the HVTL corridor varying between 244 meters and 5 Km (Table 1). We considered that the study area includes sub-regions where visual impacts of lines vary sensibly due to the shape and size of the towers, the orographic trim, climatic conditions and the presence of vegetation. In our case, considering the high variability of the land crossed by HVTLs, we consider a corridor of 1200 meters, so that the maximum distance between households and power lines is 600 meters[2].

Table 1. Corridor widths in the literature

Studies	**Valuation method**	**Corridor width**
Atkinson *et al. (*2004)	Choice experiments	5 km
Colwell (1990)	Hedonic price	244 m.
Bond (1995)	Hedonic price	244 m.
Haider *et al.* (2001)	Hedonic price	300 m.
Des Rosiers (2002)	Hedonic price	488 m.
Rosato *et al.* (2004)	Contingent Valuation	1200 m.

We overlaid the 1200 meters corridor around the HVTLs on the regional cartography (the set of Technical Regional maps, at a scale 1:10000). This procedure provided a sample stratification on the basis of geographic strata which allowed us to define ten macro-areas: this zoning in sub-regional areas has been obtained from previous geographical analysis of the Piedmont region (ITATEN, 1996). These macro-areas are considered as strata that have an appreciable degree of internal homogeneity from the point of view of landscape typologies, vegetation, orography.

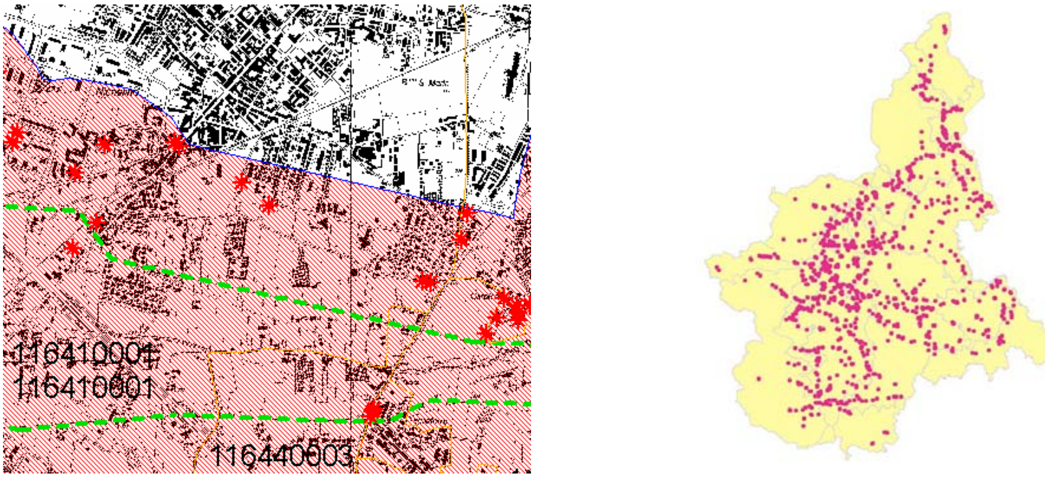


Figure 3. Locations of respondents

The number of interviews was divided between the macro-areas on the basis of the number of residents within the HVTL corridor. In each macro-area the number of interviews was calculated using a density index of power lines[3].

[2] In the application of CV to the estimation of costs and benefits from power lines burial by Rosato *et al.* (2004), a corridor of 1200 meters was defined as: “the corridor interested by direct impacts of visual encumbrance and health risk” (our translation).

[3] The density index of power lines was defined as the adimensional ratio between the corridor area and the municipal area.

3.3 Pre-test and differentiated levels of damage

An exploratory pre-test was used to define bid vectors for the elicitation questions. *Open-ended* questions have been employed in the pre-test to obtain some preliminary indication on the maximum WTP values. Most respondents living in close proximity to the lines indicated WTP values near to 10-20.000 Euro, signalling a negative impact of the infrastructure on real estate property values. All other respondents, who generally referred generic forms of nuisance linked to visual encumbrance, expressed much lower WTP. The overall damage could be thought of as modulated in function of a progressive overlapping of impacts. We therefore predisposed three different vectors for bid values to account for differentiated levels of damage.

In the first part of the survey we inserted a few questions aimed at providing information on the type of individually perceived damage, in order to assign the appropriate bid vector to the respondent. Typologies and intensity of the perceived impacts, proximity to the lines, and depreciation of the house were used as assignment criteria. In this way the respondent him/herself selected the bid vector. Vector 1 is associated with a diffused and ordinary condition. The others two vectors are used in a limited number of observations to identify an intermediate (vector 2) or heavy damage (vector 3).

4. Data analysis: global models

The elicitation question has a dichotomous double-bounded format (Hanemann, Loomis and Kanninen, 1991). In the dichotomous formats the answers at the first and the second elicitation question allowed to define closed intervals within which the unknown WTPs lies.
We consider, in the evaluation of HVTLs externalities, that the infrastructure has an effect in the determination of individual perception of own utility. We assume that each respondent, resident in the HVTLs proximity, has a state variable that assume value *0* when he is exposed at the effect of power line and value *1* when the infrastructure is not present (removed). So utility levels can be expressed as:

$U_0(0,y_0,s)$ in presence of infrastructure
$U_1(1,y_1,s)$ without infrastructure

with y the income and s the vector of observable characteristics that influence the willingness to pay.

The respondent will choose to pay the amount proposed (BID) to remove the power line if:

$U_1>U_0$ that is $v_1(1, y_1- WTP, s) + \varepsilon_1 > v_0(0, y_0,s) + \varepsilon_0$ [1]

Considering the single-bounded model, one dichotomous variable is required to estimate WTP. From expression [1] a relationship between utilities and probabilities to obtain positive responses is derived (Hanemann, 1984):

$$\Pr\{Yes\} = \Pr \{ v_1(1, y_1- WTP, s) + \varepsilon_1 > v_0(0, y_0,s) + \varepsilon_0\} \quad [2]$$
$$\Pr\{Yes\} = \Pr \{\varepsilon_1-\varepsilon_0 < v_1(1, y_1- WTP, s) - v_0(0, y_0,s)\} \quad [3]$$

The [3] can be rewritten as:

$\Pr\{Yes\} = \Pr\{\eta < \Delta v\}$ [4]

with η the difference ε_1-ε_0 (the stochastic components) and Δv the difference between the observable components of utility [$\Delta v = v_1(1, y_1 - WTP, s) - v_0(0, y_0, s)$]

The probability to obtain a positive response can be read as the probability that the stochastic component (η) is lower or equal to the deterministic component (Δv), that is the cumulative distribution function of the stochastic variable.

$\Pr\{Yes\} = \Pr\{\eta < \Delta v\} = F_\eta(\Delta v)$ [5]

We chose to use the logistic distribution for $F\eta$ (logit and probit distributions are pointed out as suitable to estimate double-bounded models (McFadden D., 1994)).
The [5] can be rewritten as:

$\Pr\{Yes\} = 1/(1+e^{-\Delta v})$ [6]

To express the deterministic component we adopt the linear form:

$v_0 = \alpha_0 + \beta_y$ [7]
and
$v_1 = \alpha_1 + \beta\,(y - WTP)$ [8]

We are interested to estimate the amount of WTP that corresponds with the median value of the distribution (in this case the probability of response positive equal to 0,5). On the basis of [5],

$F\eta\,(\Delta v) = 0{,}5$ [9]

If we use a logistic distribution for η, $F\eta$ assume value 0,5 when η is equal to 0.

$F\eta\,(0) = 0{,}5$ so $\Delta v = 0$

We can calculate the WTP using the [10] and the [11]:

$\alpha_0 + \beta\, y = \alpha 1 + \beta\,(y - WTP)$
$\alpha_0 - \alpha 1 = -\beta\, WTP$

$WTP = -(\alpha_0 - \alpha 1)/\beta$ [10]

To solve the case of double-bounded format we used the Hanemann, Loomis and Kanninen approach (Hanemann, Loomis and Kanninen 1991). As we have two dichotomous responses, four combinations of answer can be verified; to each of these answer is associated a probability:

$\Pr\{Yes,Yes\} = \Pr\{WTP \geq BID^H \geq BID\}$ [11]

$\Pr\{Yes,No\} = \Pr\{BID \leq WTP \leq BID^H\}$ [12]

$$\Pr\{\text{No,Yes}\} = \Pr\{BID^L \leq WTP \geq BID\} \quad [13]$$

$$\Pr\{\text{No,No}\} = \Pr\{WTP \leq BID^L \leq BID\} \quad [14]$$

To estimate the parameters (α and β), we use the maximum likelihood function, and we insert four indicator variables representative of the four possible combinations of answer YY, YN, NY, NN. Given our assumptions, the log-likelihood function of the sample is:

$$\text{Ln}(\theta) = \sum_{i=1}^{N} \{d_i^{yy} \ln Pr^{yy} + d_i^{yn} \ln Pr^{yn} + d_i^{ny} \ln Pr^{ny} + d_i^{nn} \ln Pr^{nn}\} \quad [15]$$

where d_i^{yy}, d_i^{yn}, d_i^{ny}, d_i^{nn} are the indicator variables. Estimates of parameters (vector θ) are obtained maximizing the log-likelihood function.

