Ann. Inst. Statist. Math.
Vol. 47, No. 3, 401–414 (1995)

REAL ESTATE PRICE PREDICTION UNDER ASYMMETRIC LOSS

MICHAEL CAIN[1] AND CHRISTIAN JANSSEN[2]

[1] *Department of Economics, University of Salford, The Crescent, Salford M5 4WT, U.K.*
[2] *Department of Finance and Management Science, University of Alberta, Edmonton, Canada T6G 2R6*

(Received January 5, 1994; revised December 12, 1994)

Abstract. This paper deals with the problem of how to adjust a predictive mean in a practical situation of prediction where there is asymmetry in the loss function. A standard linear model is considered for predicting the price of real estate using a normal-gamma conjugate prior for the parameters. The prior of a subject real estate agent is elicited but, for comparison, a diffuse prior is also considered. Three loss functions are used: asymmetric linear, asymmetric quadratic and LINEX, and the parameters under each of these postulated forms are elicited. Theoretical developments for prediction under each loss function in the presence of normal errors are presented and useful tables of adjustment factor values given. Predictions of the dependent price variable for two properties with differing characteristics are made under each loss function and the results compared.

Key words and phrases: Asymmetric loss, Bayesian prediction, real estate valuation.

1. Introduction

Valuation of real estate is a world-wide phenomenon. A wide range of properties are evaluated for a multitude of purposes. With respect to single family residences the predominant purposes are purchases and sales, tax assessment, expropriation, divorce, inheritance or estate settlement, and mortgaging. The valuations are performed by members of different professions—real estate agents, appraisers, assessors, mortgage lenders, and, not infrequently, by various specialists or consultants. Sometimes adjudication or arbitration by boards, tribunals, courts and other legal or administrative bodies is involved. A variety of techniques have been developed for carrying out the valuations, one of the principal approaches being the market comparison—essentially establishing the value of a property by comparisons with similar properties. Relatively little attention, however, has been paid to the loss associated with over or under-valuation.

Considered here is the situation of a real estate agent being engaged by a vendor to provide a market valuation of a house. In practical terms this involves prediction of the selling price with different losses being associated with different prediction errors. The loss from overestimation would generally be different from that of underestimation by the same amount, and its magnitude typically different from the monetary amount of the error.

The consequences of over and underestimation are discussed and an attempt is made to elicit the real estate agent's loss function. It is shown that if the loss function is asymmetric linear, or asymmetric quadratic, or of the LINEX type, and the distribution of prediction errors normal, the expected loss is minimised by additively adjusting the predictive mean—the adjustment being a fraction of the standard deviation, or variance, of the prediction errors. Although the paper deals with real estate, its applicability is much wider.

The paper is organised as follows. The theory concerning the adjustment term is developed and presented in Section 2. The particulars of the application including the listing process, and how losses occur, are set out in Section 3. The prediction model is given in Section 4, and the elicitation process for the prior information in Section 5. Regression results are presented in Section 6. The whole methodology is brought together in Section 7 to give specific predictions for two properties. Conclusions are in Section 8.

2. Theoretical developments

Let F be the predictive (cumulative) distribution function of the continuous random variable Y for which it is required to find the optimal prediction, h, taking into account the loss $g(y-h)$ associated with the predictive error. The loss function, g, which is not necessarily symmetric, can be written in the form (with $u = y - h$):

$$g(u) = \begin{cases} g_1(u), & u \geq 0 \\ g_2(u), & u < 0 \end{cases}$$

where it is assumed that $g_1(0) = 0 = g_2(0)$, $g_1'(u) > 0$ for $u > 0$ and $g_2'(u) < 0$ for $u < 0$. It is required to minimise the expected predictive loss $\int_{-\infty}^{\infty} g(y-h)dF(y)$ and the optimal prediction is a solution $h = h^*$ of

$$\int_h^{\infty} g_1'(y-h)dF(y) + \int_{-\infty}^{h} g_2'(y-h)dF(y) = 0. \tag{2.1}$$

Granger (1969) obtained a solution for prediction with an asymmetric linear loss (cost) function of the form

$$g(u) = \begin{cases} au, & u \geq 0 \\ -bu & u < 0 \end{cases}$$

where $a > 0$, $b > 0$, and gave conditions which essentially ensure that the predictive mean is optimal. The solution of (2.1) in this case is

$$h^* = F^{-1}\left(\frac{a}{a+b}\right). \tag{2.2}$$

If the loss function is symmetric ($a = b$) the predictive median, M, is optimal and the corresponding minimal mean loss is $aE(|Y - M|)$.

