Quality, Information and Wine Labelling

Experiences from the British Wine Market

Bodo STEINER

Cahiers d'économie et sociologie rurales, n° 60-61, 2001

Bodo STEINER*

Qualité, information et étiquetage du vin : les enseignements du marché britannique

Résumé – Cet article utilise la méthode des prix hédonistes pour identifier les valeurs que les consommateurs donnent aux informations fournies par les étiquettes des bouteilles sur les caractéristiques du vin qu'elles contiennent. L'analyse vise à mettre en évidence l'effet des différentes combinaisons possibles entre ces attributs. L'étude porte en particulier sur les cas où des variétés identiques de raisin proviennent de régions ou de pays d'origine différents. Les résultats économétriques montrent que la variété de raisin est d'une grande importance dans le choix des vins italiens et australiens, alors que, pour les vins français, c'est la provenance régionale qui constitue le critère de choix le plus important. On étudie également les implications commerciales des informations données par les étiquettes de bouteilles de vin pour les détaillants. L'étude des combinaisons possibles d'attributs peut en effet aider les négociants à gérer les stocks de vins, en déterminant les coûts imputables au remplacement d'une catégorie de vin par une autre.

Mots-clés:
vin, méthode des prix hédonistes, information et qualité, Grande-Bretagne

Quality, information and wine labelling. Experiences from the British wine market

Summary – *This paper applies hedonic price analysis to identifying the values which consumers and marketers place on labelling attributes. We analyse data from British wine retailers. Interaction terms are used to highlight the differential effects between attributes, and where these are found to be relevant, consumers are considered to regard attribute bundles as imperfect substitutes. Results suggest that grape varieties are highly valued in the case of Italian and Australian wines, whereas the ability and willingness of consumers to differentiate between regional specificities is most prevalent in the case of France. In accounting for the relative significance of both grape varieties and regional origins, the results suggest an asymmetry between Australia, possibly the most classical 'New World' wine producer, and France, the most classical of 'Old World' wine producers. Econometric results suggest that consumers attach great value to information concerning the originating retailer. Marketing implications are also investigated. By considering interactions between attributes, we demonstrate the usefulness of studying the valuation on the attribute level and identify the cost changes that a wine marketer would face if a stock-transfer of particular attribute combinations were intended.*

Key-words:
wine, hedonic price method, information and product quality, Great Britain

* *University of Kiel, Department of Agricultural Economics, Olshausenstr.40, 24118 Kiel, Germany*
e-mail: bsteiner@agric-econ.uni-kiel.de

I would like to thank Ian McAvinchey, the co-editors and two anonymous referees for their helpful comments.

ECONOMISTS may observe choices made by consumers and may also observe demand, but they do not observe preferences. For a given product, the analyst could, therefore, attempt to reconstruct the hierarchy of consumer attitudes towards product attributes, within his or her utility structure, starting from the behavioural intentions stated by consumers. But instead of relying on stated intentions, as do multi-attribute attitude models and conjoint analyses, we could use revealed preference analyses, which produce predictions by combining observations of realised choices with assumptions about underlying decision processes (Rosen, 1974; McFadden, 1974). When observing consumer choices of heterogeneous bundles of attributes in a retail market for wine, we could rely on hedonic price analysis to explore the implicit valuation of market participants concerning the components of heterogeneous attribute bundles.

In relying on observed choices, Frederick V. Waugh (1928) pioneered the development of hedonic price analysis in agricultural economics. His analysis of vegetable prices is based on the hypothesis that vegetable quality is related to measurable specification variables. In a study on automobile demand, Court (1939) essentially incorporated the hedonic hypothesis that heterogeneous goods are aggregations of attributes (in today's Gorman (1980)-Lancaster (1966) sense), and that economic behaviour is related to these attributes [(1)]. He was the first to define constructed price indices as 'hedonic price indices' [(2)]. However, the fact that, up to now, hedonic analysis has been applied to a wide range of quality-related issues is largely due to the work of Zvi Griliches and Sherwin Rosen. The foundations were laid by Griliches (1961) in his characteristics approach to the construction of price indices and in his subsequent work, and also by the unifying approach of Rosen, in which varying marginal implicit prices are derived from both a distribution of marginal rates of substitution and marginal rates of transformation [(3)]. Applications range from correcting consumer price indices for changes in quality of the 'constant' commodity basket, to valuing amenity changes (Cropper and Oates, 1992), estimating the value placed by consumers on avoiding risks of losses from hazardous events (MacDonald, 1990) or studying the implicit valuation of market participants for attributes (Berndt and Griliches, 1993). But hedonic studies have been motivated by two main concerns. First, the need

[(1)] Though Gorman's paper was written in 1956, it was not published until 1980.

[(2)] See Goodman (1998) for an interesting revisit of Court's 1939 analysis.

[(3)] The generalised commodity approach to demand analysis (Houthakker, 1952) was the first one to present the hedonic function as a market phenomenon. Existing literature on hedonic quality measurement before Lancaster (1971) had already proved that analysis of consumption at the level of characteristics is more powerful than traditional analysis (Triplett, 1971). A study of Gorman's (1980) theory of linear consumption activities shows that Lancaster (1971) followed Gorman in specifying hedonic contours.

to identify implicit attribute prices. And second, to investigate welfare impacts by analysing the structure of demand for attributes (Follain and Jiminez, 1985; Bresnahan and Gordon, 1997).

Several recent studies including Golan and Shalit (1993), Oczkowski (1994), Nerlove (1995) and Combris *et al.* (1997) have applied hedonic price analysis to identify the implicit valuation placed by consumers, producers and marketers on wine attributes. In Golan and Shalit's study on hedonic grape and wine pricing, the authors aim to identify and evaluate wine quality characteristics of Israeli grapes. Assuming that the Californian wine market is perfectly competitive, wine prices are presumed to reflect both consumer preferences and the value of grape quality attributes. If, therefore, Californian and Israeli wine consumers have the same preferences, the competitiveness assumption can be used to derive hedonic prices for the Israeli market. By estimating the relative contribution of grape characteristics to wine quality, and using the monetary values from the Californian market, the authors are able to value individual grape characteristics so as to provide a producer pricing schedule for Israel. This quality-based pricing schedule could then help reduce the production of poor-quality wines by giving Israeli farmers an appropriate incentive to supply grapes of higher quality.

Based on recommended retail prices for Australian premium table wine, Oczkowski identifies the implicit valuation of table wine attributes for consumers and retailers. On the producer side, the author suggests that the hedonic functions estimated provide important information on which longer-term investment decisions could be made. Oczkowski includes dummy variables for producer size in the hedonic regression and argues that this allows for two effects: first, for possible price-making strategies and second, for measuring the characteristic of 'exclusiveness'. Indeed, some consumers will seek particular wines from small producers for their limited availability, rarity and 'trendiness'. The author's innovative approach to the underlying dummy variable model allows for explicit estimation of coefficients for all dummy variables. Also, for the Australian wine economy, the author concludes that between 1991 and 1992 there were general downward pressures on wine prices, despite inflationary pressures throughout the economy.

Nerlove builds on a generalisation of the 'pure repackaging' case labelled 'variable repackaging' case of quality differences by Fisher and Shell (1971), and in which the amount of repackaging is allowed to depend on the quantity of the good. Due to state intervention in wine pricing Nerlove does not follow standard hedonic regression, but assumes that variety prices are exogeneously determined and that consumer preferences are expressed by the purchased quantity of each variety. Therefore, variety supplies are viewed as perfectly elastic for the considered group of consumers and the quantities of each variety consumed are regressed on the unit variety price and on the measures of quality attributes which characterise that variety. He estimates a hedonic price function for wine using Swedish data from 1989-1991. The price elasticity is estimated to be about -1.65. This suggests that Swedish consumers are highly sensitive to prices. Estimates of implicit valuations of quality attributes are shown to greatly differ from those obtained from the classical hedonic regression where price is the dependent variable.

In their study on Bordeaux wine, Combris *et al.* (1997) use a stepwise regression procedure to investigate whether quality 'matters' or not in explaining market prices. The authors suggest that for their data set, quality as measured by a jury grade attributed by professional wine tasters can mainly be explained through the wine's 'subjective' sensory characteristics, which can not be observed when consumers choose the wine. The authors derive implicit prices for a wine tasting panel that has no access to any of the 'objective characteristics' during the tasting process (grape variety, vintage year etc.), including price, of the wines they judge.

In this article, we aim to comprehend wine variants by estimating hedonic price functions for still light wine on offer on the British off-licence market[(4)]. Since the following empirical analysis relies on data from 1994, we will briefly introduce the developments that took place in supply and demand at that period. The United Kingdom (UK) wine market was and is dominated by still light wine imports (more than 90 %, 1994 value). English and Welsh wine, produced from fresh grapes, accounts for only 0.3 % (1997 value) of domestic consumption. Imports, primarily from the EU but increasingly from a wide range of third countries, make up the balance. Currently, more than 25 countries of origin are represented in the UK wine market. Two types of licences allow for alcoholic beverages to be sold in the UK. 'Off-licences', where products are consumed outside the premises in which they were purchased (*e.g.* retail outlets), and 'On-licences', where alcohol is consumed *in situ* (*e.g.* pubs, clubs and restaurants). With more than 45 000 points of sale and 70 % of total wine sales in 1993 (value), the UK wine market is dominated by the off-licence sector. Regarding the evolution of sales by country of origin, the four great traditional suppliers, *i.e.* France, Germany, Italy and Spain, still dominate even though collectively, if not in all cases individually, their share has eroded. Their combined share declined from 89 % of the volume of imported wine from fresh grapes in 1983 to 78 % in 1993. Most countries heavily depend on off-licence sales, with France and, to a lesser extent, Germany, disproportionally depending on the on-licence trade.

