Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria (INIA)
Available online at www.inia.es/sjar
Spanish Journal of Agricultural Research 2007 5(2), 119-129
ISSN: 1695-971-X

A multiple indicators and multiple causes model for valuation

L. Ambrosio Flores*, L. Iglesias Martínez, C. Marín Ferrer,
V. Pascual Gallego and A. Serrano Bermejo

Departamento de Economía y Ciencias Sociales Agrarias. Escuela Técnica Superior de Ingenieros Agrónomos. Universidad Politécnica de Madrid. Avda. Complutense, s/n. 28040 Madrid. Spain

Abstract

Value assessment (valuation) of certain real properties is commonly required for a variety of purposes. In this article an approach to define an estimator of the value of a real property, with good statistical properties is proposed. This estimator reaches a consensus between the various available methods. The estimator proposed uses all available information optimally, including the information which value indicators contain on the effect of contextual factors, and this furnishes it with a greater relative efficiency with regard to alternative estimators. It is also shown how the proposed approach is more general and flexible than the usual hedonic approach. The approach is illustrated using data on agricultural real properties and the relative efficiency of the compromise estimator is assessed using this same data.

Additional key words: hedonic price modeling, land valuation, latent variables, measurement errors, unbiased and optimal linear estimators.

Resumen

Un modelo de indicadores y causas múltiples para la valoración de fincas

Para una gran diversidad de propósitos frecuentemente se requiere la valoración de bienes inmuebles, tales como fincas. En este artículo se propone un estimador del valor de un bien inmueble dado, con buenas propiedades estadísticas. Este estimador es un compromiso entre los métodos propuestos en la literatura. El estimador utiliza de un modo óptimo toda la información disponible, incluyendo la que los indicadores de valor contienen sobre el efecto de factores contextuales, lo que da por resultado una mayor eficiencia relativa respecto a estimadores alternativos. También se muestra cómo la aproximación propuesta es más general y flexible que la aproximación basada en los precios hedónicos. Tanto para ilustrar el procedimiento de valoración, como para evaluar la eficiencia relativa del estimador de compromiso que se propone, se utilizan datos de una muestra de parcelas agrícolas.

Palabras clave adicionales: errores de medida, estimadores lineales insesgados y óptimos, modelos de precios hedónicos, valoración de fincas, variables latentes.

Introduction

Valuation (estimation of the value) of certain real properties has been necessary for a very long time, as an appraisal or basis of appraisal set by authorities for any purposes which may require this valuation (England, 2003). In the case of commodities, it is easy to find an objective standard on which to found the valuation process: the market price. However, in order to value real estates, such as an agricultural property, there are many possibilities (e.g. net realizable value, opportunity cost, or replacement cost), and there is disagreement over the best method to measure value (Lemke, 1966).

There is agreement in value theory on considering effective demand, scarcity of supply, exchange, and transferability as the sources and bases of value. However, there is disagreement about the definition of value, due to the fact that the attributes of value identified and emphasized by each school of economic thought have often focused on the central current social, political, and economics issues. This disagreement is a serious obstacle for the valuation process, since the definition of value has a key role both in directing data collection, as well as estimation procedures leading to an appropriate estimation of value (Grissom, 1985).

* Corresponding author: luis.ambrosio@upm.es
Received: 24-05-06; Accepted: 27-02-07.

To estimate the value of a given real property, the International Valuation Standards Committee (IVSC) suggests three different approaches: the sales comparison approach, the income capitalization approach and the cost approach (IVS, 2005). All three approaches presume that economic agents are rational and compare the benefits from buying or selling a property with investment alternatives. The economic rationale of the sales comparison approach is that no informed agent would pay more for a property than other agents have recently paid for comparable properties given that the general market conditions are the same. The economic rationale of the income approach for existing properties is that no economic agent will pay more for a property than he will retrieve by holding the property. Finally, the economic rationale of the cost approach is that no rational agent will pay more for an existing property than the cost of constructing an equally desirable substitute property (Corgel *et al*., 2001).

All three valuation approaches are consistent with economic reasoning and neither of these approaches is better than the other (Shiller and Weiss, 1999). The problem is that the appraiser has to make a final reconciliation step to reach a final estimate of the market value and no methodology exists on how to mix the different market values calculated by the three approaches into one final market value.

This paper focuses on a method to achieve a consensus among the various valuation methods proposed in valuation literature, in order to reach a final estimate of the market value of a given real property. This consensus should make optimal use of the available data. Instead of a value definition, a «valuation hypothesis» is proposed as a guide during the data collection process and the value estimation process. It is assumed that the value of a given property is more than a physical phenomenon —a quality inherent to the property— and is a constructed human perception (Hadley, 1928). Human perceptions are to a large extent socially acquired, and to estimate the value of the given property, it is assumed that it is possible to establish a social consensus about the set of its measurable attributes of value, which here are called multiple-causes of value.

This «valuation hypothesis» is also assumed in the hedonic approach. Hedonic price modeling is a prominent tool in analyzing real estate prices, and it has been used for calculating real estate price indices (Bailey *et al*., 1963; Sheppard, 1999; Malpezzi, 2002) as well as calculating the impact on property values from proximity to a landfill and other sources of environmental risks (Freeman, 1979; Reichert *et al.*, 1991; Nelson *et al.*, 1992). The approach is also known as hedonic regression in the literature because it relates observed prices to property-specific characteristics by regression methods. It is the core assumption of the hedonic approach that prices are given by a function of property-specific attributes (Cropper *et al.*, 1988), which will be called the hedonic assumption from now on.

In this paper, all approaches on the valuation problem suggested in the literature are taken into account, including the hedonic approach. Our interest focuses on how to best combine the value estimates given by the approaches proposed in the literature into one final market value estimate. In addition, our approach accounts for the effect due to contextual factor. Context can play a tremendously important role in human perception (Luke, 2004). This effect of unobserved contextual factors must be taken into account, in addition to property-specific characteristics, when the value of a real estate property is estimated.