4.1 Results of base models

During the Summer 2005 we collected 1459 telephone interviews. The final dataset is composed by 1193 observations divided into the three levels of damage. Table 2 shows the frequencies of each combination of answers.

Table 2. Number of observations for each combination of answers

	Ordinary damage	**Intermediate damage**	**Heavy damage**
Yes-Yes	311	33	30
Yes-No	**218**	**31**	**12**
No-Yes	**116**	**14**	**12**
No-No	373	20	23
Number of cases	1018	98	77

The statistic base models, without covariates, relate the probability of a positive willingness to pay and the amount proposed, for the three levels of damage (Table 4).

Table 3. The base models

	β	Standard Error	b/St. Er.	P [| Z | > z]	E(DAP) (Dev.St.)
Ordinary damage					
Constant	1.02693661	0.07526426	13.644	0.0000	**€ 189**[4]
Bid	- 0.00543941	0.00025234	- 21.556	0.0000	(€11)
Intermediate damage					
Constant	1.02693661	0.07526426	13.644	0.0000	**€569**
Bid	- 0.00543941	0.00025234	- 21.556	0.0000	(€57)
Heavy damage					
Constant	0.84944524	0.23832590	3.564	0.0004	**€3.753**
Bid	- 0.00022631	0.403837D-04	- 5.604	0.0000	(€995)

The WTPs estimated in the three models point out reduced variability for ordinary damage (€ 189) and intermediate damage (€ 569), while in the case of heavy damage the estimated average

[4] The WTPs are for family.

is markedly higher (€ 3.753). For the ordinary damage the interval of the WTPs is €178-200, for the intermediate is €512-626 and for the heavy damage is € 2.758-4.748.

4.2 The extended model

The damage suffered by the few households denouncing a severe depreciation of their property as a consequence of an extreme proximity to the infrastructure is of a different nature with respect to that suffered by the large majority of other households. We therefore chose to investigate the latter separately. In the extended model we used only the observations pertaining to the ordinary and the intermediate damage, merged together. We inserted a dummy variable to identify the respondents who declared to suffer an intermediate damage. The cases used in this analysis are 1116. Beginning from the basic model we inserted sequentially groups of variables, selected from the information collected with the survey and with the data geo-referentiation; many of these indicators are presented in Table 4.

The groups of variables from which we have selected those to be introduced in the extended model were the perception variables, the individual variables and the local context variables.

Perception variables. The perception variables identify the type of impact considered as prevalent by the respondent, between environmental impact, human health risk, impact on the landscape. The components of damage that contribute to the formation of each respondent's WTP are many. In order to recognize which component prevails in the individual perceptions we inserted a question in the survey which enabled us to define three dummy variables.

Sociodemographic variables. Individual variables include income (a continuous variable, constructed as the intermediate value of the intervals of income proposed during the interviews), the presence of children in the family (a dummy variable), high education (dummy that identifies the respondent with an high school or more advanced degrees).

Local context variables. This group of variables includes features of the valuation context that we hypothesized as able to offer an explanatory contribute to the model. In particular:

- the density of power lines that crosses the municipality area (a continuous variable equal to the ratio between the corridor area and the total municipal area);
- the proximity of the house to other linear infrastructures, in particular roads and railways (these variables are codified as the natural logarithm of the euclidian distances);
- the presence of protected areas, classified as valuable areas through legislative acts: SIC, SIR, ZPS, parks, tie 1497/39 (a dummy for the presence of at least one tie in the municipality of respondent).

Table 4. Examples of GIS derivate variables

Phenomenon	Derivate Variables			
HVTLs	Distance from HVTLs	Line density (Percentage of municipal area within corridors)	Shape and size of towers	Voltage
Brownfield	Number of brownfield in the municipality	Presence of brownfield in the municipality	Presence of brownsfield within 700 m. from households	Distance from brownsfield
Airport	Presence of airport in the municipality	Presence of airport within 5000 m from households	Distance from airport	-
Roads	Presence of roads in the municipality	Presence of roads within 200 m. from households	Distance from roads	-
Highway	Presence of highway in the municipality	Presence of highway within 400 m. from households	Distance from highway	-
Rail	Presence of rail in the municipality	Presence of rail within 600 m. from households	Distance from rail	-
Environmental Protection	Municipality surface identified as protected area	Protection density (percentage of municipal area)	Presence of protected area in the municipality	-
Heritage	Number of historic buildings	-	-	-
Municipality features	Municipality surface (m^2)	Number of touristic activities	Tourists presence	Presence of woodland and its density
Features of built-up areas	Covered Ratio (Area covered/Municipality area)	Built-up areas fragmentation index (Perimeter of buildings/ Covered area)		

Table 5. The extended model

Variables	Model 1 β b/St. Er P [| Z | > z]	Model 2 β b/St. Er P [| Z | > z]	Model 3 β b/St. Er P [| Z | > z]	Model 4 β b/St. Er P [| Z | > z]
Constant	0.95330742 13.279 0.0000	0.06786684 0.466 0.6415	-0.61509767 -3.697 0.0002	-1.07817207 -2.160 0.0308
Bid	-0.00507275 -22.966 0.0000	-0.00518110 -23.013 0.0000	-0.00554886 - 23.113 0.0000	-0.00562801 -23.156 0.0000
Dv2	1.82657663 8.235 0.0000	1.87714904 7.513 0.0000	1.75686119 7.707 0.0000	1.77935773 7.705 0.0000
Environmental	*	0.71844151 2.670 0.0076	0.51539056 1.851 0.0641	0.56870162 2.029 0.0425
Health	*	1.22040014 7.477 0.0000	1.03455233 6.140 0.0000	1.05961885 6.207 0.0000
Visual	*	0.96214266 5.229 0.0000	0. 64485329 3.380 0.0007	0.67294288 3.487 0.0005
Income	*	*	0.02553965 6.472 0.0000	0.0245483 6.147 0.0000
Children	*	*	0.31940352 2.405 0.0162	0.34301245 2.553 0.0107
Education	*	*	0.60314005 4.890 0.0000	0.61040890 4.904 0.0000
Proximity	*	*	*	0.90100411 2.029 0.0089
Line Density	*	*	*	0.86245323 2.782 0.0054
LogRail	*	*	*	0.08645279 1.766 0.0774
LogRoads	*	*	*	-0.10332778 -2.126 0.0335
Protection	*	*	*	0.21330092 1.607 0.1080

4.3 Results of extended model

The four models presented in the paper are the outcome of the progressive introduction of groups of variables from the base model. First of all we have inserted the most commonly used variables – the perception variables and the individual variables. In the last model (number 4) we have inserted the LCVs. Table 5 reports the outputs of partial and extended models.

Perceptions variables. In Table 7 we show the frequencies of answers related with the different components of damage.

Table 6. Perceived damages

Variable	Frequency	Percentage
Visual impacts and landscape quality	265	21,66
Ecosystems impacts	74	6,05
Human Health impacts	667	54,53
No impacts	217	17,74

The component of damage about which the respondents have declared to be most concerned is human health risks (54,5%). For 21,6% of the respondents the visual impacts on landscape quality are the most relevant, whereas only 6% considers ecosystem impact as the most serious form of damage. A 17% of the sample declares not to suffer any damage from the infrastructure. The three dummy variables present a high statistic significance and the values of the coefficients reflect the ranking of frequencies observed, emphasizing the importance of component of damage related with human health.

Sociodemographic variables. The variable ‘income’, as generally happens, is statistically significant and with a positive coefficient, showing a positive relation with the dependent variable. People with a higher level of income are more willing to pay. High education as well has a positive and statistically significant coefficient. The dummy signalling the presence of children in the family presents also a positive and significant coefficient. People are more willing to pay to protect their children from the potential risks connected with human health.

LCV. The main problem in the use of the LCV linked to spatial data arises from the absence of a consolidate and standardized methodology for their coding. Spatial weights, based on distance measures, are used to express the proximity (in the case of the HVTL, roads or railways or other infrastructures): “On one hand, this state of affairs is not surprising, for there is no such thing as “true”, “universal” spatial weights, optimal in all situations: good candidates must reflect the properties of the particular phenomenon, properties which are bound to differ from field to field. On the other hand, this difficulty should not impede a more systematic investigation of *models* for spatial weights, starting with the question ‘which classes of models yield specified families of spatial weights, and what are the properties of the latter?’” (Bavaud, 1998).