Cain (1991) considered the minimisation of the p-th ($0 < p < 1$) quantile of predictive loss (cost) and compared the solution with that of minimising mean loss. He took as one example an asymmetric quadratic loss function of the form:

$$g(u) = \begin{cases} au^2, & u \geq 0 \\ bu^2, & u < 0 \end{cases}$$

where $a > 0$, $b > 0$. The optimal (minimising expected loss) prediction then satisfies the particular case of (2.1):

$$a[E(Y) - h] + (b - a)\left\{\int_{-\infty}^{h} ydF(y) - hF(h)\right\} = 0 \tag{2.3}$$

and to proceed further more information about F is required; however, if the loss function is symmetric it follows from (2.3) that the predictive mean is optimal and the minimal expected predictive loss is thus $aV(Y)$.

Zellner (1986) considered prediction with a particular class of asymmetric loss functions introduced by Varian (1975) and obtained estimators and predictors which are optimal relative to Varian's (LINEX) loss function. The LINEX loss function:

$$g(u) = b[e^{-au} + au - 1] \tag{2.4}$$

where $a \neq 0$, $b > 0$, has a minimum at $u = 0$.

2.1 *Normal predictive distribution*

If the predictive distribution of Y is $N(\mu, \sigma^2)$, the optimal prediction, given by (2.2), with the asymmetric linear loss function is

$$h^* = \mu + \sigma\Phi^{-1}\left(\frac{a}{a+b}\right) \tag{2.5}$$

and the corresponding minimal expected predictive loss is

$$(a + b)\sigma\phi\left(\Phi^{-1}\left(\frac{a}{a+b}\right)\right). \tag{2.6}$$

Here, ϕ and Φ are, respectively, the standard normal probability density function and (cumulative) distribution function. Note that in (2.5) the optimal solution, h^*, is an additive adjustment to the predictive mean. Table 1 gives values of both the optimal adjustment factor $\delta^* = \delta^*(a, b) = \Phi^{-1}(\frac{a}{a+b})$ and, to facilitate the evaluation of (2.6), $l_1(\delta^*) = (a/b + 1)\phi(\delta^*)$ for a range of values of $a/b \geq 1$. The results for $a/b < 1$ can be deduced by observing that $\delta^*(b, a) = -\delta^*(a, b)$ for $a > 0$, $b > 0$.

Table 1. Adjustments for linear loss.

a/b	1	1.5	2	3	4	5	6	7	10	100
δ^*	0	0.253	0.431	0.674	0.842	0.967	1.068	1.150	1.335	2.334
$l_1(\delta^*)$	0.798	0.966	1.091	1.272	1.399	1.500	1.579	1.648	1.800	2.645

Table 2. Adjustments for quadratic loss.

a/b	1+	1.5	2	2.5	3	3.5	4	4.5
ϵ^*	0	0.162	0.276	0.364	0.436	0.497	0.549	0.595
$l_2(\epsilon^*)$	1	1.218	1.391	1.537	1.663	1.774	1.874	1.966
a/b	5	5.5	6	6.5	7	7.5	10	100
ϵ^*	0.636	0.673	0.707	0.737	0.766	0.792	0.902	1.721
$l_2(\epsilon^*)$	2.050	2.127	2.200	2.268	2.331	2.392	2.653	5.222

The optimal prediction with the asymmetric quadratic loss function ($a \neq b$) is now, from (2.3),

$$h^* = \mu + \sigma\epsilon^*$$

where the adjustment factor ϵ^* is a solution for ϵ of

$$\phi(\epsilon) - \epsilon\left(\frac{a}{a-b} - \Phi(\epsilon)\right) = 0. \tag{2.7}$$

Observe that $\epsilon^* < 0$ if $a < b$ and $\epsilon^* > 0$ if $a > b$. The corresponding minimal expected predictive loss is

$$(a-b)\sigma^2\phi(\epsilon^*)/\epsilon^*. \tag{2.8}$$

Equation (2.7) can be solved numerically; Table 2 gives values of both $\epsilon^* = \epsilon^*(a,b)$ and, in (2.8), $l_2(\epsilon^*) = (a/b-1)\phi(\epsilon^*)/\epsilon^*$ for a range of values of $a/b > 1$. Note that $\lim_{a/b\to 1+} l_2(\epsilon^*) = 1$. The results for $a/b < 1$ can be deduced by observing that $\epsilon^*(b,a) = -\epsilon^*(a,b)$ for $a,b > 0$; note that

$$\phi(\epsilon) - \epsilon\left(\frac{b}{b-a} - \Phi(\epsilon)\right) \equiv \phi(-\epsilon) + \epsilon\left(\frac{a}{a-b} - \Phi(-\epsilon)\right).$$

The optimal prediction with the LINEX loss function of (2.4) is

$$h^* = \mu - \frac{a\sigma^2}{2} \tag{2.9}$$

and the corresponding minimal expected predictive loss is $\frac{1}{2}ba^2\sigma^2$; see (5) of Varian (1975), (3.3) and (3.5) of Zellner (1986). Note that the adjustment to the predictive mean in this case is a multiple of the predictive variance and not the standard deviation as with the linear and quadratic asymmetric loss functions. In contrast, the minimal value of expected predictive loss is a multiple of the predictive standard deviation in the case of asymmetric linear loss and a multiple of the predictive variance with asymmetric quadratic loss or LINEX loss.