In this study, a statistical approach is proposed to estimate the value of an agricultural property as the sum of two components: one stemming from measurable property-specific value attributes, and the other as a consequence of unobserved contextual factors, using data describing the indicators of value. The term «indicator» is used here to denote an uncertain measure of value, when this uncertainty is not measurable. In this sense, both subjective measures as well as some objective measures are considered indicators; for instance, the present discounted value of the future cash flow of the agricultural property being valued (capitalization approach) is considered an indicator, since there is uncertainty about the discount rate, and this uncertainty is rarely measurable. When it is possible to quantify the uncertainty of a measure of value, this measure will be called «estimate».

To derive both an estimator of the value of a given agricultural property and the standard error of this estimator, a multiple indicators and multiple causes model (MIMCM) is proposed (Robins and West, 1977; Skrondal and Rabe-Hesketh, 2004). It is a statistical model which describes the relationship between a set of indicators on the value of a given agricultural property (multiple-indicators), and the measurable attributes (multiple-causes) of this property. This model allows us to «borrow strength» from data relative to properties related to the property to be valued, in order to obtain the best estimate of the value of the property under consideration. Based on this model, a compromise among the various valuation methods proposed in the valuation literature is established and this consensus should be that which makes an optimal use of the available data, relative to multiple indicators and causes (attributes).

This compromise valuation could be useful to reduce the traditional disagreement on the best method to measure a defined value. The procedure is illustrated using a case study from agriculture, but it can be used in the remaining economic sectors as differences among valuation methods used in each sector are of a more practical than theoretical nature (ASFMRA-AI, 2000; IVS, 2005).

The purpose of this study is to find a statistical approach to the problem of estimating the value of an agricultural property, and the approach suggested consists in expressing this value as the sum of two components. The first component stems from measurable property-specific value attributes, which is the usual procedure in the hedonic approach. The other component evaluates unobserved contextual factors by using data on the indicators of value, for example, valuations arising from methods suggested by the IVSC.

The model

The relationship between the indicator y_{ij} of the value of an agricultural property i, and the measurable attributes of this property x_i, is established using the structural model:

$$y_i = x_i^T \beta + v_i \qquad [1]$$

$$y_{ij} = y_i + e_{ij} \qquad [2]$$

where [1] is the *behavioral equation* and [2] is the *measurement equation*.

The *behavioral equation* specifies that the value y_i is perceived as the sum of two components: one, $x_i^T \beta$, due to the measurable attributes included in the vector x_i (β is the vector of parameters representing the changes on y_i due to changes on x_i), and the other, v_i, is the effect due to unobserved contextual factors. An agricultural property with x_i attributes is expected to have a value $E(y_i|x_i) = x_i^T \beta$. Nevertheless, unobserved contextual factors can have an effect v_i on value. For measurement purposes, it is considered that this effect is a realization of a random zero mean variable $Ev_i = 0$, with variance σ_v^2 and covariance null $Cov(v_i, v_{i'}) = 0$ for $i \neq i'$. In other words, it is assumed that because of unobserved contextual factors the value y_i of a given agricultural property (i.e. given x_i) varies around its expected value with variance $V(y_i|x_i) = V(v_i|x_i) = \sigma_v^2$: the hedonic assumption is that σ_v is low with respect to $x_i^T\beta$.

The *measurement equation* specifies that each indicator of y_i in the set $\{y_{ij}; j = 1,2,...,J\}$ deviates from y_i in a certain (unobservable) quantity, $e_{ij} = y_{ij} - y_i$, due to the difficulties involved in perceiving β and v_i, as well as to measurement errors of the attribute x_i. The disagreement between valuation measurement results is modeled through this deviation, assuming that this deviation is a realization of a random variable of mean zero, $E(e_{ij}|y_i) = 0$ (so that $E(y_{ij}|y_i) = y_i$, and assuming the indicators are unbiased), and variance $V(y_{ij}|y_i) = V(y_{ij}|x_i, v_i) = V(e_{ij}|y_i) = \sigma_e^2$. In addition, it is assumed that the deviation e_{ij} is independent of the value y_i, that is $Cov(v_i, e_{ij}) = 0$, and that the deviation of two indicators are independent, $Cov(e_{ij}, e_{ij'}) = 0$; $\forall j \neq j'$.

Replacing [2] in [1] the reduced form is obtained,

$$y_{ij} = x_i^T \beta + u_{ij} \qquad [3]$$

where $u_{ij} = v_i + e_{ij}$ is a random perturbation of mean zero, $Eu_{ij} = Ev_i + Ee_{ij} = 0$, which makes y_{ij} vary around its expected value, $E(y_{ij}|x_i) = x_i^T \beta$, with variance $V(y_{ij}|x_i) = V(y_i|x_i) + V(y_{ij}|x_i, v_i) = V(u_{ij}) = \sigma_v^2 + \sigma_e^2$; where σ_v^2 and σ_e^2 are the so-called variance components.