Proximity. For codifying the variable referred to the proximity of households to the HVTLs we departed from the hypothesis that this variable could be important in the formation of perception only within a delimited corridor. We chose to use distance decay models, using a positive, decreasing, smooth function for the square of the euclidean distance (1/D^2) (Bavaud, 1998). The perception of damage decreases rapidly with increasing distance from HVTLs, while over a certain distance threshold (the dotted rectangle in Figure 4) the curve grows asymptotically close to the abscissas axis, settling down at a substantially homogeneous level of damage. The variable presents a positive coefficient (0.90100411) and a high statistic significance (P value= 0.0089).

This result is particularly interesting in relation to what has been previously observed in the valuation literature. In their survey, Kroll and Prestley (1992) assert that on the average the existing valuation studies do not find a significant decrease in the property values linked to the presence of HTVLs. Rosato *et al.* (2004) is among the CV studies in which the distance from the power line does not appear to influence the willingness to pay. In other studies a linkage is pointed out between the intensity of the damage and the distance from the lines. According to Colwell and Foley (1979), for instance, the decreasing effect is present only for houses within 60 meters from the power line, in particular for those situated within15 meters. In Sims and Dent (2005) the negative impacts diminish gradually and disappear for houses situated at 250 meters, whereas living within 100 meters from HVTLs causes a decrease in the property value between 6% and 17% with respect to a property with the same features situated far from the line. Overall the estimated depreciation varies sensibly: it can generally be included between 2% and 10% (Hamilton and Schwann, 1995), although it can reach values between 16% and 29% for houses near the towers (Boyer, 1978; Bond and Hopkins, 2000; DeRosiers, 2002).

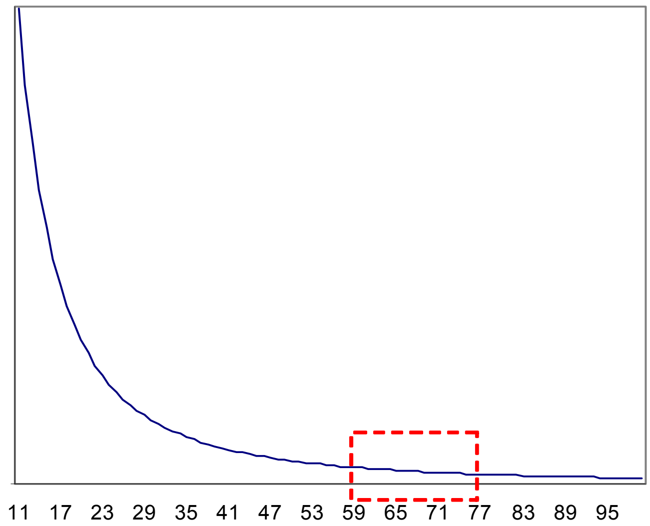


Figure 4. Spatial weight of the distance from the HVTL

Line density. The variable is built as the ratio between the municipal area in the HVTLs corridor and the municipal area. The choice to use the administrative boundaries (the municipal border) can cause some imprecision. Individual preferences stated in an *economic space* are determined by perceptions, so that perceived spatial context may not overlap with administrative boundaries. In spite of this, to distinguish the situations with a higher concentration of infrastructure we have inserted this variable as a proxy measure of quantity of lines in proximity of the respondents. The variable is significant (0.0054) and presents a positive coefficient (0.86245323), *ceteris paribus*, at increase of line density in proximity of house, increase the willingness to pay.

Interaction between linear infrastructures. The infrastructures for which we measured the distance from the respondent households are highway, railway, roads, airports and brownfields. Several ways to parameterize the distances are used in order to identify the proximity of respondents to the infrastructures considered. Only the variables referred to roads and railways, constructed as the log-natural of the euclidean distance from the houses, resulted statistically significant. The coefficients present opposite signs: positive for the railway, negative for the roads. The influence of the simultaneous presence of other infrastructures on the WTPs to remove the power line is not easy to explain univocally. On the one hand, the positive sign of the railway coefficient is interpretable as an effect of substitutability, in which the proximity of another infrastructures causes that WTP for getting rid of the HVTL to decrease. On the other hand, the negative sign of the coefficient connected with roads could identify a complementarity effect (proximity to another infrastructure causes a perceived damage higher than the sum of the two separate impacts).

In general the difficulty in recognizing a univocal interpretation could derive from the fact that aspects not capturable through the measure of the distances contribute to the formation of perceptions. The house position can be, for example, a relevant element in describe this phenomenon. In Figure 5, household A, even if at same distance from the railway as household B, is situated between two infrastructures, with a direct visual impact from both. We can hypothesize that the perceptions of household A are different from those of household B, whose perception of the damage from the HVTL may be mitigated or substituted by the interposition of the railway.
In section 5 (the local model) we investigate the stationarity of these coefficients in order to check whether a local scale of analysis is better suited to analyze such kind of variables.

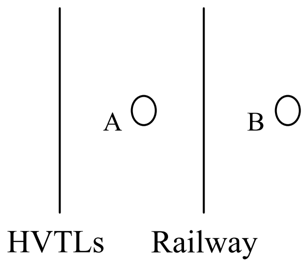


Figure 5. An example of interaction between infrastructures.

Environmental and landscape quality. Factors concerned with environmental and landscape quality may also influence the perception of damages from HVTLs. The geo-referenced indicator used is the presence of protection programmes and ties. We have defined a dummy that identifies the presence of at least one of these elements in the municipality of the respondent, The ties considered are:
- sites of communitarian interest (SIC and Natura 2000),
- sites of regional interest (SIR),
- special protection area (ZPS),
- parks,
- environmental ties (Italian law 1497/1939)[5].

The variable is significant at the 90% level and presents a positive coefficient (0.21). The damage caused by HVTLs is more strongly perceived in municipalities with protected areas, which we may reasonably think of as those recognised of high environmental and landscape value.

[5] The main legislative references are::
- Regional Law 20/89 (Environmental and landscape tie);
- Law n. 431/85;
- Ministerial Decree 01/08/85.

4.4 Goodness of fit

As first measure of goodness of fit and to comparing the basic model to the models with an incremental number of covariates, we use the ‘sequential classification procedure’, an extension of standard classification approach (Kanninen and Khawaja, 1995), which was used to estimate percentage of fully, correctly, classified cases (FCCC).

Table 7. The ’sequential classification procedure’.

	ICCY	ICCN	ICCT	ICCC	FCCY	FCCN	FCCT	FCCC
Model 1(basic model)	353	315	668	**59.80%**	226	235	461	**41.27%**
Model 2 (basic model + perception variables*)	341	369	710	**63.56%**	212	289	501	**44.85%**
Model 3 (basic model +perception variables +sociodemographic variables**)	380	362	742	**66.42%**	240	278	518	**46.37%**
Basic model + perception variables + sociodemographic variables + Local Context Variables***	382	373	755	**67.59%**	247	285	532	**47.62%**

** The perception variables used are: Ecosystem Impact, Health Impact, and Landscape Impact.*
*** The sociodemographic variables used are: Income, Presence of children, High education.*
****Local Context Variables used are: Power lines proximity, Density of pipelines, Railway proximity, Highway proximity, and Environmental and landscape quality.*

The outputs point out the importance of the introduction of LCV in the extended model. In particular the percentage of FCCC for model 3 and for model 4 confirms an increase in the explanatory power of the model with LCV.

In addition to the ‘sequential classification procedure’, that does not represent a hypothesis test, we have performed a LR test to verify further the explanatory contribute of the local context variables. Also this second test confirms the validity of the introduction of the LCV in the model, refusing the null hypothesis with a probability above 99%.

5. Local model: the GWR analysis

In the analysis presented at paragraph 4.2, the vector of the parameters $\boldsymbol{\beta}$ identifies the set of relationships between the observed values of the explanatory variables and the interval data of WTP. We are now interested in investigating whether these relationships may be considered invariant with respect to the location of the observation, or rather they are locally specified. If these relationships remain the same along the regional HVTL network, the analysis could be appropriately described at the global level by a single vector $\boldsymbol{\beta}$. Now we are interested in shifting our attention from global parameters to local heterogeneity. Households of the HVTL survey data are located all along the routes, in the infrastructure corridor. As they are geo-referenced, *regions* of data may present different local values of $\boldsymbol{\beta}$ with respect to the global dataset, and it is also possible to observe changes in the set of relevant explanatory variables. Local variations of $\boldsymbol{\beta}$ can be estimated through the GWR approach (Fotheringham, 1997). For each single observation GWR is capable of accounting for the effective relative spatial distribution of all other observations, estimating a set of $\boldsymbol{\beta}$ that is location specific**.** In the next sub-section we briefly describe the GWR model, in the linear and the binomial (logistic) case, to point out the spatial perspective of the model and what it can contribute.