Thus, a real estate agent valuing a house, and being subject to losses given by one of the above forms, should make the corresponding adjustment to the predictive mean.

3. The application

The situation considered is that of a real estate agent providing an estimate of market value of a property for the owner. In a competitive urban market, where many houses in the neighbourhood are sold each year, there is generally enough data to permit comparisons with similar houses recently sold or presently for sale; a prerequisite for the market comparison approach to valuation.

3.1 *Consequences of over or underestimation of selling price*

An agent's responsibility to the principal under agency law is strict in most jurisdictions. The agent therefore needs to be aware of possible consequences for the property owner of errors in the determination of market value. If the valuation is in preparation for a future sale, under-estimation may lead to the owner losing money and over-estimation to market resistance. The house might not sell, or only sell after a protracted period. The economic consequences for the owner may well be different from the amount of the error, and if serious there might be a case for compensation. In some instances the issue of professional negligence could arise. The loss to the agent may then exceed the amount by which the market value was incorrectly estimated.

3.2 *The loss function*

Conceptually we think of each house in a neighbourhood as having a certain market value. The selling price is an estimate of the unknown market value and, strictly speaking, it may not be possible to determine if a house sold above or below its market value. For the purpose of this paper we will focus on the prediction error—the difference between the observed (Y) and predicted ($\hat{Y}$) selling prices. Underestimation is then associated with positive errors and overestimation with negative ones.

The determination of the precise form of the loss function requires substantial information about the likelihood of a vendor seeking damages, the amount sought, and awarded, legal costs, and the likelihood of various settlements. However, extensive discussions with a real estate agent provided typical expected losses for a series of assumed cases, and it transpired that for a fairly broad range of errors the loss could be adequately represented by a general linear, or a general quadratic, function with different parameters chosen for over and underestimation to incorporate possible asymmetry. Varian (1975) arrived at a similar conclusion with respect to the LINEX loss function, even if in his application (property tax assessment) the consequences of prediction errors are more closely defined.

Table 3. Definition of variables.

X_1	Intercept
X_2	Age in years
X_3	Floor area in m^2
X_4	Number of car spaces in the garage (No garage = 0)
X_5	Garage attached to house (Yes = 1, No = 0)
X_6	Basement development (0 = None to 3 = Complete)
X_7	Number of woodburning fireplaces
X_8	Month of sale
X_9	Located in the Aspen neighbourhood (Yes = 1, No = 0)
X_{10}	Driving time from neighbourhood to central business district (cbd)
Y	Selling price ($)

Table 4. Sample statistics.

Variable	Mean	Smallest observation	Largest observation	Standard deviation
X_1	1	1	1	0
X_2	19.6	3	32	5.6
X_3	121	86.5	179.1	17.8
X_4	1.8	0	4	0.60
X_5	0.29	0	1	0.46
X_6	2.5	0	3	0.72
X_7	0.63	0	3	0.74
X_8	12.7	6	18	3.7
X_9	0.06	0	1	0.24
X_{10}	20.3	15	25	2.6
Y	109954	74000	172000	16081

3.3 *The example*

The approach is illustrated by application to a sample of 133 single family homes sold through a multiple listing system during a one-year period. Each house is described by some forty characteristics in the listing information. A subset of these was selected in cooperation with a real estate company active in the market being studied. The features chosen were those that the agents regarded as important and would generally use for the purpose of estimating the market value of a home.

To obtain a large enough sample it was necessary to include sales over a certain period—in this case a year. Consideration was given to expanding the study area, but other factors weighed against this, in particular issues pertaining to basic comparability of neighbourhoods. At present the neighbourhoods are

Table 5. The $\boldsymbol{X}'\boldsymbol{X}$ data matrix.