The contextual factors effect v_i induces positive covariance among the indicators in the set $\{y_{ij}; j = 1,2,...,n_i\}$, and the covariance structure is:

$$Cov(y_{ij}, y_{i'j'}) = \begin{cases} \sigma_v^2 + \sigma_e^2; & \forall i = i'; j = j' \\ \sigma_v^2; & \forall i = i'; j \neq j' \\ 0; & \forall i \neq i' \end{cases}$$

The correlation coefficient between indicators of the value of a same property, $\rho = \sigma_v^2/(\sigma_v^2 + \sigma_e^2)$, is a measurement of the «level of agreement» among indicators [it is called «reliability ratio» in the measurement error models theory (Fuller, 1987; Robinson, 1991)], since $\rho = 1 - \sigma_e^2/(\sigma_v^2 + \sigma_e^2)$ and $\sigma_e^2/(\sigma_v^2 + \sigma_e^2)$ represent the proportion of the total variability $V(y_{ij}|x_i)$, due to the disagreement among indicators. In fact, in our approach any disagreement among the indicators $\{y_{ij}; j = 1,2,...,n_i\}$ of a given agricultural property i (i.e. given x_i) in a given context (i.e. given v_i) is considered to be due to the difficulties involved in perceiving β and v_i and to measurement errors of the attribute x_i, which is measured by σ_e^2. When there is total agreement among the indicators, then $\sigma_e^2 = 0$ and ρ takes its maximum value, $\rho = 1$ (for any $\sigma_v^2 > 0$). When the disagreement among indicators increases in such a way that the ratio σ_v^2/σ_e^2

approaches zero, then ρ decreases approaching its minimum value, zero. The «degree of agreement», ρ, will be used to predict v_i and to exploit the relationship between y_{ij} and x_i more efficiently.

The usual hedonic price model is specified by equation [1] making y_i equal to the market price and assuming that the variability of y_i is mainly due to measurable attributes x_i, and hence the variability, $V(y_i) = \sigma_v^2$, of y_i around its expected value, $x_i^T\beta$, is assumed to be low. It is also assumed that there is no autocorrelation among perturbations, $Cov(y_i, y_{i'}) = 0$ for all $i \neq i'$ and that the attributes are measured without error.

The empirical best linear unbiased predictor

It is assumed that associated with the i^{th} property there is a set of n_i indicators $\{y_{ij}; j = 1.2,...,n_i\}$ of y_i as well as a vector x_i of measurable attributes of the said property. The set $\{(y_{ij}, x_i); j = 1,2,...,n_i; i = 1,2,...,m\}$ can be observed for a sample of m properties. The consensus estimator proposed to estimate y_i is the empirical best linear unbiased predictor (EBLUP), $\hat{y}_i$, based on model [3]:

$$\hat{y}_i = x_i^T\hat{\beta} + \hat{v}_i \qquad [4]$$

where $\hat{\beta}$ is the empirical generalized least squared estimator (EGLS) of β, $\hat{v}_i = g_i(\bar{y}_i - x_i^T\hat{\beta})$ is the predictor of the contextual factors effect, $g_i = \hat{\sigma}_v^2/[\hat{\sigma}_v^2 + (\hat{\sigma}_e^2/n_i]$, and $\bar{y}_i = \frac{1}{n_i}\sum_{j=1}^{n_i} y_{ij}$.

Due to the fact that the two error components, (v_i, e_{ij}), induce autocorrelation among the perturbations in [3], the ordinary least square (OLS) estimator of β is inefficient: the more efficient estimator is the generalized least square (GLS). However, the GLS estimator is not feasible because it depends on the unknown variance components, σ_v^2 and σ_e^2. The EGLS used in [4] is the GLS once σ_v^2 and σ_e^2 have been replaced by consistent estimators (see Appendix for details). The usual hedonic price model assumes that there is no autocorrelation among perturbations and, under this assumption, OLS and GLS coincide.

The EBLUP has optimal statistical properties (Goldberger, 1962; Robinson, 1991). However, it is necessary to verify the assumptions of the model since the statistical properties of the EBLUP are optimal only if these assumptions are correct. A statistical test for this predictor is given in the Appendix.

Equation [4] is defined as an estimator and not an indicator of value, given that it is possible to measure its standard variation. To this end the mean squared error of the empirical predictor, $M\hat{S}E(\hat{y}_i)$, proposed by Prasad and Rao (1990) (also see Ghosh and Rao, 1994) will be used, and the standard error of $\hat{y}_i$ will be estimated as the square root of $M\hat{S}E(\hat{y}_i)$.

The prediction of the effect due to unobserved contextual factors, $\hat{v}_i = g_i(\bar{y}_i - x_i^T\hat{\beta})$, is proportional to the residual value $(\bar{y}_i - x_i^T\hat{\beta})$, which is a result of the subtraction from the indicator's mean, $\bar{y}_i$, of the component due to the measurable attributes $x_i^T\hat{\beta}$. The coefficient of proportionality, g_i, depends on both the variability due to contextual factors, σ_v^2, and the agreement among indicators, σ_e^2. When there is total agreement among the indicators, then $\sigma_e^2 = 0$ and $g_i = 1$ and the whole residual value is estimated to be due to contextual factors. When there is a certain disagreement among indicators then $\sigma_e^2 > 0$ and only a proportion $0 < g_i < 1$ of the residual value is due to the contextual factors; the remainder is considered to be due to the disparity in the perception of β. When σ_v^2 approaches zero, the effect of contextual factors tends to zero.

Relative efficiency of the empirical best linear unbiased predictor estimator

The estimator in [4] can be written as $\hat{y}_i = g_i\bar{y}_i + (1 - g_i)x_i^T\hat{\beta}$, that is a weighted mean of the extreme estimators $\hat{y}_i(0) = x_i^T\hat{\beta}$ and $\hat{y}_i(1) = \bar{y}_i$, corresponding to $g_i = 0$ and to $g_i = 1$, respectively, with the weights being g_i and $(1 - g_i)$. The value $g_i = 0$ corresponds to the case where it is assumed that the variability of the property value due to unobserved contextual factors is null ($\sigma_v^2 = 0$), and the whole variability is due to measurable attributes and the value $g_i = 1$ corresponds to the case where it is assumed that the whole variability of the observed properties value is due to unobserved contextual factors.

The estimator $\hat{y}(0)$ is known as the synthetic regression estimator in statistics literature. When the hedonic assumption holds ($\sigma_v^2 \rightarrow 0$), then $\hat{y}_i(0)$ coincide with the usual hedonic estimator and this is why it is called hedonic regression estimator from now on. When the hedonic assumption holds, then the hedonic regression estimator approaches the EBLUP. However, when the hedonic assumption does not hold and the variability of the real properties value is mainly due to the contextual factors effect, then the EBLUP approaches the mean estimator, $\hat{y}_i(1) = \bar{y}_i$.