The annual per capita consumption of wines in Britain has increased threefold from the early 1970's to the early 1990's (CFCE, 1994). Nevertheless, apart from Northern Ireland, Great Britain is characterised

(4) Still light wine is defined as the product obtained exclusively from the total or partial alcoholic fermentation of fresh grapes or fresh musts, with a total alcoholic strength usually not exceeding 15 %.

by the lowest consumption level in the EU. With a 64.5 litres annual per capita wine consumption in 1992, France is the world leader, whereas in the UK only 12.4 litres were consumed in the same year (Robinson, 1994). Considering the consumption pattern in relation to colour, the 1993 sale shares per volume of still light wine imported in the UK were 63.7 % for white wine, 33.2 % for red wine and 2.9 % for rosé (EIU, 1994).

We rely on a survey covering 3 940 bottles of still wine that were uniquely identified by objective attributes (region of origin, vintage etc.). The article both contributes to and departs from the existing hedonic price literature on wine markets in several ways. First, we expand the existing econometric approach. Theoretical restrictions of the hedonic approach and of heteroscedasticity are explicitly addressed by using a General Least Squares (GLS) estimator. We elaborate on a dummy variable approach pioneered by Kennedy (1986) and Oczkowski (1994) to obtain a distinct and comparable contribution to the variation of wine prices for each attribute. Secondly, in contrast to previous papers such as Combris *et al.* (1997), we do not rely on sensory characteristics. Indeed, we place our hypotheses on two sets of variables. We consider objective attributes, which consumers may observe on the label, and which are thus considered to determine the use value and tasting qualities of the wine. But we also consider retailer traits as an additional choice variable with no direct impact on tasting qualities. Thirdly, more than 14 000 observations are used in this study, reflecting the significance of 3 940 identified bottles of wine in the retailers' samples. Finally, we use a specific example to demonstrate the importance of price estimates in marketing a particular range of wine qualities. Likewise, in contrast to previous hedonic wine-related studies, we rely on actual retail prices rather than on recommended retailer prices.

The article is organised as follows. The firts section develops the theoretical and empirical framework needed to describe the valuation of wine attributes by agents, based on previous models of product differentiation. The second provides a description of the survey data employed but also a presentation and discussion of the econometric results. The third section presents the marketing implications of the econometric model while the last section concludes.

A HEDONIC FRAMEWORK

Theoretical considerations

The Lancaster model and its variants consider hedonic price functions as only reflecting consumer behaviour and assume a discrete spectrum of alternative qualities. Lancaster contrasts with another influential model of consumer choice based on product characteristics, *i.e.* that of Houthakker (1952) and Theil (1952) respectively. Houthakker (1952), who assumes a continuous spectrum of product qualities, was the first to develop a market notion of hedonic prices. In Rosen's product differentiation model – upon which this article relies – this market notion is developed further. Market clearing conditions determine the set of hedonic prices, where hedonic prices are defined as the implicit prices of attributes as disclosed to economic agents from observed prices and specific amounts of those characteristics which they are associated with. What is being estimated in Rosen's description of a competitive equilibrium is the locus of intersections between the demand curves of different consumers with varying tastes and the supply functions of different producers with possibly varying production technologies. The estimated implicit prices for quality therefore give us the implicit marginal valuation placed by consumers and producers on a vector of attributes. When choosing one unit (bottle) of wine, the maximisation of utility U subject to a non-linear budget constraint requires the choice of a composite good with vector x and a vector of wine attributes $(z_1, ..., z_n)$ to satisfy the constraint $y = p(z) + x$ and the first-order conditions [5],

$$\frac{\partial p}{\partial z_i} = p_i = \frac{\partial U / \partial z_i}{\partial U / \partial x}, i = 1,\ldots,n \quad . \tag{1}$$

The marginal rate of substitution between wine attribute z_1 and x equals, therefore, the marginal price of wine attribute z_1.

This paper departs from a one-period model of the choosing behaviour in wine consumers, in which the agent chooses one wine attribute bundle at a time from a number of different wine attribute bundles. In a setting of perfect competition, we assert that market equilibrium conditions are reflected in the valuation of the attributes. Although certain attribute combinations can only be selected in a reshuffled form (it may be possible to find a Cabernet Sauvignon 1992, aged for two years in barrels, of *either* French *or* Chilean origin), we assume that any quantity can be supplied to match consumer demand. Hence, we conjecture perfect divisibility.

Objectives and hypotheses

In the following empirical section, we attempt to investigate the relationship between varietal and regional origin and quality as perceived by consumers. To do so, we examine implicit wine attribute prices

[5] Jones (1988) elaborates on Rosen to demonstrate the origin and consequences of non-linearity. Since attributes cannot be traded directly, bundling constraints will always lead to trade (within the Edgeworth Box) where certain attribute combinations are not mutually advantageous. Hence we observe trade that is not mutually advantageous, although the marginal utilities of attributes do not equate.

through the estimation of hedonic price functions. In the absence of consumer-specific characteristics, we cannot investigate welfare impacts that could be derived in instance where the structure of demand for attributes could be identified.

Consumers are seen as facing a choice of wines from different countries of origin when confronted with the labels of the bottles on the shelf. It is asserted that a first group of attribute categories (quality designation, grape variety, vintage, region and country of origin) determine the use value, hence the tasting qualities, of the wine. Another category group, *i.e.* the originating retailer, is deemed to have no bearing on this use value and is therefore taken as not entering the consumer's utility function for tasting qualities. The consumers' willingness-to-pay should, therefore, be determined by variables from the first group of categories only, unless retailer traits enter the utility function in an indirect way. However, since we have no information concerning individual retailer traits, we assume that the valuation placed by the consumer on the retailer's name reflects the aggregate valuation of relevant retailer traits to the consumer. We consider that retailer traits are relevant to the consumer when they indirectly enter his or her utility function. Since our implicit prices are assumed to reflect an equilibrium price relationship, they can be given both a user value and a resource cost interpretation. Hence, we assume that retailers themselves incur costs to build a reputation based upon their own traits. Since they receive a competitive return on their reputation investment[(6)], they regard reputation as an asset. If we consider reputation linked to the retailer's name as a more abstract wine attribute, a second step in consumption technology is then required to obtain 'pleasure from reputation' from the retailer traits. Two broad categories of objective and abstract qualities emerge, the dividing line adopted here being the measurement on an objective, cardinal scale (Lucas, 1975). If utility is directly derived from 'pleasure of reputation', and, if this attribute is in turn acquired from other objective attributes, a utility function can then be written, substituting the second step of the consumption technology into the expression of utility:

$$U = U = U\,(r^1(z_1, z_2),\, r^2\,(z_3, z_4),\ldots) \tag{2}$$

where $r_1 = r^1(z_1, z_2)$ would indicate the amount of 'pleasure from reputation' obtained from the retailer traits.

In a market where reputation effects may be important, we assume that the degree of information on wines possessed by the consumer will be reflected in his or her degree of product involvement. This degree of product involvement can be identified by analysts from the willingness of wine consumers to differentiate between, and pay for, different attributes

(6) Shapiro (1983) demonstrates that the introduction of reputation as an asset that must initially be built up allows the construction of an equilibrium model that includes perfect competition, free entry, and quality choices by firms under imperfect information.

within the total attribute bundle. We therefore assume that the further down their decision trees consumers are willing to proceed, the more distinct attributes they are willing to pay for, and the higher their level of information about the attributes which they are comparing must be.

Specification and interpretation of the hedonic price functions

Given the qualitative nature of the data (dummy variables only; see the section on the survey data) and the need to retain comparability across attributes, variables first have to undergo a modification that only alters interpretation of the estimates. This modification does not alter the underlying meaning of implicit price estimates as 'missing prices', in a hypothetical market where both consumers and producers are asked to attribute their valuation to the existence of a particular wine attribute, *ceteris paribus*.

As a result of this modification, and after adjusting the coefficient estimates with the estimated variances, the final interpretation is that coefficient estimates measure the relative impact of the presence of the attribute *ceteris paribus* on the dependent variable (unit price evaluated at the sample means).