5.1 Methodology for the linear GWR model

The GWR approach considers **β** as a function of (u,v), the geographical coordinates that identify a point (location) in space. Following Fotherihgham *et al*.(2002, chapter 2) we note that GWR assumes that for each observation *i* a local-specific set of **β** has to be estimated. A global linear regression model,

$$y_i = \beta_0 + \sum_k \beta_k x_{ik} + \varepsilon_i \qquad [16]$$

can be rewritten, in the local GWR approach, as

$$y_i = \beta_0(u_i, v_i) + \sum_k \beta_k(u_i, v_i) x_{ik} + \varepsilon_i \qquad [17]$$

Each of the k parameters could be graphically described as a surface, whose height is represented by the local value of β_k. This is "geographically weighted" in the sense that, for each observation i , GWR considers nearby data as more relevant than distant ones by implementing spatial weights. Hence, proximity is an indicator of the relevance of data points around i in the estimation of the model to emphasize the local value of parameters in the regression.
The matrix form for the case of fixed parameters, with no spatial dependency of **β**,

$$\boldsymbol{\beta} = (\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{Y} \qquad [18]$$

in GWR is substituted from the estimate of the locally-specific

$$\boldsymbol{\beta}(i) = (\boldsymbol{X}^T\boldsymbol{W}(i)\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{W}(i)\boldsymbol{Y} \qquad [19]$$

where $\boldsymbol{W}(i)$ is the weighting matrix relative to any observation *i* that is a square matrix with all off-diagonal values set to zero and, on the diagonal, the terms w_{ij} , as an indicator of proximity of other *js* observations to *i*. Fotheringham *et al.* (2002) suggest, as a rule to express proximity, to specify a parameter, the bandwidth, which is a spatial distance that measures the spatial area that is considered as the local spatial context of *i*. Spatial kernels are specified to assign weights to data that lie inside the local spatial context of *i*, while other data outside the boundary parameterized from the bandwidth are assigned a weight of 0.

The results of GWR are sensitive to the specification of the bandwidth and the choice of the weighting function. A step further in the realism of modelling local spatial contexts has been achieved with adaptive kernels, that make it possible to calibrate the extension of the local area depending on the density of data, that can meaningfully vary over space. The bandwidth can be adjusted by measuring the distance of the *Nth* nearest data. For a continuous weighting function, Fotheringham *et al.* suggest this structure of the bi-square kernel: $w_{ij} = [1-(d_{ij}/b)^2]^2$, if *j* is one of the N nearest data respect to *i*, where d_{ij} is the distance between *i* and *j* , and *b* is the distance between *i* and the Nth nearest neighbour.

5.2 Logistic GWR regression for the analysis of HVTL data

Beside the GWR linear models, non-linear (logistic and Poisson) regression models have been developed, and are now available inside software packages[6]. Set in a generalized linear model

[6] Our models are estimated with the software GWR release 3.0.

framework, these models are not estimated by OLS. They are specified as general linear models, using the *Iterative Reweighted Least Squares* algorithm (Green, 1984) and the pointwise estimation of parameters is obtained via log-likelihood.

We use this procedure to perform a geographic weighted logistic regression (GWLR) on HVTL data. The standard global model presented in 4.2 used information about WTP specified as interval data, from the sequence of the two dichotomous elicitation questions. The GWLR analysis is performed on single bounded interval data. The probability that a respondent answers "yes" to a given amount of the bid and the other explanatory variables, in the general case is (Hanemann, 1984):

$$\Pr_i(Yes) = \frac{e^{(\beta_0+\beta_1 BID+\beta_2 x_2+...+\beta_k x_k)}}{1+e^{(\beta_0+\beta_1 BID+\beta_2 x_2+...+\beta_k x_k)}} \qquad [20]$$

The equivalent, in the geographically weighted regression approach estimates a specific vector $\boldsymbol{\beta}_i$ for each location $i(u_i, v_i)$:

$$\Pr_i\{Yes|(u_i,v_i)\} = \frac{e^{(\beta_{i0}+\beta_{i1} BID+\beta_{i2} x_2+...+\beta_{in} x_n)}}{1+e^{(\beta_{i0}+\beta_{i1} BID+\beta_{i2} x_2+...+\beta_{in} x_n)}} \qquad [21]$$

it can be rewritten using the appropriate logit link function in linear form as:

$$\log\left(\frac{\Pr_i\{Yes|(v_i,u_i)\}}{1-\Pr_i\{Yes|(v_i,u_i)\}}\right) = \beta_{0i} + \beta_{1i} BID + \beta_{2i} x_{2i} + ... + \beta_{ki} x_{ki} \qquad [22]$$

The assumption of independent and identically residuals does not affect the GWR binomial specification, as here it is derived from the generalized linear model; hence, it is possible to avoid spatial autocorrelation bias, and endogeneity from omitted variables may be more efficiently controlled.[7]

Output of a simple GWR logistic model that use the bid and a constant term is visualized as a map that represent unconditional median WTPs (Hanemann,1984; Haab, McConnell, 2002). The spatial weights are given from an adaptive bi-square kernel with a bandwidth relative to the nearest 200 observation. The extended models calibrated in the last section (model 1-2-3-4) are re-estimated with the same local approach pointing out the predictive performance and the spatial non-stationarity of the relationship between WTPs and their determinants. Results are presented in the next paragraph, with numerical outputs that summarize the statistical distribution of estimated coefficients, and maps to visualize the spatial patterns of values.

5.3 Results of local models with GWR

Estimates of the $\boldsymbol{\beta}_i$ for each location $i(u_i, v_i)$ obtained with the logistic GWR algorithm are not presented as a single output (it would require 1116 tables). We are mainly interested in testing the spatial stationarity of variables that, in the global approach, are statistically significant. GWR, in non-linear models, does not *create* information where we have no data. So, smoothed surfaces which spatially model variations of $\boldsymbol{\beta}_i$ or of the standard error are not available. However, local pointwise values of $\boldsymbol{\beta}_i$, standards error and residuals are produced in a tabular form, in which each row provides parameters referred to each local logistic model. As these location are geo-referenced, GIS maps are created to visualize tabular outputs of GWR (see figures 6,7,10,11,12,13). Themes (we use this term meaning GIS maps) are proposed to

[7] A more complete and detailed explanation of the generalized linear model approach and IRLS applied to GWR framework can be found in chapter 8 of Fotheringham *et al.* (2002).

illustrate, with graduated colours or symbols, spatial patterns in the distribution or β_i's values.

Four models corresponding to those estimated applying the global approach were calibrated, and summary results are presented in Tables 9,10,11,12. A bid-constant model is also presented as preliminary test (model 0 in Table 8). A conventional logit estimation of a global model (corresponding to a standard single bounded approach) is presented in the first part of each table. The distribution of the local estimates is represented by percentiles in the lower part of the output table. Fotheringham *et al.* (2002) propose to consider the interquartile range as a measure of the dispersion of local estimates of β, to compare it with the double of the global standard error of β.

If the dispersion of the local model is less or equal to the range described by the global model, the relationship can be considered as spatially invariant. It is not a formal test like those based on parametric testing of hypothesis, and other methods are available to test spatial stationarity[8]. For each model are specified the result of this "informal" stability test is reported in the last column on the right. We report also the Akaike information criterion and the percentage of total correct prediction of the model to compare goodness-of-fit among the four specifications of local models. The sequential classification procedure, used to calculate the percentage of fully corrected classified cases (FCCC) of the double bounded model in the global analysis (see section 4.4) is replaced here by the percentage of initially correct classified cases (ICCC) because this GWR model does not implement the sequential structure of the follow-up question about WTP. We are now estimating a sort of single bounded model set in a GWR framework.
To describe the performance of this simple logit model, we use this ICCC index.

Table 8.

0 - Global model

	β (global)	**Std.Err**	**T**	**Exp(β)**	
Constant	0.632	0.104	6.075	1.882	**ICCC of GWR model: 60,39%**
Bid	-0.002	0.000	-5.994	0.998	**Akaike:1508,882617**

0 - Local model

	Minimum	**25th percentile**	**50th percentile**	**75th percentile**	**Maximum**		**Stationarity**
Constant	0.071154	0.499205	0.663383	0.836941	1.396561	0,337736	(no)
Bid	-0.004781	-0.002876	-0.002189	-0.001130	0.000691	0,001746	(no)

As well as the composition of the set of explanatory variables, the local approach of GWR confirms the results of the global approach: adding blocks of variables concerned with perceptions (model 2), socio-demographic characteristics (model 3), and LCV (model 4) the explanatory power of the model is improved, as shown by the ICCC index. Also the Akaike Information Criterion, adjusted for the different number of variables, confirm that the extended model with LCV works better than models 1,2,3. Income is actually the only spatially invariant variable, while all the other variables exhibit relationships that are function of the geographic location (Tables 11 and 12).