Results: A case study

A sample of $m = 38$ agricultural properties from *Castilla y León* (Spain) is considered, and a set of five indicators $\{y_{ij}; j = 1,2,3,4,5\}$ of the value of each agricultural property is given. The land of the agricultural properties in the sample is currently being used for barley crops or vineyards, either using unirrigated or irrigated systems. The attributes of value measured in each agricultural property are the size of the property (x_{1i}), the distance of the property both to the nearest road (x_{2i}) and to the nearest town (x_{3i}), and the net income per hectare (x_{4i}). The choice of attributes is inspired by the economic rational of the three usual approaches to the valuation problem as well as in the land valuation literature (Platinga and Miller, 2001). The net income (gross income minus expenses, without taking opportunity costs into account) is calculated assuming that the technical and economic conditions are those corresponding to the average farm in the area.

One of the five value indicators (i) is the valuation by an expert appraiser operating in the studied area and the other four were calculated using common methods: (ii) the so-called rental capitalization approach, defined as the present discounted value of the rate of rental, (iii) the so-called income capitalization approach, defined as the present value of the net incomes, (iv) a valuation based on the ratio between the valuation of the expert and the crop net income, which consists in multiplying the observed ratio in the set of properties with a same crop and under the same system (unirrigated or irrigated) by the yield of the crop observed in the property, and, finally, (v) a variant of this last method, consisting in using the mean of individual ratios, instead of the ratio of the set.

Table 1 shows the attributes of the average agricultural property in the sample. Agriculture in *Castilla y León* is similar to that of many regions and it is expected that the attributes collected in Table 1 will also be similar and that the empirical results arising from this article will also be of use in regions which are similar to the Spanish region mentioned above. Furthermore, in order to broaden the usefulness of results, the procedure suggested has been applied to a wide range of data simulated on the basis of the model adjusted to observed data.

The F-Fisher & Snedecor test for the difference between means of indicators of value across the same crop type under a similar system (unirrigated or irrigated) does not reject the equality of means hypothesis. This result supports the assumption that the indicators are unbiased.

Land use is considered as explanatory to the value of a property, and a dummy variable is associated with each observed category of land use observed in the sample. The significance of the eight variables considered as explanatory to value (four dummy and four attributes) is tested using a Wald test and the estimates $\hat{\beta}$ and $Var\hat{\beta}$ defined in the Appendix [A.1]. The net income per hectare and the dummy variables associated with unirrigated barley, irrigated vineyards and unirrigated vineyards are statistically significant at the 1% level, while the remaining four variables considered are not significant at the 5% level, possibly due to the

Table 1. Attributes of the average agricultural property

Land use	Sample mean and standard deviation of attributes			
	Surface (ha) $\bar{x}_1$ (S_1)	Distance to the road (km) $\bar{x}_2$ (S_2)	Distance to the town (km) $\bar{x}_3$ (S_3)	Net income (€ ha^{-1}) $\bar{x}_4$ (S_4)
Irrigated barley	11.71 (9.47)	0.84 (0.77)	5.31 (2.43)	128.50 (18.61)
Unirrigated barley	18.78 (10.40)	0.24 (0.21)	2.63 (0.89)	68.41 (8.16)
Irrigated vine	15.22 (15.69)	0.6 (0.56)	4.65 (1.20)	2,156.88 (278.10)
Unirrigated vine	4.67 (5.23)	0.91 (0.61)	5.36 (0.50)	1,900 (368.39)

Table 2. Estimates coefficients

Variable to explain	Dummy variables			Net income per hectare
	Unirrigated	Irrigated vineyard	Unirrigated vineyard	
Indicator	–3,107.93 (566.17) [5.48]	23,080 (835.02) [16.91]	14,116.83 (755.36) [30.56]	2.26 (0.11) [21.15]

(): standard error of the estimates coefficients. []: Wald statistics of the estimates coefficients.

low variability among properties in the sample, as shown in Table 1.

Table 2 shows the estimates of β and (in brackets) the estimates of the standard error of the estimators, together with [in square brackets] the Wald statistics results, for the significant variables.

Table 3 shows the estimates of the variance components.

The estimation of variability due to the contextual factors' effect is $\hat{\sigma}_v^2 = 894{,}951.25$ and the estimation of variability due to differences between indicators and difficulties involved in perceiving both β and v_i, as well as the measurement errors in the attributes is $\hat{\sigma}_e^2 = 6{,}606{,}256.9$. The «level of agreement» between indicators is estimated to be $\hat{\rho} = 1 - \hat{\sigma}_e^2/(\hat{\sigma}_v^2 + \hat{\sigma}_e^2) = 0.12$, so that there is a notable disagreement, $1 - \hat{\rho} = 0.88$, among the value indicators used.

The value of λ_{lm} observed when using [A.3] (see appendix) is 4.23 and since $\chi^2_{0.95}(1) = 3.84$, the null hypothesis, $\sigma_v^2 = 0$, is rejected at a 5% significance level. Hence, it is admitted that g_i differs from zero, so that the EBLUP differs from both the hedonic regression and the arithmetic mean of the estimators.

Relative efficiency between estimators

Table 4 shows the EBLUPs value per hectare of the sample properties, using expression [4], and the typical standard error of the EBLUP (square root of the estimated mean squared error of the estimator computed by using (A.4) (see appendix)], as well as the remaining estimators. The EBLUP is the most precise (the lowest standard error), while the worst estimator is the indicators' mean arithmetic.

The relative efficiency of the EBLUP with regard to the arithmetic mean ranges between 1.64 and 1.83 (e.g. the variance of the arithmetic mean is 64% to 83% higher than the mean square error of the EBLUP estimator), and with regard to the hedonic regression estimator, ranges between 1.17 and 1.47.