Although the choice of the functional form for the hedonic price function should remain an empirical matter, theory suggests that non-linear functional forms frequently provide a more appropriate alternative. Likewise, on pragmatic grounds and with respect to heteroscedasticity, non-linear forms such as the semilogarithmic (log-lin) model could be preferable. In this instance, the coefficient of a dummy variable measures the percentage effect of the presence of the factor represented by the dummy variable on the dependent variable. However, Kennedy (1981) objects to interpretations by Halvorsen and Palmquist (1980) who estimate the percentage effect on asymptotic grounds as their suggested procedure leads to a biased estimator for the dummy variable[7]. Instead of estimating g by

[7] Halvorsen and Palmquist demonstrate that in the instance where the dependent variable is lnY, the coefficient of the dummy expresses change in $ln\ Y$ units since it reflects the difference in subgroup means between the designated group and the reference group in units of the dependent variable. The authors report the general form of a log-lin equation as,

$$\ln Y = a + \sum_i b_i X_i + \sum_j c_j D_j \qquad \text{(F.1)},$$

where X_i reflects continuous variables and D_j represent dummy variables. In the above simplified case of a single dummy variable, the interpretation of the coefficient of the dummy variable is revealed by transformation of equation (F.1):

$$Y = (1 + g)^D \exp(a + \sum_i b_i X_i),$$

where

$$g = (Y_1 - Y_{ref})/Y_{ref}.$$

Y_1 and Y_{ref} are the predicted values of the dependent variable when the dummy variable

$$\hat{g} = \exp(\hat{c}) - 1 \quad (3)$$

Kennedy suggests following Goldberger (1968) and estimating g by

$$g^* = \exp\left(\hat{c} - \frac{1}{2}\hat{V}(\hat{c})\right) - 1, \quad (4)$$

where $\hat{V}(\hat{c})$ is an estimate of the variance of $(\hat{c})$, which is asserted to have less bias than $\hat{g}$.

Suits (1984) suggests a procedure to adjust dummy variable coefficient estimates keeping all variables in the equation. He interprets the estimates as deviations from average behaviour[8]. Following Suits, we impose identifying restrictions, but instead of employing Kennedy's laborious extension of Suits, we expand on Oczkowski and substitute the full constraint into the original equation. Following symmetrical estimations, it is possible to obtain all coefficient estimates. If, for example, the objective was to get coefficient estimates for wine colours (red, white, rosé: C_1, C_2, C_3) and, say, all four producer regions of a given county (R_1, R_2, R_3, R_4), both constraints (5) and (6) could be substituted into the original hedonic price model as follows:

$$\begin{aligned} \alpha_1 Pc_1 + \alpha_2 Pc_2 + \alpha_3 Pc_3 &= 0 \\ \alpha_1 &= [-(\alpha_2 Pc_2)/Pc_1 - (\alpha_3 Pc_3)/Pc_1] \end{aligned} \quad (5)$$

where Pc indicates the mean, hence the proportion of non-zero's, in the colour categories for each bottle of wine. And

$$\begin{aligned} \beta_1 Pr_1 + \beta_2 Pr_2 + \beta_3 Pr_3 + \beta_4 Pr_4 &= 0 \quad (6) \\ \beta_2 &= [-(\beta_1 Pr_1)/Pr_2 - (\beta_3 Pr_3)/Pr_2 - (\beta_4 Pr_4)/Pr_2] \end{aligned}$$

where Pr reflects the proportion of non-zero's for each bottle of wine in the region categories. When substituted into the original equation, this gives:

$$\begin{aligned} P = {} & [-(\alpha_2 Pc_2)/Pc_1 - (\alpha_3 Pc_3)/Pc_1] C_1 + \alpha_2 C_2 + \alpha_3 C_3 + \beta_1 R_1 \quad (7) \\ & + [-(\beta_1 Pr_1)/Pr_2 - (\beta_3 Pr_3)/Pr_2 - (\beta_4 Pr_4)/Pr_2] R_2 + \beta_3 R_3 + \beta_4 R_4 \end{aligned}$$

---

is equal to one and stands for the reference group, respectively. The coefficient of the dummy variable in the (F.1) equation is therefore $c = \ln(1 + Y_1 - Y_{ref})/Y_{ref}$. Since g displays the relative effect on Y of the presence of the factor represented by the dummy variable, the percentage effect of the dummy variable on Y, in units of Y, is found by applying the antilog function, $100.g = 100 . \{\exp(c) - 1\}$, which is therefore the percentage difference associated with being in group 1 rather than being in the reference group. Thus, if for example -.246 emerges, after taking the antilog to the base e and subtracting 1, the expected value of Y for the designated group is found to be 24.6 % lower than the value for the reference group. The expected implicit price for a bottle of wine from region A would thus be 24.6 % lower than the expected price for a bottle from region B, *i.e.* the reference group.

[8] Instead of forcing one of the coefficients of the dummy variables to be zero, all could be restricted to zero and the resulting intercept be interpreted as the average of intercepts of all observations in the sample.

and,

$$P = \alpha_2[C_2 - (Pc_2/Pc_1)C_1 + \alpha_3[C_3 - (Pc_3/Pc_1)C_1] \quad (8)$$
$$+ \beta_1[R_1 - (Pr_1/Pr_2)R_2] + \beta_3[R_3 - (Pr_3/Pr_2)R_2] + \beta_4[R_4 - (Pr_4/Pr_2)R_2]$$

The corresponding hedonic model therefore assumes the following form,

$$p^* = \alpha_2[X_{a2}] + \alpha_3[X_{a3}] + \beta_1[X_{b1}] + \beta_3[X_{b3}] + \beta_4[X_{b4}] + \varepsilon, \quad (9)$$

where p^* is a $N \times 1$ vector of transformed observations on the dependent variable (the price per bottle). There are vectors of **X** of five $N \times 1$observations, α and β denote the unknown parameters and ε is a $N \times 1$ vector of unknown stochastic disturbances. In the example given in (9), symmetrical substitution generates estimates for the missing α_1 and β_2 coefficients ('symmetrical regressions').

However, the above specification embodies an equivalence effect. The effect of grape variety, for example estimated implicit price differences between Cabernet Sauvignon and Shiraz, is assumed to be the same across all regions [9]. Therefore, a model providing sufficient flexibility to allow differential effects should be specified. Interaction terms are therefore introduced, which enable us to test for these differential effects [10].

EMPIRICAL IMPLEMENTATION

Functional form specification

Since there is little theoretical guidance on the functional form in hedonic regressions, the objective should be to include all forms that theory shows are plausible. However, as all our explanatory variables are of a qualitative nature, the choice of the functional form is limited to the linear and the log-lin, *i.e.*, semilog, specification. Nevertheless, the use of interaction terms allows us to gain additional flexibility.

The log-lin hedonic function produces non-constant marginal Engel prices (the prices paid for incremental units of characteristics when purchased as part of the same bundle), but it exhibits constancy of relative

(9) This assumes that the traditional way of dropping one category to avoid perfect multicollinearity is pursued.

(10) The interaction terms of primary interest are those for *region/variety*. Coefficient estimates for those product variables reflect the differential effect of region for each variety. For example, the interaction term for grape variety and region estimates the extent to which, say, the effect of being Chardonnay differs for Hunter Valley versus Napa Valley.

prices with respect to changes in proportions of characteristics (Triplett, 1975). This log-lin specification therefore assumes homotheticity of the utility function, hence homogeneity of degree zero of demand equations for attributes. Since only relative prices matter, the imputed price is independent of the level of the characteristic, which appears to be a realistic and convenient assumption, since only qualitative variables are used as explanatory variables in the present model. Moreover, since the log-lin form allows each marginal implicit price to be a non-linear function of the entire set of characteristics, it appears as an attractive alternative hypothesis, since it accommodates the idea that bundling constraints are present for wine attributes in a bottle of wine.

The survey data

The data originate from a survey undertaken in August 1994 in 94 retail outlets of different commercial forms in England and Scotland (see Appendix I). The aim is to give a representative sample of foreign still wines sold in those regions in the off-licence sector, as retail outlets were selected according to the market share of a company in 1993-1994.

Each price per wine bottle is, where appropriate, described by a combination of the following dimensions:

- country of origin	- importer
- colour	- brand (*e.g.* Gallo)
- category (AOC, QbA, DO, etc.)	- vintage
- region of origin	- place of bottling
- appellation (Chianti, Rioja, etc.)	- volume
- producer	- grape variety

Thus, all information appearing on the bottle labels was collected, except for the alcohol degree. The survey shows the number of outlets per company in which a bottle was found. We use this information as quantity proxy. On the whole, the survey included 14440 bottles (prices) from 13 countries of origin that are identified by 575 attributes. The large number of bottles is due to the fact that 3940 uniquely identified bottles of still wines appear on average in 3.7 retail outlets of the same commercial form.

Specification search and data analysis

As there appears to be no single, readily available econometric approach for estimating the above functional (hedonic) relationship, the following modelling strategy borrows from several methodologies, namely from those frequently associated with David Hendry and Edward Leamer.

The present analysis follows Leamer's (1990) 'classical' references to sensitivity analysis and subsequent attempts to simplify the models by incorporating the insights gained from specification uncertainty diagnostics and measurement error diagnostics. Although the Hendry methodology is based on time series, Hendry's 'general-to-specific' approach and the related steps are thought to be appropriate in the present cross-sectional context[11]. The evaluation of the resulting model by extensive analysis of residuals and predictive performance is borrowed from the final step of Hendry's analysis. We expand the above approach by applying the diagnostic framework suggested by Belsley *et al.* (1980), and Belsley (1986), to uncover statistical problems in an OLS framework. By proceeding in this fashion, we hope that the strengths of the above approaches can be applied together, so as to ensure robust estimation procedures providing stable implicit price estimates. We follow Leamer (1990) in distinguishing three phases in data analysis: (1) estimation, (2) sensitivity analysis and (3) simplification.