133.0	2608.0	16038.3	236.0	39.0	339.0	84.0	1683.0	8.0	2698.0
2608.0	55248.0	306586.2	4571.0	618.0	6820.0	1538.0	33202.0	190.0	51734.0
16038.3	306586.2	1975886.7	28707.2	5157.1	40496.2	10504.5	202553.3	902.4	327431.9
236.0	4571.0	28707.2	466.0	68.0	611.0	155.0	3004.0	13.0	4797.0
39.0	618.0	5157.1	68.0	39.0	90.0	29.0	496.0	2.0	808.0
339.0	6820.0	40496.2	611.0	90.0	933.0	221.0	4301.0	18.0	6816.0
84.0	1538.0	10504.5	155.0	29.0	221.0	126.0	978.0	7.0	1737.0
1683.0	33202.0	202553.3	3004.0	496.0	4301.0	978.0	23061.0	116.0	34127.0
8.0	190.4	902.4	13.0	2.0	18.0	7.0	116.0	8.0	168.0
2698.0	51734.0	327431.9	4797.0	808.0	6816.0	1737.0	34127.0	168.0	55602.2

quite comparable with respect to schools, parks, shopping facilities, street layout, traffic patterns, bus service, and general access to transportation corridors. The houses are all one-storeyed of the same type.

One issue that arose was how to account for the passage of time. General economic conditions changed only slightly over the study period, but the market for resale homes did experience some inflationary demand pressure. This made it important to keep track of when during the year a property was for sale and sold. This was done by including a 'time' variable, being the number of months from the beginning of the year when sold. This way the effect of time was accounted for inside the model. Applying various deflators to the selling price appeared to be less effective.

The regression model for selling price included nine explanatory variables: age, floor area, garage size, whether garage is attached or detached, basement development, fireplace, month, distance to the city centre, and whether located in the Aspen neighbourhood—a somewhat exclusive area due to natural surroundings like wooded ravines. These variables are fairly typical of those commonly used in studies of hedonic price estimation; see e.g. Wang *et al.* (1991) and Varian (1975). The variable definitions are given in Table 3, sample statistics are given in Table 4, and the $\boldsymbol{X}'\boldsymbol{X}$ data matrix in Table 5.

4. The model

The $n \times 1$ vector $\boldsymbol{Y} = (Y_1, Y_2, \ldots, Y_n)'$ of prices is assumed to have a multivariate normal distribution with mean vector $\boldsymbol{X\beta}$ and variance-covariance matrix $\sigma^2\boldsymbol{I}$, where $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_k)$ is a $k \times 1$ vector of unknown parameters ($k < m$), σ^2 is an unknown scalar and $\boldsymbol{I}$ is the $n \times n$ identity matrix. The $n \times k$ design matrix $\boldsymbol{X}$ of observed characteristics has full column rank and includes a column of ones to account for an intercept. A realisation of $\boldsymbol{Y}$ is denoted by $\boldsymbol{y}$.

A further value of price, Y_{n+1}, corresponding to a vector of characteristics $\boldsymbol{x} = (X_1, X_2, \ldots, X_k)'$ is assumed to have a normal distribution with mean $\boldsymbol{x}'\boldsymbol{\beta}$, variance σ^2 and Y_{n+1} independent of $\boldsymbol{Y}$ (given $\boldsymbol{\beta}$ and σ^2).

A conjugate normal-gamma prior distribution for $\boldsymbol{\beta}$ and σ^{-2} is adopted:

$$\boldsymbol{\beta} \mid \sigma^2 \sim \mathrm{MVN}(\boldsymbol{m}_0, \sigma^2 \boldsymbol{D}_0), \qquad \sigma^{-2} \sim \gamma\left(\frac{1}{2}d_0, \frac{1}{2}g_0\right) \tag{4.1}$$

where the prior parameters $\boldsymbol{m}_0, \boldsymbol{D}_0, d_0 > 0$ and $g_0 > 0$ are assumed to have been elicited and are thus known; with $\boldsymbol{D}_0$ positive definite. Interest centres on the prediction of Y_{n+1} given $\boldsymbol{y}$.

The predictive distribution of Y_{n+1} with a diffuse prior was derived by Zellner and Chetty (1965) and is given by Zellner (1971) as univariate-t with $n-k$ degrees of freedom, mean $\boldsymbol{x}'\boldsymbol{b}$ and variance

$$v_0 = v_0(\boldsymbol{x}) = \left(\frac{n-k}{n-k-2}\right) s^2[1 + \boldsymbol{x}'(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{x}], \tag{4.2}$$

where $\boldsymbol{b}$, s^2 are the ordinary least squares (OLS) estimates of $\boldsymbol{\beta}$ and σ^2, respectively. This corresponds to the classical result that

$$\frac{Y_{n+1} - \boldsymbol{x}'\hat{\boldsymbol{\beta}}}{\hat{\sigma}[1 + \boldsymbol{x}'(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{x}]^{1/2}} \sim t_{n-k}$$

where $\hat{\boldsymbol{\beta}}$, $\hat{\sigma}^2$ are the OLS estimators of $\boldsymbol{\beta}$, σ^2, with realisations $\boldsymbol{b}$, s^2. The predictive distribution with an informative prior of the form (4.1) is given by