Table 3. Estimates of the variance components

Variance components	
$\hat{\sigma}_v^2$ (Standard error of $\hat{\sigma}_v^2$)	$\hat{\sigma}_e^2$ (Standar error of $\hat{\sigma}_e^2$)
894,951.25 (574,275.4)	6,606,256.9 (760,294.74)

Parameterization of the «degree of agreement» and the contextual factors effect

The relative efficiency of the EBLUP with regard to the hedonic regression and arithmetic mean estimators, depend on the ratio between the variance components σ_v^2 and σ_e^2. In order to study this dependency, both the «degree of agreement» among indicators, σ_e^2, and the variability due to contextual factors, σ_v^2, were parameterized, assigning ρ values between 0.1 and 0.9, with an increase of 0.1. The values of σ_v^2 y σ_e^2, corresponding to each of ρ's values have been calculated, with the restriction that the total variation should be the same as that observed in the sample.

A set of values $\{y_{ij};\, j = 1,2,3,4,5;\, i = 1,2,...,38\}$ for each parameterized value of ρ were simulated using model [3] and the data relative to value attributes observed in the sample. These simulations have been calculated adding the value of v_i, simulated from a normal variable of mean zero and variance σ_v^2 and the values $\{e_{ij};\, j = 1,2,3,4,5\}$, simulated from a normal variable of mean zero and variance σ_e^2 to $x_i^T\hat{\beta}$.

The relative efficiency of EBLUP with regard to both hedonic regression and arithmetic mean estimators calculated taking the simulated values is shown in Figure 1. In order to represent the graph, a box-and-whiskers graph has been associated with each of the values assigned to ρ. The box length is the inter-quar-

Table 4. Estimates based on EBLUP and relative efficiencies

Crop type	Property number	Estimation of value (€ ha^{-1})			Standard error (€ ha^{-1}) of the total estimations			Relative efficiency of EBLUP with regard to	
		Attributes' effect	Contextual effect	Total	EBLUP	Hedonic regression	Arithmetic mean	Hedonic regression	Arithmetic mean
Irrigated barley	1	13,569.61	–141.15	13,428.46	894.40	1,046.06	1,145.67	1.37	1.64
	2	13,569.61	–657.23	12,912.38	894.40	1,046.06	1,145.67	1.37	1.64
	3	11,308.01	44.17	11,352.19	869.06	1,034.43	1,145.67	1.42	1.74
	4	13,569.61	–1,003.96	12,565.66	894.40	1,046.06	1,145.67	1.37	1.64
	5	11,308.01	44.17	11,352.19	869.06	1,034.43	1,145.67	1.42	1.74
	6	9,046.41	856.63	9,903.04	847.75	1,024.82	1,145.67	1.46	1.83
Unirrigated barley	1	3,450.72	102.07	3,552.79	851.89	1,026.67	1,145.67	1.45	1.81
	2	4,807.68	–114.95	4,692.73	851.81	1,026.64	1,145.67	1.45	1.81
	3	4,807.68	–114.95	4,692.73	851.81	1,026.64	1,145.67	1.45	1.81
	4	5,712.32	–315.18	5,397.14	852.71	1,027.04	1,145.67	1.45	1.81
	5	4,129.20	21.57	4,150.77	851.63	1,026.56	1,145.67	1.45	1.81
	6	3,676.88	173.20	3,850.08	851.76	1,026.61	1,145.67	1.45	1.81
	7	4,807.68	–108.33	4,699.35	851.81	1,026.64	1,145.67	1.45	1.81
	8	2,546.08	363.01	2,909.09	852.89	1,027.13	1,145.67	1.45	1.80
	9	4,355.36	–72.18	4,283.18	851.65	1,026.57	1,145.67	1.45	1.81
	10	4,355.36	–72.18	4,283.18	851.65	1,026.57	1,145.67	1.45	1.81
	11	3,450.72	137.91	3,588.63	851.89	1,026.67	1,145.67	1.45	1.81
Irrigated vineyard	1	37,780.61	627.28	38,407.88	855.83	1,028.45	1,145.67	1.44	1.79
	2	38,911.41	621.84	39,533.25	856.54	1,028.77	1,145.67	1.44	1.79
	3	38,911.41	1,114.50	40,025.91	856.54	1,028.77	1,145.67	1.44	1.79
	4	36,649.81	–1,071.33	35,578.47	856.31	1,028.66	1,145.67	1.44	1.79
	5	38,911.41	–101.00	38,810.41	856.54	1,028.77	1,145.67	1.44	1.79
	6	38,911.41	872.21	39,783.62	856.54	1,028.77	1,145.67	1.44	1.79
	7	34,388.20	–1,678.22	32,709.98	860.80	1,030.69	1,145.67	1.43	1.77
	8	38,911.41	908.55	39,819.96	856.54	1,028.77	1,145.67	1.44	1.79
	9	38,911.41	262.44	39,173.85	856.54	1,028.77	1,145.67	1.44	1.79
	10	34,388.20	–1,556.27	32,831.93	860.80	1,030.69	1,145.67	1.43	1.77
Unirrigated vineyard	1	23,163.24	–185.17	22,978.08	858.35	1,029.58	1,145.67	1.44	1.78
	2	27,686.45	346.73	28,033.17	851.87	1,026.67	1,145.67	1.45	1.81
	3	25,424.85	–44.20	25,380.64	852.74	1,027.06	1,145.67	1.45	1.81
	4	26,555.65	–117.32	26,438.33	851.71	1,026.59	1,145.67	1.45	1.81
	5	27,686.45	104.03	27,790.48	851.87	1,026.67	1,145.67	1.45	1.81
	6	31,078.85	–164.51	30,914.34	859.48	1,030.09	1,145.67	1.44	1.78
	7	28,817.25	–66.41	28,750.84	853.23	1,027.28	1,145.67	1.45	1.80
	8	23,163.24	–185.17	22,978.08	858.35	1,029.58	1,145.67	1.44	1.78
	9	25,424.85	202.13	25,626.97	852.74	1,027.06	1,145.67	1.45	1.81
	10	28,817.25	–66.41	28,750.84	853.23	1,027.28	1,145.67	1.45	1.80
	11	28,817.25	176.29	28,993.54	853.23	1,027.28	1,145.67	1.45	1.80