(1) Estimation

Model selection

Since the hedonic price theory does not provide further guidance to the inclusion of variables in the present application (it is assumed that all pre-selected variables have a resource cost/user value interpretation), the following estimation and testing procedure is of high importance.

For the initial hedonic regression, we carry out a subjective pre-selection of attribute categories based on information from broadly available UK magazines for wine marketers and consumers, such as *"Decanter"*, *"The Sainsbury's Magazine"*, and weekly magazines and supplements of daily newspapers (*Observer*, *Independent*, *Financial Times*). As the mean price in the sample is 551 pence, we suggest that highly reputed 'appellations' as well as highly reputed producers are less important. Therefore, we exclude 'appellations' and producer names. All the remaining attribute categories were included in the initial regressions, jointly with a subset of interaction terms: interactions for colour/country of origin, colour/region of origin, category/country of origin, grape variety/region, and grape variety/country of origin. Following this pre-selection of

[11] Hendry suggests to first formulate a general model which is consistent with what economic theory postulates and which is subsequently reparameterised to obtain explanatory variables near orthogonal (Davidson *et al.*, 1978; Hendry, 1980). Leamer suggests the formulation of a general family of models and, in a second step, to decide what inferences are interesting and to express these in terms of parameters. In addition, in this estimation stage, Leamer recommends the use of specification uncertainty diagnostics and measurement error diagnostics (see Pagan, 1995, for a more detailed discussion).

regressors, the subsequent selection procedure, based on the single equation hedonic approach, does not follow a purely mechanical procedure – such as stepwise regression – as dangers of doing so are well established (*e.g.* Wallace and Ashar, 1972; Judge and Bock, 1983; Leamer, 1983b; Greene, 1997).

Specification tests

First, we test for the equality of implicit price contributions. This is done in two ways, while relying most heavily on the second one. We first follow Berndt and Griliches (1993) and compensate for the large sample size by choosing very tight significance levels for the standard F-tests (.01 significance level). Secondly, we follow Ohta and Griliches (1975, 1986), who suggest to consider the difference in fit between unconstrained and constrained regressions specifically for hedonic models, and not to reject the simpler hypotheses unless they are very different. Hence, we compare the standard errors of both regressions. However, we only consider the null hypothesis of parameter equality as relevant if it is based on economic significance rather than statistical significance. If the difference in standard error of regression is smaller than or equal to .01 in the system being tested, the null hypothesis will not be rejected on practical grounds. As the regression is semilogarithmic, a .01 increase in SE implies an increase of about 1 % in the standard deviation of the unexplained price component (the fit to actual price data is smaller by one per cent in the constrained regression than in the unconstrained regression)[12]. Looking for the most parsimonious specification, we follow Berndt and Griliches (1993) in rejecting the null-hypothesis when the root mean squared errors under the alternative results in a more than 5 per cent reduction in the standard deviation of the unexplained log price variation. The following specification tests were applied:

(a) Tests for Heteroscedasticity

In this article, we rely on the Breusch-Pagan (1979) test and its extension by Koenker (1981). We apply weighted regressions as this has a two-fold advantage. First, it enables us to correct for heteroscedasticity

[12] Assume a difference of 0.01 in the standard errors in the constrained and unconstrained regressions and a standard error (SE) of 0.1 of the constrained regressions. The implication is that the lack of fit of the constrained regression is increased by 10 per cent compared with that of the unconstrained regression (0.01/0.1 = 0.1). Equally, if the SE was 0.2, the 0.01 criterion implies willingness to accept up to 5 per cent deterioration in the fit of the model as measured by the standard error of its residuals.

by transforming the error terms[13]. But it also complies with hedonic theory, as each attribute should be accounted for in terms of its market significance. Hence, using weighted regressions where weights reflect a proxy for the demanded quantity, should provide meaningful results.

(b) Specification tests for collinearity

Atkinson and Crocker (1987) emphasise two consequences of multicollinearity in hedonic models. First, the mean squared error of the estimator may cause substantial instabilities in coefficient signs and magnitudes as independent variables are added or removed from the model. Secondly, measurement error bias may be partly transferred to collinear variables measured without error and alter their signs.

(i) As in standard analysis, we consider *F*-, *t*-values and corrected *R*-square together, and ask whether there is a lack of individual significance despite overall significance and high corrected *R*-square. At the same time, a standard model selection criterion – the Akaike information criterion (AIC) – is selected here in an attempt to judge trade-off between model complexity and goodness of fit. The AIC criterion is preferred to the Schwarz one in the present context of numerous potential variables, because the latter penalises model complexity much more heavily than other criteria.

(ii) Auxiliary regressions are considered, as collinearity can appear both in the form of linear dependence between variables, and as a lack of variation in the values of a control variable about its mean. Thus, auxiliary regression R square and the sum of squared least squares residuals from the auxiliary regression are considered together (Griffiths *et al.*, 1993).

(iii) We examine the *condition number* of the data matrix (Belsley *et al.*, 1980)[14]. Judge *et al.* (1985) suggest that moderate to strong near

[13] The Goldfeld-Quandt test may lack power if an error variance that is related to more than one variable is present. Since the present analysis employs GLS, only one form of heteroscedasticity is tested. Given the weights in the present study, it is assumed that the error variance varies with the expected price. The consequence is that White's (1980) heteroscedastic-consistent covariance matrix estimation, which corrects the estimates for an unknown form of heteroscedasticity, cannot be used.

[14] The condition number of a non-singular matrix provides a measure of the potential sensitivity of the solution vector **z** of a linear **Az** = **c** system of equations to changes in the elements of **c** and **A** of the linear system. Therefore, the conditioning of any square matrix **A** can be summarised by a condition number, defined as the product of the maximal singular value of **A** (its spectral norm) and the maximal singular value of $\mathbf{A}^{-1}$ (the square root of the ratio of its largest to its smallest eigenvalue). Hence, the condition number of any matrix with orthonomal columns is unity. For a general linear model, where the OLS estimator is the Best Linear Unbiased Estimator with variance-covariance matrix $V(b) = \sigma^2(X'X)^{-1}$, collinearity among the columns of X causes crucial elements of the cross-product matrix $(X'X)$ to be large and unstable, thus affecting both the b and $V(b)$ estimates. For $V(b)$, the effect is to blow up its diagonal elements, creating inflated variances. And the effect on b is to render this solution very sensitive to small changes in X (Belsley, 1986).

exact linear dependencies are associated with condition indices between 30 and 100[15].

(2) Sensitivity Analysis

In seeking a robust estimation procedure that can produce estimates which are insensitive to model misspecifications, we follow Leamer (1990) in his 'classical approach' to sensitivity analysis. Leamer suggests investigating whether econometric results can be attributed to their modelling assumptions, hence whether inference is fragile and not believable (Leamer, 1983a). We apply techniques to discover influential observations, as developed by Belsley *et al.* (1980). These techniques are complemented by applying the trimmed least square estimation method as performed in SHAZAM. In this robust estimation method, a desired proportion of trimmed observations is discarded and OLS is then applied to the remaining observations.

Belsley *et al.* (1980) suggest three means for deletion diagnostics. First, we examine single-row diagnostics. We investigate changes that would occur in estimated regression coefficients if the *i*th observation were deleted. This DFFIT diagnostic measure has the advantage of being independent from the particular co-ordinate system used to form the regression model. Scaling this measure with the standard deviation of the fitted values displays a scaled row-deleted change in fit (DFFITS)[16]. Secondly, we examine the *hat matrix* by studying the diagonal elements of the least-square projection[17]. Finally, we also run a Lagrange-Multiplier test for normality (Jarque-Bera). We exploit the link between *hat matrix* and residual variance by investigating the standardised residual, which is frequently called *studentised residual*[18]. If the observation

[15] See Judge *et al.* (1985, p. 904) for a discussion of the limitations due to the fact that the exact and separate nature of linear dependencies among explanatory variables cannot be determined.

[16] First, sort both DFITS and log price, then plot them to get an impression of the relative magnitudes of the DFITS and notions of price clusterings. For details on the equations used, see Belsley *et al.* (1980) equations 2.10 and 2.11.

[17] This *hat matrix* (equation 2.15 in Besley *et al.*, 1980) determines the fitted or predicted values. The influence of the response value on the fit is most directly reflected in its impact on the corresponding fitted value. Since the diagonal elements of the *hat matrix* have a distance interpretation (*ibid.*, p. 16), they provide a basic starting point for revealing 'multivariate outliers' which would not be revealed by scatter plots when $p > 2$. Based on the Gaussian assumption, Belsley *et al.* (1980, p. 17) identify when a value of the *hat matrix* is large enough, *i.e.* far enough from average, to warrant attention. The authors define the *i*th observation as a *leverage point* when the *hat matrix* exceeds $2p/n$. Although leverage only measures the influence of the regressors and not the influence of outlying observations caused by large absolute errors, we do not consider measures that are designed to detect large errors (see Judge *et al.*, 1988, p. 894).

[18] The RSTUDENT option in SHAZAM. In seeking significantly large residuals, it is required to standardise them by dividing by the appropriate standard error for that particular residual. In contrast to Judge *et al.* (1988, p. 984) and Belsley *et al.* (1980, p. 20), Greene (1997, p. 445) suggests that the standardised residual is approximately distributed as standard normal, and not distributed as t.

conforms to the model estimated through other observations, this standardised residual should be small (the calculation is repeated for each observation). Absolute values under two are acceptable in terms of model specification. Others are regarded as outliers. However, since some of the most influential data points can have relatively small studentised residuals, row deletion and residual analysis are studied together and on equal footing (Belsley *et al.*, 1980, p. 21).