Hence, we can now deal with the problem of designing and choosing the "right" scale for valuation, by examining where local results meaningfully diverge from those of the global model results from the local models differ significantly from those of the global one. We note that relationships vary across space in magnitude and sign. If the true estimates of some parameter are zero, this implies that locally the model does not correspond to the structure of preferences stylized by the global analysis.

[8] For example Fotheringham *et al.*(2002) suggest a Montecarlo analysis to iteratively re-estimate the GWR model, testing the stability of β with respect to changes in the geographical coordinates of the data.

Table 9

1 - Global model

	β (global)	Std.Err	T	Exp(β)
Constant	0.716	0.106	6.772	2.047
Bid	-0.003	0.000	-7.641	0.997
Dv2	1.398	0.258	5.419	4.048

ICCC of GWR model: 61,55%
1477,914834

1 - Local model

	Minimum	25thpercentile	50thpercentile	75thpercentile	Maximum		Stationarity
Constant	-0.013148	0.618707	0.793332	0.919158	1.673678	0,300451	(no)
Bid	-0.007778	-0.004026	-0.003206	-0.001700	0.000304	0,002326	(no)
Dv2	-1.199736	0.834077	1.721674	2.293456	7.257594	1,459379	(no)

Table 10

2 - Global model

	β (global)	Std.Err	T	Exp(β)
Constant	-0.306	0.175	-1.744	0.736
Bid	-0.003	0.000	-7.528	0.997
Dv2	1.224	0.265	4.625	3.401
Health	1.352	0.180	7.506	3.866
Visual	1.238	0.205	6.043	3.449
Environmental	0.853	0.295	2.895	2.347

ICCC of GWR model: 69,17%
Akaike:1419,523008

2 - Local model

	Minimum	25thpercentile	50th percentile	75th percentile	Maximum		Stationarity
Constant	-2.575661	-0.631701	-0.249701	0.081092	0.840548	0,712793	(no)
Bid	-0.008659	-0.004161	-0.003235	-0.001649	0.000163	0,002512	(no)
Dv2	-1.257181	0.603745	1.627916	2.229595	7.011445	1,62585	(no)
Health	0.249086	1.121476	1.441108	1.882395	3.277228	0,760919	(no)
Visual	-0.127792	0.944051	1.303579	1.721445	3.992821	0,777394	(no)
Environmental	-3.429490	0.473787	1.068087	1.399580	3.197007	0,925793	(no)

Table 11

3 - Global model

	β (global)	Std.Err	T	Exp(β)
Constant	-0.942	0.198	-4.757	0.390
Bid	-0.003	0.000	-8.155	0.997
Dv2	1.310	0.275	4.769	3.705
Health	1.193	0.189	6.315	3.296
Visual	0.948	0.216	4.399	2.581
Environmental	0.621	0.311	1.999	1.861
Income	0.020	0.004	4.447	1.020
Education	0.659	0.140	4.711	1.933
Children	0.558	0.152	3.661	1.747

ICCC of GWR model: 75,08%
Akaike:1338,428135

3 - Local model

	Minimum	25th percentile	50th percentile	75th percentile	Maximum		Stationarity
Constant	-3.173268	-1.177027	-0.935473	-0.623370	0.209765	0,553657	(no)
Bid	-0.008771	-0.004953	-0.004088	-0.002094	-0.000239	0,002859	(no)
Dv2	-0.846745	0.780388	1.840181	2.337680	7.370223	1,557292	(no)
Health	-0.178167	0.967054	1.350071	1.723945	3.298590	0,756891	(no)
Visual	-0.822305	0.513881	1.062694	1.459734	3.901174	0,945853	(no)
Environmental	-2.514896	0.291855	0.644526	1.019167	3.195106	0,727815	(no)
Income	-0.016439	0.009933	0.022942	0.029282	0.064806	0,019349	(si)
Education	0.894006	0.488019	0.745740	1.004784	1.656454	0,516765	(no)
Children	-0.663663	0.087319	0.607288	0.888594	1.579931	0,801275	(no)

Table 12

4 - Global model

	β (global)	Std.Err	T	Exp(β)
Constant	-1.706	0.575	-2.968	0.182
Bid	-0.003	0.000	- 8.093	0.997
Dv2	1.320	0.279	4.724	3.743
Health	1.202	0.192	6.275	3.328
Visual	0.961	0.218	4.401	2.613
Environmental	0.685	0.313	2.187	1.984
Income	0.019	0.004	4.266	1.019
Education	0.636	0.142	4.493	1.890
Children	0.585	0.154	3.788	1.795
LineDensity	1.224	0.359	3.409	3.402
LogRail	0.079	0.056	1.414	1.082
LogRoads	-0.062	0.055	-1.123	0.940
Protection	0.121	0.152	0.794	1.128
Proximity	0.811	0.390	2.082	2.251

ICCC of GWR model: 79,03%
Akaike:1326,564302

4 - Local model

	Minimum	25th percentile	50th percentile	75th percentile	Maximum		Stationarity
Constant	-5.300623	-2.799407	-1.964340	-0.353405	1.766502	2,446002	(no)
Bid	-0.009508	-0.005286	-0.004270	-0.002289	0.000202	0,002997	(no)
Dv2	-0.727697	0.887917	1.992770	2.519351	7.868767	1,631434	(no)
Health	-0.193938	1.100674	1.394348	1.819803	3.743329	0,719129	(no)
Visual	-0.950226	0.661996	1.184300	1.607356	3.996915	0,94536	(no)
Environmental	-1.449835	0.460547	0.797623	1.160309	3.495641	0,699762	(no)
Income	-0.014372	0.010525	0.023153	0.030404	0.075621	0,019879	(si)
Education	-0.871480	0.471953	0.753430	1.041817	1.727084	0,569864	(no)
Children	-0.786942	0.081254	0.649636	0.990228	1.759903	0,908974	(no)
LineDensity	-3.094300	0.693939	1.366477	2.477696	7.511284	1,783757	(no)
LogRail	-0.518308	0.00062	0.095801	0.208158	0.660176	0,207529	(no)
LogRoads	-0.857616	-0.205835	-0.050447	0.024237	0.363307	0,230072	(no)
Protection	-1.661092	-0.173333	0.155619	0.458711	1.889950	0,632044	(no)
Proximity	-12.589854	-0.188157	0.647534	1.654037	10.630685	1,842194	(no)

The same happens for relationships that locally may have a sign different from what we found in the global analysis. This implies that attributes may have two opposite effects on the values of WTP. Consider, as example, Figure 6: the unconditional values of WTP estimated from the simple bid-constant specification (Table 8) are illustrated in a pointwise theme.

In Figure 7 we report the output of a GIS procedure that produces an interpolation of scatter point values, so as to enable a visualization of the WTP values. We use an *inverse weighted distance* algorithm (see Appendix B). GWR may be intended also as an explanatory technique to obtain a zoning scheme, as it makes possible to see "clusters" or sub-areas where relationships are more stable. In Figure 6 and 7 different levels of median local WTP are mapped. The shift from a scatterplot to a zoning scheme, anyway, must take into consideration that observable data are placed only within HVTL corridors.