tiles range of relative efficiency, the point within the box marks the median and the «whiskers» mark the maximum and minimum relative efficiencies. As expected, EBLUP is the most efficient (its relative efficiency is higher than 1) for every ρ value. When ρ is low (whether the reason for this is that the level of disagreement among indicators is high with regard to the variability due to contextual factors or that the aforementioned variability is low with regard to the disagreement), then EBLUP is notably more efficient than the arithmetic mean and only slightly more efficient than the hedonic regression estimator. On the contrary, when ρ is high (whether the reason is because the degree of agreement among indicators is high with respect to variability due to contextual factors or because this variability is high with regard to the level of

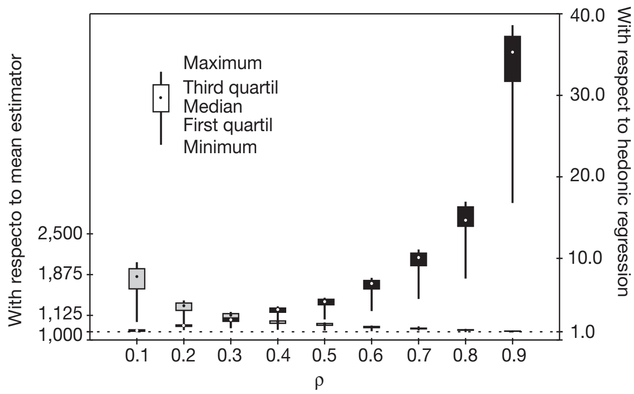

Figure 1. The relative efficiency of EBLUP with regard to both hedonic regression and mean arithmetic estimators.

disagreement), then EBLUP is much more efficient than the hedonic regression estimator and only slightly more efficient than the arithmetic mean. Hence, the EBLUP always protects against the inefficiencies of alternative estimators, especially in extreme cases. As approaches its central values, the relative efficiency of EBLUP decreases, although it always remains higher that of alternative estimators.

Discussion

There is a disagreement in the literature over the best method to measure the value of real estates, and this has traditionally been a serious obstacle, in particular for the valuation of agricultural properties. The methods proposed to date provide indicators of a property value, but not the standard error of these indicators and —as a result— there is no way to know which method is the most precise. To estimate the value of a given real property, the IVSC suggests using three different approaches: the sales comparison approach, the income capitalization approach and the cost approach. The problem is that the appraiser has to make a final reconciliation step to reach a final estimate of the market value and there is no existing methodology on how to combine the different market values calculated by the three approaches into one final market value.

The statistical approach proposed in this paper provides the best linear combination of value indicators provided by the methods of valuation proposed in the literature, including the hedonic approach, and is, thus, the best compromise among these methods. The consensus estimator proposed uses all available information optimally, including the information which value indicators contain on the effect of contextual factors. This furnishes it with a greater relative efficiency with regard to alternative estimators, such as the arithmetic mean of the indicators or the hedonic regression estimator. While the former solely uses the indicators' data —ignoring data relative to attributes— the latter uses data relative to attributes but does not use data relative to indicators and assumes that the effect of contextual factors is negligible.

A MIMCM model is estimated, using data on a sample of agricultural properties in *Castilla y León* (Spain), and the relative efficiency of the consensus estimator was assessed using this model. The empirical results confirm that the proposed estimator is more efficient than the alternative linear estimators. The inefficiency of these alternative estimators depend on both, the level of disagreement between indicators and the contextual factors' effect, which has been evaluated using simulation. When the disagreement among indicators is high with respect to the contextual factors' effect, then the proposed estimator is notably more efficient than the arithmetic mean of indicators, although only slightly more efficient than the hedonic regression estimator. However, when a high level of agreement among indicators is achieved or the contextual factors' effect is high with respect to the level of agreement among indicators, then the consensus estimator proposed is only slightly more efficient than the arithmetic mean of indicators, but much more efficient than the hedonic regression estimator.

Acknowledgement

Financial support from the Spanish Ministry for Education and Science (Project CGL2005-02589/CLI) is gratefully acknowledged. We wish to thank the editor-in-chief and the referees for their very helpful comments and suggestions.

References

AMBROSIO L., IGLESIAS L., 2000. Land cover estimation in small areas using ground survey and remote sensing. Remote Sens Environ 74, 240-248.

ASFMRA-AI, 2000. The appraisal of rural property. American Society of Farm Managers and Rural Appraisers (ASFMRA) and Appraisal Institute (AI). USA.

BAILEY M.J., MUTH R.F., NOURSE H.O., 1963. A regression method for real estate price index construction. J Am Stat Assoc 58, 933-942.

CORGEL J.B., LING D.C., SMITH H.C., 2001. Real estate perspectives. An introduction to real estate. McGraw-Hill Irwin, New York.