With respect to row deletions, Belsley *et al.* (1980, p. 22) suggest employing the COVRAT statistic and comparing the covariance matrix using all data with the covariance matrix resulting from deletion of the *i*th row[19]. Since this magnitude is a ratio of the estimated generalised variances of the regression coefficients with and without the *i*th observation deleted from the data, it can be interpreted as a measure of the effect of the *i*th observation on the efficiency of coefficient estimation (Belsley *et al.*, 1980, p. 48). Since the two matrices differ only in the inclusion of the *i*th row in the sum of squares and cross products, near unity values of this ratio can be taken to indicate that the two covariance matrices are close, or that the covariance matrix is insensitive to the deletion of row *i*. Therefore, a value of COVRAT greater than one indicates that the absence of the associated observation impairs efficiency.

However, in order to discover which observations are the most strongly influential, we follow Belsley *et al.* (1980) and apply external scaling, where cut-off values are determined by having recourse to statistical theory. If observations have a high leverage *and* significant influence on the estimated parameters, there is enough evidence to view them as presenting potentially serious problems. Accounting for the above measures, we consider about 2.4 % of the observations to be occasionally influential. However, results from the trimmed least square estimation also suggest that parameter estimates are sufficiently stable to continue with weighted least squares regressions.

Hedonic price models: estimation and results

The summary statistics can be found in Appendix II. The testing procedure starts with assessing the hedonic price functions using a General Least Squares (GLS) estimator[20]. Hence, the regressions are implemented as weighted least squares regressions, where ordinary least squares (OLS) are applied to a transformed model. Consider the following linear regression model,

$$Y_t = X_t' \beta + \varepsilon_t \quad \text{where } E[\varepsilon_t' \varepsilon_t] = \sigma^2 / W_t \quad \text{for } t = 1, \ldots, N \qquad (10)$$

[19] Belsley *et al.* (1980) equation 2.36 and COVRAT in SHAZAM.

[20] The following regressions are performed by using SHAZAM, version 7.0 (White, 1993).

and the dependent variables Y_t, the exogenous variables X_t and the W_t weights are observed with β and σ^2 as the unknown parameters. Each observation of the dependent and independent variables is replicated W_t times and multiplied by the square root of the weight variable. The weighted least squares estimator is then obtained by applying OLS to the transformed model:

$$\sqrt{W_t}\, Y_t = \sqrt{W_t}\, X_t' \beta + v_t \qquad (11)$$

The resulting GLS regressions were performed for two reasons. First, and most importantly, GLS rather than OLS is used as an estimation rule on the ground that each attribute (and its price) in the context of hedonic market studies is important only to the extent that it captures some relevant fraction of the market (Griliches, 1971). Here, the weights applied in the GLS regressions reflect the number of retail outlets of each retailer type (*e.g.* Marks & Spencer) in which a uniquely identified bottle was found. It is therefore implicitly assumed that the sample fractions are directly proportionate to the number of bottles sold, although this is clearly an imperfect assumption. Secondly, the implementation of GLS allows us to account for heteroscedasticity due to omitted variables and/or to misspecification.

It was suggested that certain categories of attributes (quality designation, grape variety, region and country of origin, vintage) could determine the use value (tasting qualities) of the wine, and therefore be part of the utility function of the consumer. Another category was asserted not to have any bearing on this use value (*i.e.* retailers). The consumer's willingness-to-pay would therefore be determined by variables from the first group of categories. However, results in Table 1 suggest that retail outlets where the bottle is chosen (and thus the retailer traits) affect consumer choice in significant ways. Although it was not possible to compare exact attribute bundles for 'non-taste attributes' (namely retailers), distinct and significant valuation of retailers were identified. Results indicate that consumers attach a high value to the information provided on the label. In all cases where conditional effects between attributes were found to have a significant impact on price, consumers are viewed as regarding these attribute bundles as imperfect substitutes. In these instances of more than overall impacts, outstanding grape varieties are shown to have a strongly positive or negative regional impact on price, just as outstanding regions have a similar grape varietal impact.

The estimates are interpreted as follows. The valuation which a consumer is assumed to place on the wine colour is as anticipated, since the parameter estimates for red (+2.2), white (-1.6) and rosé (-11.2) take the expected signs[21].

---

[21] Since the majority of top quality wines is red rather than white, the relative valuation is as expected (assuming we do not regard reds and whites as complements).

Table 1. Econometric results of the log-lin hedonic model

VARIABLE DESCRIPTION	VARIABLE NAME	RELATIVE[(a)] IMPACT %	ESTIMATED COEFFICIENT	STANDARD ERROR	T-RATIO[(b)]
	CONSTANT		1.5234	1.73E-02	479.90
COLOUR	*RED	2.20	2.17E-02	3.33E-03	6.52
COLOUR	WHITE	-1.63	-1.65E-02	3.39E-03	-4.86
COLOUR	ROSÉ	-11.23	-0.11894	1.60E-02	-7.43
COUNTRY OF ORIGIN	*ARGENTINA	-26.19	-3.02E-01	5.29E-02	-5.71
COUNTRY OF ORIGIN	AUSTRALIA	5.32	5.18E-02	9.39E-03	5.52
COUNTRY OF ORIGIN	GERMANY	-28.04	-0.32895	1.16E-02	-28.49
COUNTRY OF ORIGIN	BULGARIA	-39.17	-0.49689	1.79E-02	-27.82
COUNTRY OF ORIGIN	CHILE	-10.47	-0.11033	2.17E-02	-5.09
COUNTRY OF ORIGIN	SPAIN	-20.29	-0.2267	1.46E-02	-15.54
COUNTRY OF ORIGIN	HUNGARY	-34.36	-0.42088	1.90E-02	-22.15
COUNTRY OF ORIGIN	ITALY	-3.99	-4.07E-02	1.07E-02	-3.82
COUNTRY OF ORIGIN	NEW ZEALAND	25.55	0.22765	1.52E-02	14.99
COUNTRY OF ORIGIN	PORTUGAL	-19.89	-0.22167	1.61E-02	-13.75
COUNTRY OF ORIGIN	ROMANIA	-48.44	-0.66143	4.33E-02	-15.28
COUNTRY OF ORIGIN	SOUTH AFRICA	-16.31	-0.17792	1.65E-02	-10.78
COUNTRY OF ORIGIN	FRANCE	12.26	0.11566	3.00E-03	38.57
GRAPE VARIETY	CABERNET SAUVIGNON	7.26	7.01E-02	1.20E-02	5.83
GRAPE VARIETY	CHARDONNAY	15.33	0.14265	1.09E-02	13.13
GRAPE VARIETY	CHENIN BLANC	-8.91	-9.25E-02	3.84E-02	-2.41
GRAPE VARIETY	GEWÜRZTRAMINER	35.06	0.30119	3.59E-02	8.38
GRAPE VARIETY	PINOT NOIR	25.73	0.22924	2.41E-02	9.52
GRAPE VARIETY	RIESLING	36.56	0.31192	2.51E-02	12.42
GRAPE VARIETY	SANGIOVESE	-34.12	-0.41385	8.41E-02	-4.92
GRAPE VARIETY	SEMILLON	-28.82	-0.33959	2.48E-02	-13.70
GRAPE VARIETY	SAUVIGNON	11.69	0.11072	1.59E-02	6.97
INTERACTION TERM	RIESLING-AUSTRALIA	-34.93	-0.42895	3.91E-02	-10.97
INTERACTION TERM	SEMILLON-FRANCE	12.10	0.11463	2.85E-02	4.02
INTERACTION TERM	SAUVIGNON-FRANCE	-22.19	-0.25047	2.82E-02	-8.90
INTERACTION TERM	CHENIN BLANC-NEW ZEALAND	-31.99	-0.38472	4.05E-02	-9.51
INTERACTION TERM	SAUVIGNON-CHILE	-20.04	-0.22339	2.38E-02	-9.37
INTERACTION TERM	CHARDONNAY-SPAIN	21.23	0.19389	5.24E-02	3.70
INTERACTION TERM	CHARDONNAY-ITALY	-18.67	-0.20566	4.57E-02	-4.50
VINTAGE	1983	69.98	0.53356	7.83E-02	6.82
VINTAGE	1985	35.89	0.30769	4.51E-02	6.82
VINTAGE	1986	52.44	0.4222	3.55E-02	11.91
VINTAGE	1987	27.55	0.24358	2.14E-02	11.38
VINTAGE	*1988	28.80	0.25322	1.75E-02	14.50
VINTAGE	1989	14.60	0.13636	1.18E-02	11.61
VINTAGE	1992	-10.53	-0.11126	3.99E-03	-27.87
VINTAGE	1994	-15.35	-0.16615	3.07E-02	-5.42
RETAILER NAME	ASDA	-11.33	-0.12022	1.22E-02	-9.90

Table 1 (continued)