$$[(g/d)(1 + \boldsymbol{x}'\boldsymbol{D}\boldsymbol{x})]^{-1/2}(Y_{n+1} - \boldsymbol{x}'\boldsymbol{m}) \mid \boldsymbol{y} \sim t_d, \tag{4.3}$$

a Student's t with d degrees of freedom; see Broemeling (1985) and Cain and Owen (1990). Here,

$$\begin{aligned}
\boldsymbol{D} &= (\boldsymbol{D}_0^{-1} + \boldsymbol{X}'\boldsymbol{X})^{-1}, \\
\boldsymbol{m} &= \boldsymbol{m}_0 + \boldsymbol{D}\boldsymbol{X}'\boldsymbol{X}(\boldsymbol{b} - \boldsymbol{m}_0), \\
d &= d_0 + n, \\
g &= g_0 + (n-k)s^2 + (\boldsymbol{b} - \boldsymbol{m}_0)'\boldsymbol{X}'\boldsymbol{X}\boldsymbol{D}\boldsymbol{D}_0^{-1}(\boldsymbol{b} - \boldsymbol{m}_0),
\end{aligned}$$

and $\boldsymbol{b}$, s^2 are the OLS estimates

$$\boldsymbol{b} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y}, \qquad s^2 = (\boldsymbol{y}'\boldsymbol{y} - \boldsymbol{b}'\boldsymbol{X}'\boldsymbol{y})/(n-k).$$

Note that (4.3) implies that the predictive mean of Y_{n+1} is $\boldsymbol{x}'\boldsymbol{m}$ and the predictive variance is

$$v = v(\boldsymbol{x}) = \frac{g}{(d-2)}[1 + \boldsymbol{x}'\boldsymbol{D}\boldsymbol{x}]. \tag{4.4}$$

The lack of prior knowledge about $\boldsymbol{\beta}$ and σ^2 can be expressed by a diffuse prior in which $\boldsymbol{\beta}$ and $\ln \sigma^2$ are independent and uniform; see Jeffreys (1961) and Savage (1962). This corresponds to the choice of (improper) prior parameters $d_0 = -k$, $g_0 = 0$, $\boldsymbol{D}_0^{-1} = \boldsymbol{0}$ and non-specific $\boldsymbol{m}_0$; cf. the previously given result of Zellner (1971).

5. The elicitation process

It is necessary to elicit the parameters of both the loss function and the prior distribution, and this is considered below.

5.1 *Eliciting the loss function*

Detailed discussions with a real-estate agent were held with a view to eliciting the parameters of the loss function in relation to each of the postulated forms: asymmetric linear, asymmetric quadratic, LINEX (see Varian (1975), Zellner (1986)). The values obtained were derived by OLS fitting to a series of suggested points on the loss function curve and are given below.

Asymmetric Linear. With an asymmetric linear loss function of the form:

$$g(u) = \begin{cases} au & \text{if } u \geq 0 \\ -bu & \text{if } u < 0 \end{cases}$$

the elicited value of a was 0.993, b was 1.465 and hence $b/a = 1.475$.

Asymmetric Quadratic. With an asymmetric quadratic loss function of the form:

$$g(u) = \begin{cases} au^2 & \text{if } u \geq 0 \\ bu^2 & \text{if } u < 0 \end{cases}$$

the elicited value of a was 0.0000483, b was 0.0000696 and hence $b/a = 1.441$.

LINEX. With a LINEX loss function of the form:

$$g(u) = b(e^{-au} + au - 1)$$

the elicited value of a was 0.0000212 and b was 258500.

The above values of a and b identify for each loss function the adjustment factor from Tables 1, 2 and equation (2.9) to be applied to the predictive mean, μ, in the minimal expected loss prediction of Y.

5.2 *Eliciting the prior parameters*

The elicitation of the prior parameters in (4.1) was more difficult and it involved a lengthy process of discussion and interchange of ideas between the authors and the subject real estate agent. The parameters of the (gamma) distribution for σ^{-2} were obtained by eliciting the median and possible range of σ^2, converting these to statements about σ^{-2} and obtaining the standard deviation of σ^{-2} by dividing its range by 5; the approximate length of a 95% HPD interval being 5 standard deviations for a gamma distribution. This produced $d_0 = 8$ and $g_0 = 4 \times 10^8$. Next, $\boldsymbol{m}_0$ and the associated standard deviations were elicited. This was accomplished by first obtaining mean values and ranges; and then dividing each range by 4 to obtain the standard deviation, as would be reasonable for a normal distribution.