CROPPER M.L., DECK L.B., MCCONNELL K.E., 1988. On the choice of functional form for hedonic price functions. Rev Econ Stat 70, 668-675.
ENGLAND R.W., 2003. State and local impacts of a revenue-neutral shift from a uniform property to a land value tax: results of a simulation. Land Econ 79, 38-43.
FREEMAN A.M., 1979. Hedonic prices, property values and measuring environmental benefits: a survey of the issues. Scand J Econ 81, 154-173.
FULLER W.A., 1987. Measurement error models. Wiley, New York.
GHOSH M., RAO J.N.K., 1994. Small area estimation: an appraisal. Stat Sci 9, 55-93.
GOLDBERGER A.S., 1962. Best linear unbiased prediction in the generalized linear regression model. J Am Stat Assoc 57, 369-375.
GRISSOM T.V., 1985. Value definition: its place in the appraisal process. The Appraisal Journal 53, 217-225.
HADLEY A.T., 1928. The meaning of valuation. The American Economic Review 18, 173-180.
HARTER R.M., 1983. Small area estimation using nested-error models and auxiliary data. Doctoral Thesis. Iowa State University.
IVS, 2005. International valuation standards [online]. Available in http://www.ivsc.org/standards/index.html [27 November, 2006].
JUDGE G.G., GRIFFITHS W.E., HILL R.C., LÜTKEPOHL H., LEE TC., 1985. The theory and practice of econometrics. Wiley, New York.
KHURI A.I., SAHAI H., 1985. Variance components analysis: a selective literature survey. Int Stat Rev 53, 279-300.
LEMKE K.W., 1966. Asset valuation and income theory. The Accounting Review 41, 32-41.
LUKE D.A., 2004. Multilevel modelling. Series: Quantitative applications in the social sciences. Sage University Papers. Thousand Oaks, California.
MALPEZZI S., 2002. Hedonic pricing models and house price indexes: a selective and applied review. In: Housing economics and public policy (O'Sullivan A., Gibb K., eds). Blackwell, Oxford. pp. 67-89.
NELSON A.C., GENEREUX J., GENEREUX M., 1992. Price effects of landfills on house values. Land Econ 68(4), 359-365.
PLATINGA A.J., MILLER D.J., 2001. Agricultural land values and the value of rights to future land development. Land Econ 77, 56-67.
PRASAD N.G.N., RAO J.N.K., 1990. The estimation of the mean squared error of small-area estimators. J Am Stat Assoc 85, 163-171.
REICHERT A.K., SMALL M., MOHANTY S., 1991. The impact of landfills on residential property values. The Journal of Real Estate Research 7(3), 297-314.
ROBINS P.K., WEST R.W., 1977. Measurement errors in the estimation of home value. J Am Stat Assoc 72, 290-294.
ROBINSON G.K., 1991. That BLUP is a good thing: the estimation of random effects. Stat Sci 6, 15-51.
SHEPPARD S., 1999. Hedonic analysis of housing markets. In: Handbook of regional and urban economics, Volume 3: Applied urban economics (Cheshire P, Mills E.S., eds). Handbooks in Economics 7, North-Holland, Amsterdam. pp. 1595-1635.
SHILLER R.J., WEISS A.N., 1999. Evaluating real estate valuation systems. Journal of Real Estate Finance and Economics 18, 147-161.
SKRONDAL A., RABE-HESKETH S., 2004. Generalized latent variable modelling: multilevel, longitudinal and structural equation models. Chapman&Hall/CRC.

Appendix

The best linear unbiased predictor of y_i

The BLUP of y_i based on model [3] is taken from Goldberger (1962) and Robinson (1991) and has many applications (Ambrosio and Iglesias, 2000):

$$\hat{y}_i = x_i^T \hat{\beta} + \hat{v}_i \qquad \text{[A.1]}$$

where $\hat{\beta} = (X^T V^{-1} X)^{-1} (X^T V^{-1} Y)$ is the estimator of β, $X = [1x]$ where 1 is a column vector $(n \times 1)$ of ones and x the matrix $(n \times p)$ of the attribute values, $V^{-1} = diag\,(V_1^{-1}, V_2^{-1}, ..., V_i^{-1}, ... V_m^{-1})$ is a block diagonal matrix with $V_i^{-1} = \frac{1}{\sigma_e^2} I_{(n_i)} - \frac{g_i}{(n_i \sigma_e^2)} 1_{(n_i)} 1_{(n_i)}^T$, where $I_{(n_i)}$ is the identity matrix of order n_i and $1_{(n_i)}$ is a column vector $(n \times 1)$ of ones. The covariance matrix of $\hat{\beta}$ is $Var\ \hat{\beta} = (X^T V^{-1} X)^{-1}$.

$\hat{v}_i = g_i (\bar{y}_i - x_i^T \beta)$ is the BLUP of v_i (assuming that σ_v^2 and σ_e^2 are known), where $g_i = \sigma_v^2 / [\sigma_v^2 + (\sigma_e^2 / n_i)]$ and $\bar{y}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} y_{ij}$. Y is the column vector $(n \times 1)$ of the y_{ij} values in the sample. Substituting the BLUP in $\hat{v}_i$ in [A.1], y_i's BLUP becomes:

$$\hat{y}_i = x_i^T \hat{\beta} + g_i\,[\bar{y}_i - x_i^T \hat{\beta}] \qquad \text{[A.2]}$$

Testing model assumptions

In the specified model, the basic assumption is that y_i is not uniquely determined by the measurable attributes, x_i, but also by unobserved contextual factors, which makes y_i vary around its expected value, $x_i^T \beta$, with variance $\sigma_v^2 > 0$. The null hypothesis $\sigma_v^2 = 0$

versus the alternative hypothesis $\sigma_v^2 > 0$ can be tested using the statistic:

$$\lambda_{lm} = \frac{n}{2(\bar{n}-1)}\left[\frac{\sum_{i=1}^{m}\left(\sum_{j=1}^{n_i}\hat{u}_{ij}\right)^2}{\sum_{i=1}^{m}\sum_{j=1}^{n_i}\hat{u}_{ij}^2} - 1\right]^2 \qquad [A.3]$$

where $\hat{u}_{ij}$ is the residual of model [3] adjusted by ordinary least squares and taking $v_i = 0$; $n = \sum_{i=1}^{m} n_i$ and $\bar{n} = \frac{n}{m}$. This statistic is distributed asymptotically as χ^2 with one degree of freedom (Judge *et al.,* 1985). The rest of the cases are verified using classic tests.