VARIABLE DESCRIPTION	VARIABLE NAME	RELATIVE[a] IMPACT %	ESTIMATED COEFFICIENT	STANDARD ERROR	T-RATIO[b]
RETAILER NAME	CO-OP	12.03	0.11389	2.41E-02	4.74
RETAILER NAME	M&S	23.11	0.20807	1.83E-02	11.37
RETAILER NAME	CWS	-12.01	-0.1276	2.71E-02	-4.71
RETAILER NAME	*SAFEWAY	-6.25	-6.44E-02	1.59E-02	-4.06
REGION OF ORIGIN	HUNTER VALLEY	23.41	0.21069	2.56E-02	8.23
REGION OF ORIGIN	*BREEDE	-25.02	-0.27812	1.40E-01	-1.99
REGION OF ORIGIN	COONAWARA	39.06	0.33012	2.67E-02	12.38
REGION OF ORIGIN	LA MANCHA	-26.98	-0.3135	4.21E-02	-7.44
REGION OF ORIGIN	RIOJA	18.89	0.17332	2.28E-02	7.59
REGION OF ORIGIN	VALENCIA	-25.66	-0.29628	2.46E-02	-12.07
REGION OF ORIGIN	VENETO	-25.96	-0.30042	1.85E-02	-16.20
REGION OF ORIGIN	DOURO	-12.19	-0.12939	3.53E-02	-3.67
REGION OF ORIGIN	SONOMA VALLEY	-16.74	-0.18303	1.73E-02	-10.56
REGION OF ORIGIN	BORDEAUX	-17.33	-0.19021	1.09E-02	-17.49
REGION OF ORIGIN	LANGUEDOC	-28.33	-0.33309	7.57E-03	-44.03
REGION OF ORIGIN	LIBOURNE	43.02	0.35815	2.63E-02	13.62
REGION OF ORIGIN	MEDOC	52.29	0.42075	1.66E-02	25.36
REGION OF ORIGIN	PROVENCE	-29.79	-0.35315	3.15E-02	-11.21
REGION OF ORIGIN	SAUTERNES	133.33	0.84878	5.49E-02	15.47
REGION OF ORIGIN	CÔTE CHALONNAISE	45.72	0.3774	4.13E-02	9.14
REGION OF ORIGIN	CÔTE DE BEAUNE	148.41	0.91027	2.61E-02	34.88
REGION OF ORIGIN	CHABLIS	48.01	0.3924	2.35E-02	16.73
REGION OF ORIGIN	CÔTE DE NUIT	152.10	0.92502	2.63E-02	35.16

[a] The impact of the attribute on price is measured as in equation (4).
[b] 14 380 degrees of freedom
Notes: Adjusted R-square: 0.52
Breusch-Pagan: Chi-Square = 138 with 59 regressors [for 59 D.F., P(chi square > 77.93) = 0.05]
Variables preceded by a * are taken from symmetric regressions.
"E-02" and "E-03" mean 10^{-2} and 10^{-3}, respectively.

Among the countries of origin, being French achieves the greatest impact on price (+12.3 %), whereas Romania shows the greatest negative impact (-48.4 %). A surprising finding is the only moderate positive impact of Australian origin (+5.3 %) especially when compared to France, while wines of Chilean origin display a rather high negative impact on price (-10.5 %). Sluggish expansion of imports and the recent increase in popularity, following the introduction of new winemaking technologies, may be part of the explanation regarding Chilean wines.

The highly different impact of New Zealand origins, as compared to the average (+25.5 %), is not particularly surprising. This valuation could be explained by the fact that Chardonnay and Sauvignon Blanc probably produce the top quality whites from the Island (+15.3 % and

+11.7 % respectively). The impact on price of German (-28 %), Bulgarian (-39.2 %), Hungarian (-34.4 %) and Romanian origins (-48.4 %) are equally expected.

Consumer valuation for Italy (-3.9 %) and Spain (-20.3 %), i.e. two Mediterranean competitors, are also expected. Italy thus represents the most typical average prices amongst all countries of origin. Although Italian Merlot, Cabernet Sauvignon and Chardonnay are sold on the British market, these grape varieties do not represent the qualities for which Italy has long been known. The expected negative price contribution therefore comes about because Italy has only recently been increasing its supply of the world's most favourite grape varieties. Spain's downgrading impact on price could partly be explained by the (low) performance and high importance (in volume terms) of its wines from La Mancha (-26.9 %), where nearly half of Spain's production originates, as well as from Valencia (-25.7 %).

The impact of regions on price are, therefore, in line with expectations for La Mancha, for Rioja (+18.9 %), as well as for Provence (-29.8 %), Côte Chalonnaise (+45.7 %) and Veneto (-25.96 %). However, the negative impact of Sonoma Valley (-16.74 %) and of Douro (-12.2 %)[22] are both surprising. While wines from the Douro valley have recently become more highly valued, the low performance of Sonoma is difficult to explain.

As for French regions, an implicit valuation of the AOC system seems to show up, since the impact of Libourne (+43 %), Médoc (+52.3 %) and Sauternes (+133.3 %), all of them in the heart of the Bordeaux region, is distinctly higher than that of generic Bordeaux wines (-17.3 %). The high impact of Chablis and Côte de Beaune (+48 % and +148.4 % respectively) may, however, appear surprising. Let us also consider the high impact of the Côte de Nuit (+152.1 %) wines. The fact that top quality reds come from both Côte de Beaune and Côte de Nuit whereas Chablis is highly regarded for its Chardonnay, seems to be reflected in the relative contribution of white versus red wines.

Coefficient estimates for grape varieties are striking as consumers appear to value the price premium associated with Chardonnay from any origin more than twice as highly as for Cabernet Sauvignon (+15.3 % and +7.3 % respectively). This occurs on the background of a reverse valuation in terms of colour, and of the fact that both grapes account for the largest proportion amongst red and white wines respectively in the sample. Comparing grape varieties according to colour, the high valuation of Riesling in relation to Chardonnay also seems somewhat surprising, and in particular when considering the highly negative impact on

(22) Classic regions that have a good reputation for their quality are Rioja, Côte Chalonnaise, and Sonoma Valley.

price of Australian Riesling (-34.9 %). However, since Riesling is a rather classical grape in France and Germany, the high valuation might be associated with these countries, whereas Australia is more valued for its Chardonnay.

The highly positive impact of Sauvignon Blanc (+11.7 %) relative to Semillon and Sangiovese is to be expected, especially when taking account of its classical background from the Loire Valley and Bordeaux, and its rising success in the New World. However, it is all the more surprising that the national impact of Sauvignon Blanc is highly negative, both for France (-22.2 %) and Chile (-20 %). The impact of Chardonnay in the case of Spain (+21.2 %) and Italy (-18.7 %) may appear surprising. Given its classical roots in Burgundy and its success in Australia, its significant impact in the case of Spain seems particularly unexpected. However, the negative valuation of Italian Chardonnay seems to support the above suggestion that consumers may not consider Italy as a classical source of 'quality Chardonnay'.

As for red varietals, the highly positive impact of Pinot Noir on prices (+25.7 %) relative to Cabernet Sauvignon (+7.3 %) is not that surprising given the impact of Côte de Nuit, the heartland of Pinot Noir. It was therefore expected that Pinot Noir should show more than just an overall impact, and that consumers would value it regionally, as reflected in its interaction term. This, however, was not the case.

Consumers' valuation of different vintages should be regarded with caution, given the level of aggregation in this all-country model. However, a rather consistent pattern emerges, whereby the increasing valuation of older vintages reflects both interest rate differentials and cost of storage. Nevertheless, the 1986 and 1988 vintages stand out as being particularly valued (+52.4 % and +28.8 %, respectively).

As far as retailers are concerned, consumers value retailer traits as expected in the case of Asda (-11.3 %) and Marks & Spencer (+23.1 %)[23]. However, the rather high impact of Co-op on price is somewhat surprising (+12 %), though it may be partly explained by consumers valuing its long opening hours.

MARKETING IMPLICATIONS

If an attribute is found to explain a positive/negative price deviation from the unit price evaluated at the sample means, it indicates that retailers could investigate profitably the financial gains and losses of

[23] The fact that all wines for which the names of the retailers are given correspond to own-label wines implies that the interpretation of results follows accordingly. Thus, estimates for a given retailer do not refer to all wines sold in this particular chain, but rather to its own-labelled wines.

altering a particular range of wine qualities on offer. Supposing the retailer intends a stock-transfer of French Sauvignon Blanc (FSB) to Chilean Sauvignon Blanc (CSB), a proportionate adjustment to the mean price can be found in three steps. First, we need to identify the proportionate loss for the type of wine that is replaced, and the standard errors involved[24]. Second, in order to model the retailer's intended stock transfer, we need to account for the market share of wine to be removed from the overall sample. Third, we would have to collect the adjusted premium for the affected (grape) variety, and to weight this pivotal variable by results from the first and second steps. The following box (Box 1) aims to demonstrate the implementation of this procedure.

However, it is important to emphasise what assumptions we have made previously, so that we can consider how useful the above results could be in practice. When conducting this hypothetical stock-transfer, we have assumed a single retailer that stocks the entire sample of our wines. Given the above results, we would argue that this retailer could engage in a more efficient optimisation process. A simple price comparison across bottles will not reveal the contribution of the positive/negative price premium for certain attributes to the total change in revenue. In a dynamic setting and with knowledge of demographic variables (scanner data), the retailer would be able to control demand on the attribute level and thus undertake a dynamic optimisation problem that is more efficient than a simple price comparison over time.