The matrix $\boldsymbol{D}_0$ was elicited by first eliciting a correlation matrix for $\boldsymbol{\beta}$, converting this to a covariance matrix by making use of the previously elicited standard deviations, and then dividing by the already elicited median of σ^2. The real estate agent could not directly contemplate covariances. Even the correlations were quite difficult to obtain and this part took a considerable length of time. The elicitation of the correlation matrix was accomplished by considering a number of 2×2 tables of joint probabilities (each variable being either above or below its mean value) but in a few cases it was felt necessary to look at a 3×3 or 4×4 table to obtain further insight. Whilst this was very interesting, it was also very demanding and consequently subject to imprecision; nevertheless it was felt that the results obtained adequately reflected the beliefs of the subject real estate agent. For comparison, a diffuse prior was also considered.

The above process produced

$$\boldsymbol{m}_0 = (50000 \ \ -1000 \ \ 500 \ \ 5000 \ \ 5000 \ \ 3000 \ \ 3000 \ \ 500 \ \ 10000 \ \ -1000)'$$

with the corresponding vector of standard deviations

$$\boldsymbol{s} = (10000 \ \ 300 \ \ 50 \ \ 1000 \ \ 1000 \ \ 500 \ \ 500 \ \ 200 \ \ 1500 \ \ 250)'$$

and a correlation matrix for $\boldsymbol{\beta} \mid \sigma^2$ of

$$\mathrm{Corr}(\boldsymbol{\beta} \mid \sigma^2) = \begin{bmatrix} 1.0 & -0.2 & -0.6 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.2 \\ -0.2 & 1.0 & -0.8 & 0.0 & -0.6 & 0.2 & 0.0 & 0.0 & 0.0 & -0.6 \\ -0.6 & -0.8 & 1.0 & 0.2 & 0.2 & -0.2 & 0.2 & 0.0 & 0.0 & 0.4 \\ 0.0 & 0.0 & 0.2 & 1.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & -0.6 & 0.2 & 0.0 & 1.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.2 \\ 0.0 & 0.2 & -0.2 & 0.0 & 0.0 & 1.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.2 & 0.0 & 0.0 & 0.0 & 1.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 1.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 1.0 & 0.0 \\ 0.2 & -0.6 & 0.4 & 0.0 & 0.2 & 0.0 & 0.0 & 0.0 & 0.0 & 1.0 \end{bmatrix}$$

The evaluation of $\boldsymbol{D}_0$ was completed as

$$\boldsymbol{D}_0 = \frac{d_0}{g_0} \, \mathrm{Diag}(\boldsymbol{s}) \, \mathrm{Corr}(\boldsymbol{\beta} \mid \sigma^2) \, \mathrm{Diag}(\boldsymbol{s})$$

and yielded

$$D_0 = \begin{bmatrix} 2.00000 & -0.01200 & -0.00600 & 0.00000 & 0.00000 & 0.00000 & 0.00000 & 0.00000 & 0.00000 & 0.01000 \\ -0.01200 & 0.00180 & -0.00024 & 0.00000 & -0.00360 & 0.00060 & 0.00000 & 0.00000 & 0.00000 & -0.00090 \\ -0.00600 & -0.00024 & 0.00005 & 0.00020 & 0.00020 & -0.00010 & 0.00010 & 0.00000 & 0.00000 & 0.00010 \\ 0.00000 & 0.00000 & 0.00020 & 0.02000 & 0.00000 & 0.00000 & 0.00000 & 0.00000 & 0.00000 & 0.00000 \\ 0.00000 & -0.00360 & 0.00020 & 0.00000 & 0.02000 & 0.00000 & 0.00000 & 0.00000 & 0.00000 & 0.00100 \\ 0.00000 & 0.00060 & -0.00010 & 0.00000 & 0.00000 & 0.00500 & 0.00000 & 0.00000 & 0.00000 & 0.00000 \\ 0.00000 & 0.00000 & 0.00010 & 0.00000 & 0.00000 & 0.00000 & 0.00500 & 0.00000 & 0.00000 & 0.00000 \\ 0.00000 & 0.00000 & 0.00000 & 0.00000 & 0.00000 & 0.00000 & 0.00000 & 0.00080 & 0.00000 & 0.00000 \\ 0.00000 & 0.00000 & 0.00000 & 0.00000 & 0.00000 & 0.00000 & 0.00000 & 0.00000 & 0.04500 & 0.00000 \\ 0.01000 & -0.00090 & 0.00010 & 0.00000 & 0.00100 & 0.00000 & 0.00000 & 0.00000 & 0.00000 & 0.00125 \end{bmatrix}$$

6. The regression results

The OLS regression results are as follows:

$$n = 133, \quad k = 10, \quad s = 8498,$$
$$\boldsymbol{b} = (56147 \; -960 \; 472.5 \; 4923 \; 5282 \; 4338 \; 2690 \; 859 \; 11985 \; -936)'.$$

The coefficient estimates were all of the appropriate magnitude, had the sign expected, and generally made sense. They were also viewed as reasonable by the real estate agent participating in the study. Some seventy four percent of the variation in selling prices was explained by the regression. Detailed results are given in Table 6.