Variance components estimator

In general, the variance components σ_v^2 and σ_e^2 are unknown. For their estimation, several procedures have been proposed (Khuri and Sahai, 1985). Here the Henderson method 3 is used (Prasad and Rao, 1990):

$$\hat{\sigma}_e^2 = \hat{e}^T\,\hat{e}/(n-m-1),$$
$$\hat{\sigma}_v^2 = [\hat{u}^T\,\hat{u} - (n-2)\,\hat{\sigma}_e^2]/n_*$$

where:

$$n_* = n - trace\left\{(X^T X)^{-1}\sum_{i=1}^{m} n_i^2\, x_i^T\, x_i\right\} =$$
$$= \sum_{i=1}^{m} n_i\left[1 - n_i x_i (X^T X)^{-1} x_i^T\right]$$

$\hat{e}^T\,\hat{e}$ is the sum of squares of residues of model [3] adjusted by ordinary least squares and taking v_i as fixed, i.e., the sum of squares of residues of the dummy variable model, and $\hat{u}^T\,\hat{u}$ is the sum of squares of the residues of model [3] adjusted by ordinary least squares and taking $v_i = 0$. The dummy variable model is $y_{ij} = \mu_i + \tilde{x}_i^T\tilde{\beta} + e_{ij}$, with $\mu_i = (\beta_1 + v_i)$ and $\tilde{x}_i^T$ and $\tilde{\beta}$ excluding the column of ones and the independent term, respectively. It is assumed that the independent term, μ_i, changes from one property to another. This assumption is specified by associating a variable (dummy), D_i, to each μ_i ($i = 1, 2, \ldots, m$) the value of which is one for every indicator of the i^{th} piece of land and zero for the remaining observations ($y_{ij} = \mu_1 D_1 + \mu_2 D_2 + \ldots + \mu_i D_i + \ldots + \mu_m D_m + \tilde{x}_i^T\,\tilde{\beta} + e_{ij}\ \forall i; i = 1,2,\ldots,m$ the variable D_i takes «n» values, the n_i values corresponding to the indicators of the i^{th} property are equal to 1 and the remaining $n - n_i$ values are equal to zero).

Replacing σ_v^2 and σ_e^2 by $\hat{\sigma}_v^2$ and $\hat{\sigma}_e^2$, the result is an estimator of the predictor $\hat{y}_i$, called the empirical best linear unbiased predictor (EBLUP) estimator. And replacing σ_v^2 and σ_e^2 by $\hat{\sigma}_v^2$ and $\hat{\sigma}_e^2$ in [A.1], the result is an estimator $\hat{V}$ of the variance and covariance matrix V.

Mean squared error of the BLUP *MSE*($\hat{y}_i$)

The mean squared error (MSE) of the estimator $\hat{y}_i\,(0)$ can be expressed as a function of the mean squared error of the BLUP (Harter, 1983), $MSE\,(\hat{y}_i)$:

$$MSE\left[\hat{y}_i(0)\right] = \qquad [A.4]$$
$$= MSE\,(\hat{y}_i) + g_i^2\left[\sigma_v^2 + \frac{\sigma_e^2}{n_i} - x_i\left(X^T V^{-1} X\right)^{-1} x_i^T\right]$$

These MSE can be estimated replacing $MSE(\hat{y}_i)$ by $\hat{MSE}(\hat{y}_i)$ (Prasad and Rao, 1990; Ghosh and Rao, 1994), and V by $\hat{V}$.

Relative efficiency

(i) With regard to the arithmetic mean estimator

The mean estimator, $\hat{y}_i(1) = \bar{y}_i = \sum_{j=1}^{n_i} y_{ij}/n_i$, does not make use of attributes data, only of indicators y_{ij}. The variance of this estimator is given by $V[\hat{y}_i(1)] = S_i^2/n_i$, where S_i^2 is the variance between indicators of a same property. Assuming that S_i^2 is the same in every property and equal to S_w^2, it can be estimated by $\hat{S}_w^2 = \sum_{i=1}^{m}(n_i-1)s_i^2/(n-m)$ where $s_i^2 = \sum_{j=1}^{n_i}(y_{ij}-\bar{y}_i)^2/(n_i-1)$.

An estimator, $\hat{V}[\hat{y}_i(1)]$, of $V[\hat{y}_i(1)]$, can be defined replacing S_i^2 by $\hat{S}_w^2$. The relative efficiency of the EBLUP with regard to the mean estimator would be estimated by $\hat{V}[\hat{y}_i(1)]/\hat{MSE}(\hat{y}_i)$.

(ii) With regard to the hedonic regression estimator

The relative efficiency of the EBLUP with regard to the hedonic regression estimator would be estimated by $\hat{M}SE[\hat{y}_i(0)]/\hat{M}SE(\hat{y}_i)$. If the hedonic assumption holds and the contextual effect is low in such a way that σ_v^2/σ_e^2 is small, then g_i tends to be low so that on the basis of [A.4] $\hat{M}SE[\hat{y}_i(0)]/\hat{M}SE(\hat{y}_i)$ tend towards 1 and the hedonic regression estimator tends to be as efficient as the EBLUP. When both the contextual effect and the agreement among indicators are high enough for σ_v^2/σ_e^2 to be high, then g_i is far from zero and the BLUP tends towards the estimator, $\hat{y}_i(1)$, so that $\hat{V}(\hat{y}_i)/\hat{M}SE(\hat{y}_i)$ tends towards 1 and the mean estimator tends to be as efficient as the EBLUP.

Estimation when indicators are not available

When the set of indicators is not available, then [4] is not feasible and it is proposed to use $x_i^T\hat{\beta}$ as estimator of y_i and to estimate its mean squared error using $M\hat{S}E(\hat{y}_i) = x_i(X^T\hat{V}^{-1}X)^{-1}x_i^T + \hat{\sigma}_v^2$.