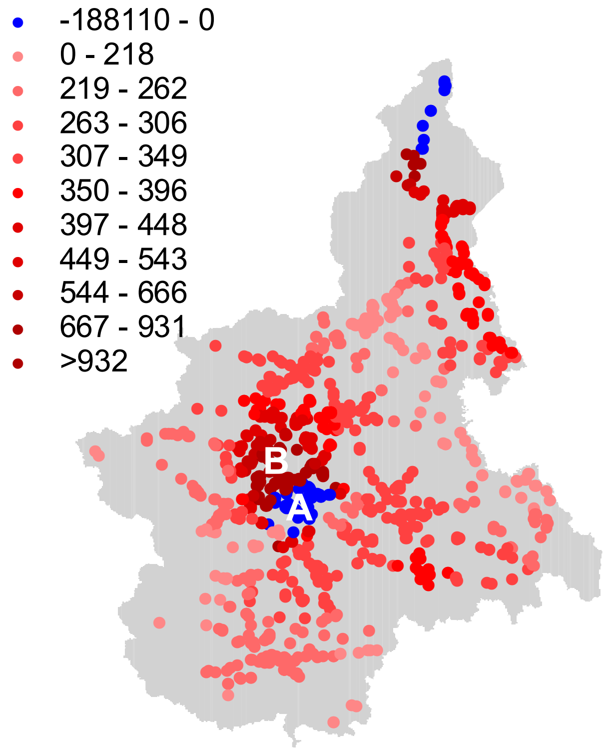


Figure 6 – Median WTP

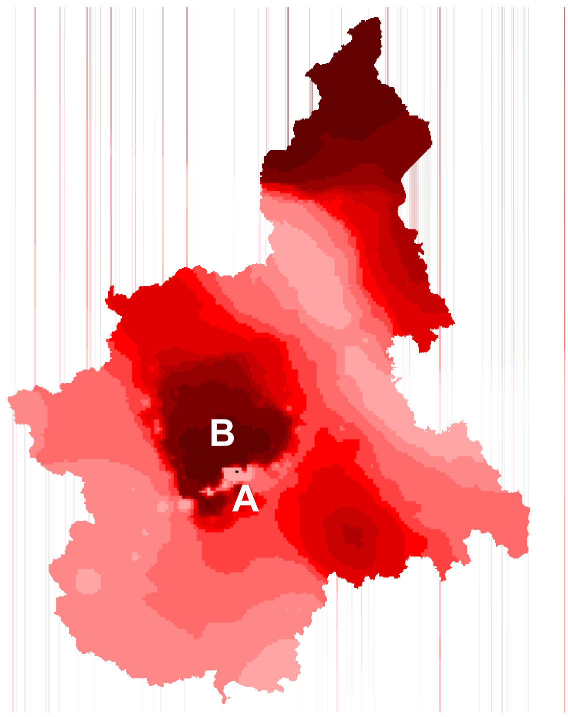


Figure 7 – Interpolation of Median WTP

Zone A highlights a group of numerous observations with negative WTP. In search of explanations for this local behavioural anomy in the values of WTP we deepened the contextual analysis. Available GIS regional maps showed that zone A is located near a XII century Abbey, and, most importantly, that there is in the area an important electric power plant towards which many HVTL converge from all directions. In Figure 9 the red broken lines are HVTL routes, and Figure 8 shows a view of the plant (see also Figure 17, 18, 19). Negative WTP could then be explained as the result of protest answers. Households in the area could have considered the elimination of the power line near to their house as an implausible scenario, and reject the experimental design. The appropriate scenario for a CV experiment in that area should be specifically designed so as to value externalities from the infrastructural complex, and should be locally implemented.

We report for all the local context variables the spatial distribution of GWR estimates. Figures 10-13 are GIS maps where observations are coloured in blue if the coefficient changes side with respect to the global model. The interval for the coefficients is specified in the legend of each map.

The proximity of the respondent to the HVTL (Figure 10) confirms at a local level the results of the global analysis. In this case, are present areas that not agree with the global analysis: two zones in the central and southern part of the region with grouped “blue points” are areas at medium altitudes where there are hills and smoothed skylines. In that areas the HTVL towers are not as large as elsewhere. Pylons are much higher (and visible) in mountain areas, while in flat rural areas even small sized tower are visible at long distances. This suggests that modelling distance decay functions for linear infrastructures is more complex than previously appreciated, and that global analysis have to carefully consider geo-morphological issues beside spatial patterns, if conducted in wide regional areas. Also the density of HVTL seems to be associated with the visual encumbrance of the infrastructure relative to the size of lines and to the presence of hill landscapes. Figure 11 points out that the central hills area gets most of "blue" observations for which the density of lines have a relationship not consistent with the results of the global model.

Figure 8

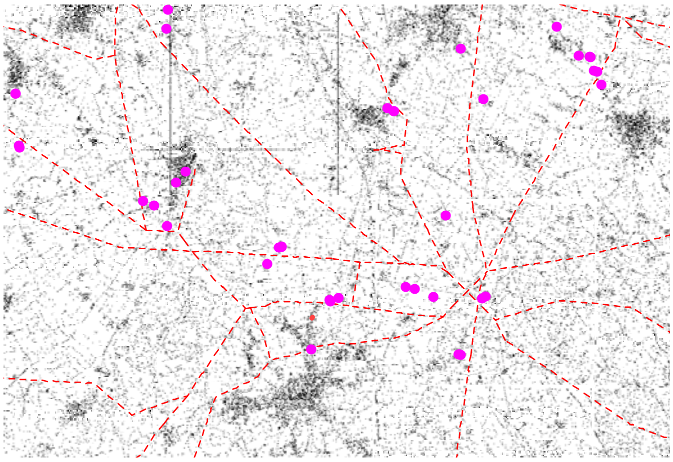


Figure 9

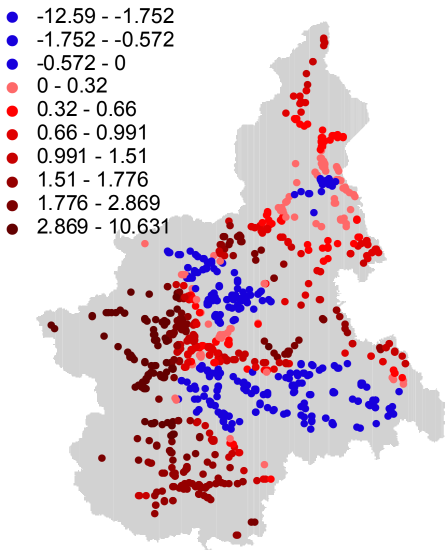


Figure 10 – Proximity to HVTL

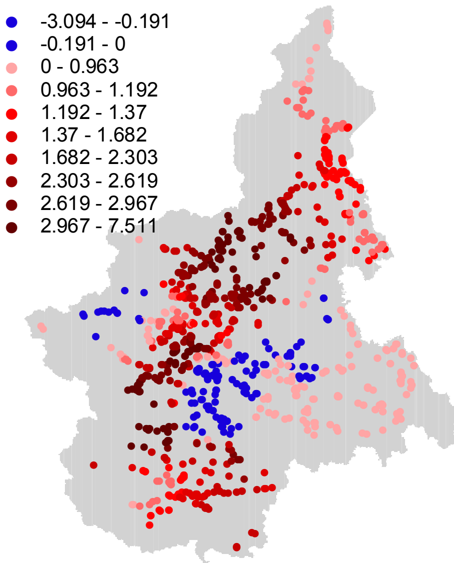


Figure 11 - Density of HTVL

Local effects on the perceived damage of HVTL determined from the presence of other linear infrastructures, formalized with a distance-based approach, are mapped in Figure 13 (log-distance from railways) and in Figure 11 (log-distance from highroads). Note (see Table 12) that the median value of local parameters both in the case of railways and in the case of highroads are close to zero. The fact that we have well balanced positive and negative effects may have different explanations. A first one is that distance-based variables are well specified as indicators of the interaction between infrastructures. Depending on individual preferences, tastes and circumstances, adding another infrastructure to the HTVL we may observe substitution or complementarity effects. On the one hand the proximity of another infrastructure may lower the

WTP for HVTL, because household allocate a part of it to another "prioritized" externality (positive signs of the log-distances). On the other hand, log-distances may show a negative sign if the proximity to another infrastructure increases the perceived damage more that the sum of the two separate impacts. Another explanation is that the log distance may be highly correlated to the proximity to the HVTL. Which effect prevails can depend on the relative location of the respondent or on spatial relationships among routes. If the two linear infrastructures are parallel to one another (Figure 15), or if two routes are crossing (Figure 14), or if there are complex multiple reticular patterns of the routes (Figure 16), then the distance is too stylized an indicator for a physical modelling approach.