Table 6. Regression results.

Variable	Coefficient	Standard error	Tolerance	Contribution to R^2
X_1	56147	14179	—	—
X_2	−960	227	.34	.038
X_3	472.5	55	.58	.159
X_4	4923	1307	.89	.030
X_5	5282	1934	.70	.016
X_6	4338	1157	.78	.030
X_7	2690	1093	.83	.013
X_8	859	213	.90	.034
X_9	11985	3406	.83	.026
X_{10}	−936	392	.54	.012

$R^2 = 0.740$, Adjusted $R^2 = 0.721$.

Note 1. 'Tolerance' (a measure of collinearity) is the proportion of the variation in a particular explanatory variable unexplained by regressing the variable on all the other explanatory variables.

2. The 'contribution to R^2' is the amount by which R^2 would be reduced were that variable to be removed from the equation.

6.1 *The error distribution*

Residual analysis and diagnostic plots were used to check for model assumption violations, and normality of the error distribution. A histogram of residual errors with a normal distribution superimposed was drawn. The general shape conformed quite well to normality, a fact not contradicated by the value of the Kolmogorov-Smirnov statistic (with the parameters estimated from the sample). The sample parameters of the error distribution were: mean 2.4 (0 without the rounding errors), standard deviation 8203, skewness 0.09 and kurtosis 3.01.

From the point of view of the real estate agent, some of the large positive and negative errors would be worrisome. On average, the predictive ability of the model is quite good with mean absolute error 6356 and mean absolute percentage error under six percent. But the larger errors indicate that the model may seriously under or overestimate the selling price of an individual house. This could then lead to the type of losses discussed earlier. It was decided to refrain from purifying the sample through outlier analysis in order to maintain realism, and investigate the loss to which real estate agents are potentially subject. After all they must on occasion encounter properties that are not very similar to the "typical" house.

7. The prediction

Thus, given the predictive distribution of selling price and the loss functions associated with over and underestimation, the problem is to modify the predictive mean in order to minimise the expected loss. The earlier theoretical developments suggest, for an asymmetric quadratic loss function, an additive adjustment of $\sigma\epsilon^*$ where ϵ^* is a solution of (2.7) and is given in Table 2 for a variety of values of $a/b > 1$; and, for asymmetric linear loss, an additive adjustment given by (2.5) with values given in Table 1. Similarly, the adjustment for LINEX loss is given by (2.9). The standard deviation, σ, corresponds to that of the predictive distribution of Y, given the particular values of the explanatory variables and the previous data. Thus

$$\sigma = \left\{ \frac{g}{d-2}[1 + \boldsymbol{x}'\boldsymbol{D}\boldsymbol{x}] \right\}^{1/2}$$

as given by (4.4). Since the number of degrees of freedom is large, $d = 141$ with an informative prior and $n - k = 123$ with a diffuse prior, we regard the predictive distribution of Y to be normal rather than t.

The value of Y was predicted for each of the following two vectors of values of explanatory variables:

$$\boldsymbol{x}_1 = (1, 20, 115, 2, 0, 3, 1, 15, 0, 21)',$$
$$\boldsymbol{x}_2 = (1, 10, 100, 0, 0, 0, 0, 8, 0, 18)',$$

$\boldsymbol{x}_1$ being quite close to the vector of sample means and $\boldsymbol{x}_2$ somewhat extreme.

With the informative prior distribution, the predictive means and variances of Y for the two chosen $\boldsymbol{x}$ vectors, are [see (4.3) and (4.4)]:

$$\boldsymbol{x}_1'\boldsymbol{m} = 111195, \qquad v(\boldsymbol{x}_1) = 69784032,$$
$$\boldsymbol{x}_2'\boldsymbol{m} = 86876, \qquad v(\boldsymbol{x}_2) = 76252072.$$

The adjustments to the predictive mean for each of the three loss functions are given in Table 7, and the predicted values of Y in Table 8. Here, and subsequently, the number of non-zero digits displayed is not indicative of the accuracy of the results.

With a diffuse prior the predictive means and variances are [see (4.2)]:

$$\boldsymbol{x}_1'\boldsymbol{b} = 110074, \qquad v_0(\boldsymbol{x}_1) = 75038763,$$
$$\boldsymbol{x}_2'\boldsymbol{b} = 83829, \qquad v_0(\boldsymbol{x}_2) = 96340922.$$

Table 7. Additive adjustments with informative prior.