We have identified the proportionate adjustment to the overall mean price under the assumption that the retailer can shift the extra wine from Chile without offering special discounts when more wine is purchased, *i.e.* demand is assumed as being perfectly price elastic. This assumption is reasonable if we are considering consumer demand for an individual retailer. When the retailer is replacing a fixed amount of French wine (shelf space) with Chilean wine, consumers' marginal willingness-to-pay for an attribute (combination) will be independent of the number of bottles replaced. Even if this assumption were challenged, the retailer could use price elasticity estimates with knowledge of the demand functions on the attribute level to predict the financial impact of particular stock-transfers.

[24] We could also consider the revenue loss of 19.88 % on sales of French Sauvignon Blanc which, given the standard error of .0429, suggests a 95 % confidence interval of 11.5 % to 28.3 %.

Box 1. Determination of the proportionate adjustment to the mean price when a retailer desires to reshuffle his wine qualities on offer

(STEP 1) It is necessary to identify the proportionate loss for the pivotal attributes that we wish to replace. Therefore, all attributes involved for which explicit coefficients have been estimated have to be identified first. As shown below, we should also account for the certainty of the joint effects, as derived from the variance covariance matrix of the estimated coefficients. However, we first proceed by computing the proportionate loss and the corresponding standard errors in three sub-steps.

1. Find the total sum of the relevant estimated coefficients:

Chile	France	CSB	FSB	TOTAL
- .1103	- .1156	- .2234	+ .2505	= -.1988

2. Compute the corresponding joint standard error (SE), assuming initially that all the parameters have zero covariances:

SE Chile	SE France	SE CSB	SE FSB	TOTAL SE
$\{(.0217)^2$	$+ (.003)^2$	$+ (.0238)^2$	$+ (.0282)^2\}^{1/2}$	= .0429

3. Find the proportionate loss or gain from the log-lin model, considering both all relevant estimated coefficients and the corresponding certainty of the joint effects:

$$g^* = \exp\left[\hat{c} - \frac{1}{2}\hat{V}(\hat{c})\right] - 1, \qquad (4)$$

where $\hat{V}(\hat{c})$ is an estimate of the variance of $\hat{c}$, the coefficient of the dummy. Therefore,

$$g^* = \exp\left[(-.1988) - \frac{1}{2}(0.0429)^2\right] - 1 = -.181$$

If we want to take an estimate of the variance into account, this can be done by pre-multiplying the corresponding segment of the variance-covariance matrix by (1,-1,1,-1), recognising positive and negative correlation between coefficients, and then post-multiplying by the transpose of this unit vector. The corresponding standard error estimate is 11.25. We can confirm this result by considering that for any random variable x and y,

$$\text{var}(x + y) = \text{var}(x) + \text{var}(y) + 2\,\text{cov}(x,y)\,.$$

In our example, we add the covariances of those variables that move together and consider that the variances of all those variables that move into the opposite direction subtract. This will result in the same standard error.

As a result, the proportionate loss accounting for the variance estimate is 18.5% (18.1% from the above), and applies to the market share of the desired attribute bundle (French Sauvignon Blanc).

$$g^* = \exp\left[(-.1988) - \frac{1}{2}(0.1125)^2\right] - 1 = -.185$$

(STEP 2) Identify the market share of the attribute bundle to be removed from the overall sample, hence the retailer's intended stock transfer: In our example, the 184 bottles of French Sauvignon Blanc correspond to 1.27% of the total sample of 14,440 bottles.

(STEP 3) Obtain the adjusted premium for the affected attribute (Sauvignon Blanc) by applying Kennedy's (1981) adjustment (Equation (4)), and weight this pivot attribute (the grape variety) by the results from Step (1) and (2):

The adjusted coefficient for Sauvignon Blanc is + 11.69% (Table 1). As a result, the monetary impact of this stock transfer, hence the proportionate adjustment to the overall mean price, is:

$$1.1169 \times .01274 \times (-.1854) = -.00264\%$$

Given the mean price of 551 pence per bottle, the proportionate adjustment to the mean price would be - 1.45 pence (-1.42, respectively) a bottle, if a stock-transfer of French Sauvignon Blanc to Chilean Sauvignon Blanc was intended.

CONCLUSION

A model of product differentiation is used to investigate the relationship between origin attributes and quality. We obtain information on wine consumer preferences for attributes contained in the label on wine bottles. By means of a parametric approach, implicit prices for these attributes are derived from prices and quantities of wines sold on the British off-licence market.

The results indicate that in some cases consumers attach a high value to the information about these attributes, namely the retailers, that were initially said to have no bearing on the use value of the wines. The corresponding price impacts are considered to be a reflection of differences in retailer traits. Interaction terms are employed in order to reveal the differential effects between attributes, and where these are found to be relevant, consumers are viewed as regarding attribute bundles as imperfect substitutes. Results suggest that this is particularly the case for identical grape varieties originating from different countries and regions within these countries. Therefore, Chardonnay from Spain, Chardonnay from Italy, Sauvignon Blanc from Chile, Sauvignon Blanc from France, Chenin Blanc from New Zealand, Semillon from France and Riesling from Australia are considered as distinctly different attribute bundles. In contrast, the consumer does not appear to value as distinct attribute bundles those grape varieties which originate from countries other than those named above. A highly distinct valuation of grape varieties according to region of origin emerges only for Australia and France. When accounting for the relative importance both of grape varieties and of regional origins, the results suggest an asymmetry between possibly the most classical 'New World' wine producer, Australia, and the most classical 'Old World' wine producer, France.' The willingness of consumers and their ability to differentiate regional differences are most prevalent in the case of France. Consumers also demonstrate a high degree of recognition of regions in the case of Australia, but consumers differentiate even more between grape varieties in the case of Australian wines. Results therefore indicate that grape varieties are highly important in the choice of Australian wines, whereas regional origins are most valued in the case of French wines.

Marketing implications are derived for retailers. Assuming a static setting, retailers can investigate potential costs or benefits of altering a spectrum of wines on offer. Since the analysis is applied to the attribute level, the proposed method has a significant advantage as compared to a direct price observation of wine bottles. Following a particular stock-transfer of wines, we can distinctively identify the reason for the lower price of, say, a bottle of Chilean wine, against other sources of origin.

REFERENCES

ATKINSON (S. E.), CROCKER (T. D.), 1987 — A Bayesian approach to assessing the robustness of hedonic property, *Journal of Applied Econometrics,* 2 (1), pp. 27-45.

BELSLEY (D.), 1986 — Centering, the constant, first-differencing, and assessing conditioning, *in:* BELSLEY (D.), KUH (E.), *Model Reliability*, Cambridge, Mass., MIT Press, pp. 117-153.

BELSLEY (D.), KUH (E.) and WELSCH (R.), 1980 — *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity*, New York, John Wiley.

BERNDT (E. R.), GRILICHES (Z.), 1993 — Price indexes for microcomputers: an exploratory study, *in:* FOSS (M. F.), MANSER (M. E.) and YOUNG (A. H.) (eds), *Price Measurements and Their Uses,* Studies in Income and Wealth, vol. 57, Chicago, National Bureau of Economic Research, pp. 63-93.

BRESNAHAN (T. F.), GORDON (R. J.) 1997— The Economics of New Goods, Studies in Income and Wealth, 58, Chicago and London, National Bureau of Economic Research, University of Chicago Press.

BREUSCH (T. S.), PAGAN (A. R.), 1979 — A simple test for heteroscedasticity and random coefficient variation, *Econometrica,* 47, pp. 1287-1294.

CHAREMZA (W. W.), DEADMAN (D. F.), 1992 — *New Directions in Econometric Practice*, Aldershot, Edward Elgar.

CFCE, 1996 — *Les circuits de distribution en Grande-Bretagne,* Paris, Centre français du commerce extérieur, Direction des produits agro-alimentaires.

CFCE, 1994 — *Vins étrangers en Grande-Bretagne: Analyse de l'offre en linéaire en août 1994,* Paris, Centre français du commerce exterieur, Direction des produits agro-alimentaires.

COMBRIS (P.), LECOCQ (S.) and VISSER (M.), 1997 — Estimation of a hedonic price equation for Bordeaux wine: does quality matter? *The Economic Journal,* 107, pp. 390-402.

COURT (A. T.), 1939 — Hedonic price indexes with automotive examples, *in*: GENERAL MOTORS CORPORATION (eds), *The Dynamics of Automobile Demand*, New York, pp. 99-117.

CROPPER (M. L.), OATES (W. E.), 1992 — Environmental economics: a survey, *The Journal of Economic Literature,* 30 (2), pp. 675-740.

DAVIDSON (J. H.), HENDRY (D. H.), SRBA (F.) and YEO (S.), 1978 — Econometric modelling of the aggregate time-series relationship between consumers' expenditure and income in the United Kingdom, *The Economic Journal,* 88, pp. 661-692.

EIU, 1994 — *Market Survey: Table Wines*, The Economist Intelligence Unit Limited, EIU Retail Business, n° 439, London.

FISHER (F.M.), SHELL (K.), 1971 — Taste and quality change in the pure theory of the true-cost-of-living index, *in:* GRILICHES (Z.) (eds), *Price Indexes and Quality Change*, Cambridge, Mass., Harvard University Press, pp. 16-54.