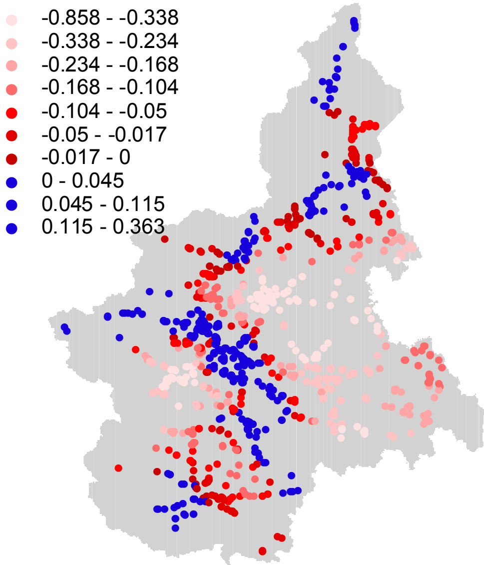


Figure 12 – Highroads

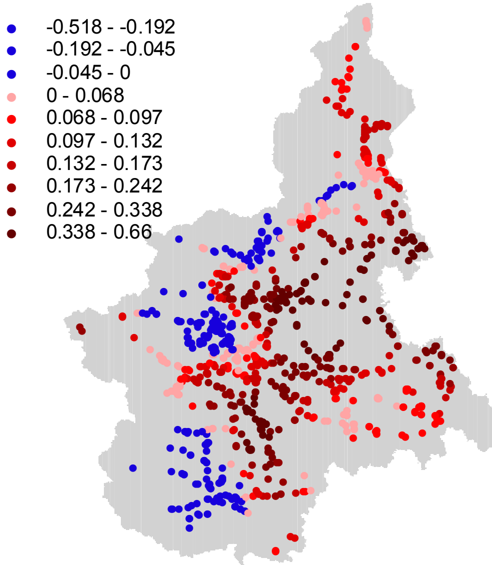


Figure 13 – Railways

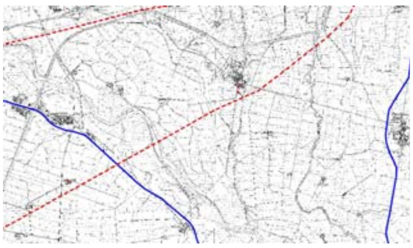


Figure 14

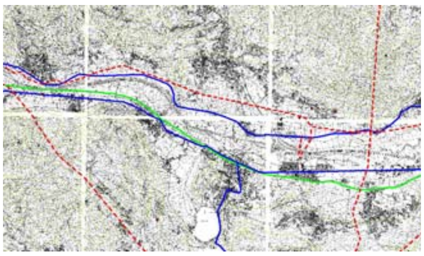


Figure 15

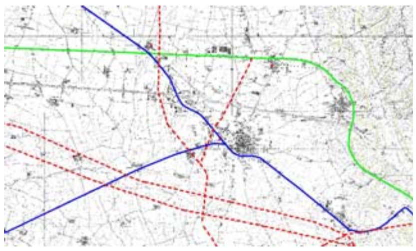


Figure 16

Figure 17

Figure 18

Figure 19

Table 13 summarizes, for the four specifications (model 1,2,3,4) values of the sequential classification index proposed by Kanninen (the initial and final percentage of correct predictions) for the double bounded approach and the GWR. These are reported to give an idea of the goodness-of-fit of the two approaches, comparing the ICCC index (the percentage of corrected predictions at the first WTP question). GWR has a meaningful higher provisional capability also in the extended models. Also, looking at the difference between ICCCs of model 3 and 4, we see that the introduction of LCV has a stronger effect in the GWR model.

Table 13

	Global logit double-bounded model		**GWR model**	
	ICCC	**FCCC**	**ICCC**	**FCCC**
Model 1(basic model)	59.80%	41.27%	61,55%	-
Model 2 (basic model + perception variables)	63.56%	44.85%	69,17%	-
Model 3 (basic model + perception variables +individual variables)	66.42%	46.37%	75,08%	-
Model 4 (Basic model + perception variables + individual variables + Local Context Variables***	67.59%	47.62%	79,03%	-

6. Conclusions

Scope of this paper was to test and compare two alternative models of CV data analysis employed in the evaluation of perceived impacts from linear infrastructures (HVTL). A well-experimented double-bounded approach to logistic regression and a single bounded approach based on the Geographically Weighted Regression have been used to analyze a dataset from an extensive survey on High Voltage Transmission Lines. Both the standard and the GWR models have demonstrated that WTP models are improved by the inclusion of variables that are concerned with characteristic of the spatial context within which the externalities are perceived.

There is no previous literature using GWR in the field of environmental economics. Our application is meant to explore a few different ways to use this methodology. The relevance of GWR stems from the fact that it was designed to estimate regressive models aimed at analysing the spatial stationarity of relationships between the dependent and the explanatory variables. The results of our analysis on the HVTL dataset show that the perceived damage expressed by the WTP of respondents are quite sensitive to the spatial localization of the household. We use a GWR estimation of the WTP parameters to test the consistency of these locally specific parameters with the spatially invariant parameters of the double bounded model, defined as global. The only variable that is definitely spatial invariant, according to our results, is ‘income’. Indicators of the goodness of fit show that GWR performs better than the conventional model.

GIS combined with GWR offers an exploratory tool which can be used to perform clustering, or to investigate areas where the conventional global model is not properly specified. In this integrated framework, GWR has been used to investigate the operative limits of the global approach. By mapping the values of WTP we have been able to identify different sub-regional areas with similar levels of median unconditional WTP. Our application also points out that the content validity of CV surveys may also be affected by a loss in plausibility due to locally determined circumstances. A GWR analysis, by revealing the geographic location of these irregularities, helps investigating and fixing such weaknesses.

The introduction of LCVs does not dramatically improve the fit of the global model, precisely because preferences are spatially grouped (clustered) as highlighted in the maps obtained with the local model. Consistency of fixed "universal" parameters has been tested employing GWR, to see if they should be included in the global analysis. The result of GWR can be used to address a *a posteriori* correction of the specification of those sub-samples that present locally specific effects, which would be biased if considered at a more aggregate scale (as it happened in our case study for density of and proximity to the HVTL). This issue is linked to the more general methodological problem of choosing the "right" scale for statistical inference, as shifting the point of view may change results that in turn will affect environmental planning and, more generally, policies and decision-making. It also raises the issue of the correspondence between the spatial viewpoint of the valuation exercise and the scale of the policies that will be informed by its results. GWR, integrated with the standard approach, allows the preference analysis to focus, as a sort of "zoom", on different scales, thus adjusting the flexibility of contingent valuation studies.

REFERENCES

ALBERINI. A, RIGANTI P., LONGO A., (2005) Public Preferences for Land uses' changes: valuing urban regeneration projects at the Venice Arsenale, paper presented at 45th Congress of the European Regional Science Association "Land Use and Water Management in a Sustainable Network Society" Vrije Universiteit Amsterdam.

ANSELIN L., (2002), "Spatial externalities, spatial multipliers and spatial econometrics", paper forthcoming in Regional Science and Urban Economics.

ATKINSON G., DAY B., MOURATO S., PALMER C. (2004), " 'Amenity' or 'eyesore'? Negative willingness to pay for options to replace electricity transmission towers". Applied Economics Letters, Vol. 11, pp. 203–208.

BATEMAN I.J., JONES A.P., LOVETT A.A., LAKE I.R., DAY B.H. (2002), "Applying Geographical Information Systems (GIS) to Environmental and Resource Economics", Environmental and Resource Economics, Vol. 22, pp.219-269.

BAVAUD F. (1998), "Models for spatial weights: a systematic look", Geographical Analysis, Vol. 30, pp. 153-171.

BOCKSTAEL N.E. (1996), "Modelling Economics and Ecology: The Importance of Spatial Perspective", American Journal of Agricultural Economics, Vol. 78, No. 5, pp. 1168-1180.

BOND S. (1995), "The impact of transmission lines on property values". New Zealand Valuers' Journal, pp. 26 – 30.

BOYER J.C., MITCHELL B., FENTON S. (1978), "The socio-economic impacts of electric transmission corridors: a comparative analysis", Report prepared for the Royal Commission on Electric Power Planning, University of Waterloo.

CAMERON T.A., JAMES M.D. (1987), "Efficient Estimation Methods for "Closed-ended" Contingent Valuation Surveys", The Review of Economics and Statistics, Vol. 69, No.2, pp.269-276.

COLWELL P.F., FOLEY K.W. (1979), "Electric transmission lines and the selling price of residential property", The Appraisal Journal, October, pp.490-499.

COLWELL P.F. (1990), "Power Lines and Land Value", The Journal of Real Estate Research, Vol. 5, No. 1, pp. 117-127.

DELANEY C.J., TIMMONS D. (1992), "High voltage power lines: do they affect residential property values?", The Journal of Real Estate Research, Vol.7, No.3, pp. 315-329.

DES ROSIERS F. (2002), "Power Lines, Visual Encumbrance and House Values: A Microspatial Approach to Impact Measurement", The Journal of Real Estate Research, Vol. 23, No. 3, pp. 275-301.

FOODY G.M. (2004), "Spatial non-stationarity and scale-dependency in the relationship between species richness and environmental determinants fro the sub-Saharan endemic avifauna", Global Ecology and Biogeography, Vol. 13, pp. 315-320.

FOODY G.M: (2005), "Clarifications on local and global data analysis", Global Ecology and Biogeography, Vol. 14, pp. 99-100.