	Asymmetric linear	Asymmetric quadratic	LINEX
x_1	−2030	−1211	−740
x_2	−2122	−1266	−808

Table 8. Predicted Y values with informative prior.

	Asymmetric linear	Asymmetric quadratic	LINEX
x_1	109165	109984	110455
x_2	84754	85610	86068

Table 9. Additive adjustments with diffuse prior.

	Asymmetric linear	Asymmetric quadratic	LINEX
x_1	−2105	−1256	−795
x_2	−2385	−1423	−1021

Table 10. Predicted Y values with diffuse prior.

	Asymmetric linear	Asymmetric quadratic	LINEX
x_1	107969	108818	109279
x_2	81444	82406	82808

The corresponding adjustments and predicted values are given in Tables 9 and 10, respectively.

Note that Tables 1 and 2 do not give the precise values of the adjustment factors δ^* (for $a/b = 1.475$) and ϵ^* (for $a/b = 1.441$); and hence these have been evaluated separately as $\delta^* = 0.243$, $\epsilon^* = 0.145$ with $l_1(\delta^*) = 0.959$ and $l_2(\epsilon^*) = 1.195$. The minimal expected predictive loss (risk) is thus $(0.959)a\sigma$ with asymmetric linear loss, $(1.195)a\sigma^2$ with asymmetric quadratic loss and $\frac{1}{2}ba^2\sigma^2$ with LINEX loss; where σ^2 is the relevant predictive variance. With the unadjusted predictive mean, the corresponding expected losses are $(0.987)a\sigma$, $(1.221)a\sigma^2$ and $b(\exp(\frac{1}{2}a^2\sigma^2) - 1)$; approximately 3%, 2% and 1%, respectively, higher than the corresponding minimum. Observe also that the adjustment term in each case is influenced by the predictive variance and not the mean, and the adjustment factor only by the loss function parameters.

The results were discussed with the real estate agent and the adjustments obtained under each fitted loss function were regarded as eminently plausible in the present context; although those for asymmetric quadratic loss and for LINEX loss seemed somewhat small in percentage terms. It is clear that the precise form of the loss function needs careful consideration. The fact that the adjustment terms for each of the three loss functions are not very different with the informative as compared with a diffuse prior, suggests that they are fairly robust with respect to changes in the parameters of the prior distribution. Likewise, the overall predictions are fairly insensitive to changes in the prior parameters.

8. Conclusion

In this paper an integrated methodology is presented for determining the optimal prediction of the response variable in a standard linear model with a conjugate normal-gamma prior distribution for the parameters. Various types of asymmetry of the loss function are considered. Theoretical developments are presented, deriving adjustments to the predictive mean in order to minimise expected loss in the presence of the asymmetry. Tables of values of adjustments are given, to be used in predicting the response. In the process, an intermediate step is the elicitation of the parameters of both the loss function and prior distribution of an expert decision maker, in this case a real estate agent active in the market. Although the particular application is to real estate valuation the methodology is applicable to a wide variety of problems of prediction.

References

Broemeling, L. D. (1985). *Bayesian Analysis of Linear Models*, Marcel Dekker, New York.

Cain, M. (1991). Minimal quantile measures of predictive cost with a generalized cost of error function, Economic Research Papers, No. 40, University of Wales, Aberystwyth.

Cain, M. and Owen, R. J. (1990). Regressor levels for Bayesian predictive response, *J. Amer. Statist. Assoc.*, **85**, 228–231.

Granger, C. W. J. (1969). Prediction with a generalized cost of error function, *Operational Research Quarterly*, **20**, 199–207.

Jeffreys, H. (1961). *Theory of Probability*, Clarendon Press, Oxford, U.K.

Savage, L. J. (1962). *Bayesian Statistics: In Decision and Information Processes*, Macmillan, New York.

Varian, H. R. (1975). A Bayesian approach to real estate assessment, *Studies in Bayesian Econometrics and Statistics in Honor of Leonard J. Savage* (eds. S. E. Fienberg and A. Zellner), 195–208, North-Holland, Amsterdam.

Wang, K., Grisson, T. V., Webb, J. R. and Spellman, L. (1991). The impact of rental properties on the value of single-family residences, *Journal of Urban Economics*, **30**, 152–166.

Zellner, A. (1971). *An Introduction to Bayesian Inference in Econometrics*, Wiley, New York.

Zellner, A. (1986). Bayesian estimation and prediction using asymmetric loss functions, *J. Amer. Statist. Assoc.*, **81** (394), 446–451.

Zellner, A. and Chetty, V. K. (1965). Prediction and decision problems in regression models from the Bayesian point of view, *J. Amer. Statist. Assoc.*, **60**, 608–616.