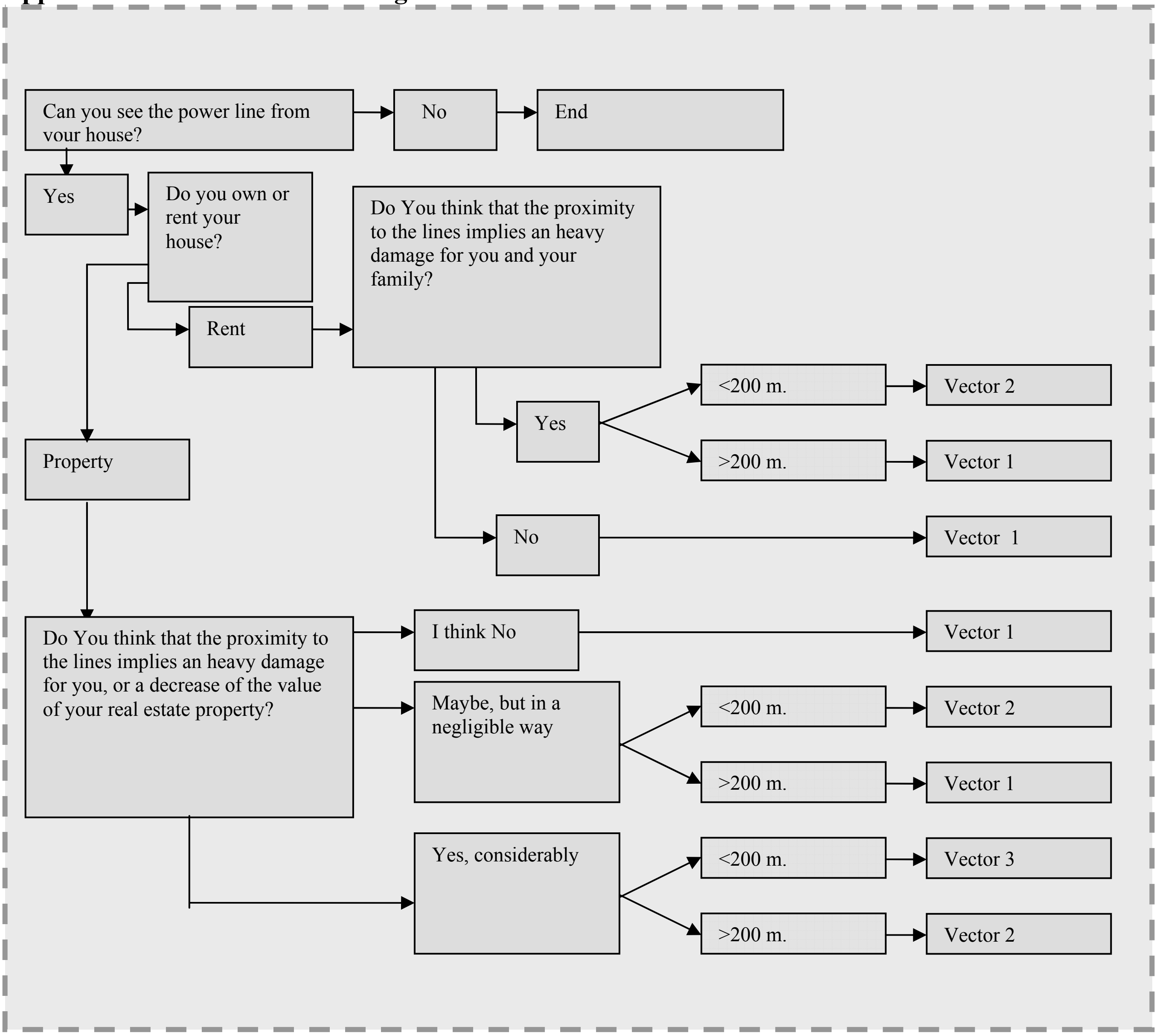
FOTHERINGHAM S. (1997), "Trends in quantitative methods: Stressing the Local", in Progress in Human Geography, Vol. 21, pp. 88-96.

FOTHERINGHAM A.S., CHARLTON M.E., BRUNDSON C. (2001), "Spatial variations in school performance: a local analysis using geographically weighted regression", Gepgraphical and Environmental Modelling, Vol. 5, pp. 43-66.

FOTHERINGHAM A.S., BRUNDSON C., CHARLTON M.E. (2002), "Geographically Weighted Regression: The Analysis of Spatially Varying Relationships", Chichester: Wiley.

GENELETTI D. (2002), “Ecological evaluation for environmental impact assessment”, PhD Thesis.
GENELETTI D. (2004), “Using spatial indicators and value functions to assess ecosystem fragmentation caused by linear infrastructures”, International Journal of Applied Earth Observation and Geoinformation, Vol. 5, pp. 1-15.
GREEN P.J., (1984), “Iterative Reweighted Least Square for Maximum Likelihood estimation and some robust and resistant alternatives”, in Journal of the royal Statistical Society, Series B (Methodological), Vol.46, No. 2, pp.149-192
GREGORY R., vonWINTERFELDT D. (1996), “The effects of electromagnetic fields from transmission lines on public fears and property values”, Journal of Environmental Management, No. 48, pp.201-214.
GUEVARA C., BEN-AKIVA M., (2005), “Endogeneity in residential location choice models”, paper submitted to the Transportation Research Board, 21st November.
HAAB T.C., MC CONNELL K.E., (2002), Valuing environmental and Natural Resources. The econometrics of Non-Market Valuation, Edward Elgar.
HAIDER M., HAROUN A., MILLER E.J. (2001), The Impact of Power Lines on Freehold Residential Property Values in the Greater Toronto Area, Canadian Regional Science Meetings 2001, Montreal, QC.
HAMILTON S.W., SCHWANN G.M. (1995), “Do High Voltage Electric Transmission Lines Affect Property Values?”. Land Economics, Vol. 81, No. 4, pp.436-444.
HANEMANN W.M. (1984), “Welfare Evaluation in Contingent Valuation Experiments with Discrete Responses”. American Journal of Agricultural Economics, Vol. 66, No. 3, pp. 332-41.
HANEMANN, W. M., LOOMIS, J. e KANNINEN, B. (1991), “Statistical Efficiency of Double-bounded Dichotomous Choice Contingent Valuation”, American Journal of Agricultural Economics, Vol. 73, pp. 1255-63.
JOHNSTON R.J., SWALLOW S.K., BAUER D.M. (2002), “ Spatial factors and Stated Preferences Values for Public Goods: Considerations for Rural Land Use”, Land Economics, vol. 74, No. 04, pp. 481-500.
KANNINEN B.J., KHAWAJA M.S. (1995), "Measuring Goodness of Fit for the Double Bounded Logit Model", American Journal of Agricultural Economics, No. 77 (November), pp. 885-890.
KINNARD W.H. Jr (1967), “Tower Lines and Residential Property Values”, The Appraisal Journal, Num. 35 (April), pp. 269-284.
KROLL C., PRIESTLEY T. (1992), The effects of overhead transmission lines on property values: a review and analysis of literature. Unpublished report prepared for the Edison Electric Institute Siting and Environmental Planning Task Force.
NAKAYA T., FOTHERINGHAM A.S., BRUNSDON C., CHARLTON M. (2005), “Geographical weighted Poisson regression for disease association mapping”, Statistics in Medicine, Vol.24, pp. 2695-2717.
ROSATO P., CANDIDO A., GALVAN A., MARTIN C. (2004), “La valutazione dei costi e dei benefici dell’interramento delle linee elettriche: Un caso di studio nella bassa pianura veneta”, Genio Rurale, Vol. 6, pp. 11-20.
SIMS S., DENT P. (2005), “High-voltage Overhead Power Lines and Property Values: A Residential Study in the UK”, Urban Studies, Vol.42, No. 4, pp. 665-694.
TEMPESTA T., MARAZZI M. (2005), “Disponibilità a pagare e disponibilità ad accettare per la riduzione dell’impatto paesaggistico delle linee elettriche dell’alta tensione”, Aestimum, No. 46, pp. 65-95.

Appendix A – Scheme for the assignment of BID vectors



Appendix B - Inverse weighted interpolation

The simplest form of inverse distance weighted interpolation is sometimes called "Shepard's method" (Shepard 1968). The equation used is as follows:

$$F(x,y) = \sum_{i=1}^{n} w_i f_i$$

where n is the number of scatter points in the set, f_i are the prescribed function values at the scatter points (e.g. the data set values), and w_i are the weight functions assigned to each scatter point. The classical form of the weight function is:

$$w_i = \frac{h_i^{-p}}{\sum_{j=1}^{n} h_j^{-p}}$$

where p is an arbitrary positive real number called the weighting exponent and is defaulted to 2. The weighting expotent can be modified by turning on the Use classic weight function option. h_i is the distance from the scatter point to the interpolation point or

$$h_i = \sqrt{(x - x_i)^2 + (y - y_i)^2}$$

where (x,y) are the coordinates of the interpolation point and (x_i,y_i) are the coordinates of each scatter point. The weight function varies from a value of unity at the scatter point to a value approaching zero as the distance from the scatter point increases. The weight functions are normalized so that the weights sum to unity.