FOLLAIN (J. R.), JIMINEZ (E.), 1985 — Estimating the demand for housing characteristics: a survey and critique, *Regional Science and Urban Economics*, 15 (1), pp. 77-107.

GOLAN (A.), SHALIT (H.), 1993 — Wine quality differentials in hedonic grape pricing, *Journal of Agricultural Economics,* 44 (2), pp. 311-321.

GOLDBERGER (A. S.), 1968 — The interpretation and estimation of Cobb-Douglas functions, *Econometrica*, 35, pp. 464-472.

GOODMAN (A. C.), 1998 — Andrew Court and the invention of hedonic price analysis, *Journal of Urban Economics,* 44, pp. 291-298.

GORMAN (W. M.), 1980 — A possible procedure for analysing quality differentials in the egg market, *Review of Economic Studies,* 47, pp. 843-856.

GREENE (W. H.), 1997 — *Econometric Analysis*, 3rd ed., Upper Saddle River, Prentice Hall.

GRIFFITHS (W. E.), HILL (R. C.) and JUDGE (G. C.), 1993 — *Learning and Practicing Econometrics*, New York, John Wiley.

GRILICHES (Z.), 1971 — *Price Indexes and Quality Change*, Cambridge, Mass, Harvard University Press.

GRILICHES (Z.), 1961 — Hedonic price indexes for automobiles: an econometric analysis of quality change, *in:* GRILICHES (Z.), *Price Indexes and Quality Change: Studies in New Methods of Measurement*, Cambridge, Mass., Harvard University Press, pp. 55-87.

HALVORSEN (R.), PALMQUIST (R.), 1980 — The interpretation of dummy variables in semilogarithmic equations, *The American Economic Review,* 70, pp. 474-475.

HENDRY (D. F.), 1980 — Econometrics – alchemy or science? *Economica,* 80, pp. 387-406.

HOUTHAKKER (H. S.), 1952 — Compensated changes in quantities and qualities consumed, *Review of Economic Studies,* 19, pp. 155-164.

JONES (L. E.), 1988 — The characteristics model, hedonic prices, and the clientele effect, *Journal of Political Economy,* 96 (3), pp. 551-67.

JUDGE (G. G.), HILL (R. C.), GRIFFITH (W. E.), LÜTKEPOHL (H.) and LEE (T. C.), 1988 — *Introduction to the Theory and Practice of Econometrics*, 2nd ed., New York, John Wiley.

JUDGE (G. G.), GRIFFITH (W. E.), HILL (R. C.), LÜTKEPOHL (H.) and LEE (T. C.), 1985 — *The Theory and Practice of Econometrics*, 2nd ed., NewYork, John Wiley.

JUDGE (G. G.), BOCK (M.), 1983 — Biased estimation, *in*: GRILICHES (Z.) and INTRILIGATOR (M.) (eds.), *Handbook of Econometrics*, vol. 1, Amsterdam, North-Holland, pp. 599-649.

KENNEDY (P. E.), 1986 — Interpreting dummy variables, *The Review of Economics and Statistics,* 69, pp. 174-175.

KENNEDY (P. E.), 1981 — Estimation with correctly interpreted dummy variables in semilogarithmic equations, *The American Economic Review,* 71 (4), p. 801.

KOENKER (R.), 1981 — A note on studentizing a test for heteroscedasticity, *Journal of Econometrics,* 17, pp. 107-112.

LANCASTER (K. J.), 1971 — *Consumer Demand: A New Approach*, New York, Columbia University Press.

LANCASTER (K. J.), 1966 — A new approach to consumer theory, *Journal of Political Economy,* 74 (2), pp. 132-157.

LEAMER (E. E.), 1990 — *Econometrics: The New Palgrave*, New York, Norton.

LEAMER (E. E.), 1983a — Let's take the con out of econometrics, *The American Economics Review*, 73, pp. 31-43.

LEAMER (E. E.), 1983b — Model choice and specification analysis, *in:* GRILICHES (Z.) and INTRILIGATOR (M.) (eds), *Handbook of Econometrics*, vol. 1, Amsterdam, North-Holland, pp. 285-330.

LUCAS (R. E. B.), 1975 — Hedonic price functions, *Economic Inquiry,* 13, pp. 157-178.

MACDONALD (D. M.), 1990 — Flood hazard pricing and insurance premium differentials: evidence from the housing market, *Journal of Risk and Insurance,* 57 (4), pp. 654-663.

MC FADDEN (D.), 1974 — Conditional logit analysis of qualitative choice behavior, *in:* ZAREMBKA (P.) (ed.), *Frontiers in Econometrics*, New York, Academic Press, pp. 105-142.

MUELLBAUER (J.), 1974 — Household production theory, quality, and the hedonic technique, *The American Economic Review,* 64 (6), pp. 977-994.

NERLOVE (M.), 1995 — Hedonic price functions and the measurement of preferences: the case of Swedish wine consumers, *European Economic Review,* 39, pp. 1697-1716.

OCZKOWSKI (E. A.), 1994 — A hedonic price function for Australian premium table wine, *Australian Journal of Agricultural Economics*, 38 (1), pp. 93-110.

OHTA (M.), GRILICHES (Z.), 1986 — Automobile prices and quality: did the gasoline price increases change consumer tastes in the US? *Journal of Business and Economic Statistics,* 4 (2), pp. 187-198.

OHTA (M.), GRILICHES (Z.), 1975 — Automobile prices revisited: extensions of the hedonic hypothesis, *in:* TERLECKYJ (N. E.) (ed.), *Household Production and Consumption,* Studies in Income and Wealth, vol. 40, New York, National Bureau of Economic Research, pp. 325-390.

PAGAN (A.), 1995 — Three econometric methodologies: a critical appraisal, *in*: OXLEY (L.), GEORGE (D. A. R.), ROBERTS (C. J.) and SAYER (S.) (eds), *Surveys in Econometrics*, Oxford, Blackwell, pp. 9-29.

ROBINSON (J.), 1994 — *The Oxford Companion to Wine*, Oxford, Oxford University Press.

ROSEN (S.), 1974 — Hedonic prices and implicit markets: product differentiation in pure competition, *Journal of Political Economy,* 82 (1), pp. 34-55.

SHAPIRO (C.), 1983 — Premiums for high quality products as returns to reputations, *The Quarterly Journal of Economics*, 98 (4), pp. 659-679.

SUITS (D. B.), 1984 — Dummy variables: mechanics V. interpretation, *The Review of Economics and Statistics,* 66 (1), pp. 177-180.

THEIL (H.), 1952 — Qualities, prices and budget enquiries, *Review of Economic Studies,* 19 (3), pp. 129-147.

TRIPLETT (J. E.), 1975 — Consumer demand and characteristics of consumption goods, *in*: TERLECKYJ (N. E.) (ed.), *Household Production and Consumption*, Studies in Income and Wealth, vol. 40, New York, National Bureau of Economic Research, pp. 305-323.

TRIPLETT (J. E.), 1971 — Book reviews: consumer demand: a new approach, *The Journal of Economic Literature,* 11 (4), pp. 77-81.

WALLACE (T.), ASHAR (V.), 1972 — Sequential methods in model construction, *Review of Economics and Statistics*, pp. 172-178.

WAUGH (F. V.), 1928 — Quality factors influencing vegetable prices, *Journal of Farm Economics,* 10, pp. 185-196.

WHITE (H.), 1993 — *Shazam User's Reference Manual Version 7.0*, Second printing, New York, McGraw-Hill.

WHITE (H.), 1980 — A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity, *Econometrica,* 48, pp. 817-838.

APPENDIX I

94 retail outlets distinguished by commercial forms

27 Supermarket outlets	37 Wine specialist outlets	18 Hypermarket outlets	5 Large retailer outlets	7 Others
7 Tesco	4 Wine Rack	6 Asda	2 Littlewoods	1 Coop
3 Co-op	14 Victoria Wines	1 Morrisons	3 Marks &	1 Cullen's
1 Somerfield	3 Unwin's	1 Safeway	Spencer	1 Europa Food
1 Kwiksave	8 Thresher	6 Sainsbury		1 Gateway
6 Safeway	2 Oddbins	1 Scotmid (Coop)		1 Independant
6 Sainsbury	2 Majestic	3 Tesco		1 Kwiksave
3 Waitrose	2 Cellar Five			1 Spar
	1 Bottom's up			
	1 Haddows			

Source: CFCE, 1994

APPENDIX II

SUMMARY STATISTICS

VARIABLE DESCRIPTION	NUMBER OF OBSERVATIONS	MEAN***	STANDARD DEVIATION	VARIANCE	MINIMUM	MAXIMUM
PRICE (£)	14440** (3940*)	(5.51)	(4.5752)	(20.932)	1.09	99.99
RED	6933	0.49137	0.49999	0.24999	0	1
WHITE	7111	0.48655	0.49988	0.24988	0	1
ROSE	396	2.23E-02	0.14779	2.18E-02	0	1
ARGENTINIA	35	2.54E-03	5.03E-02	2.53E-03	0	1
AUSTRALIA	1495	6.95E-02	0.25441	6.47E-02	0	1
GERMANY	801	6.75E-02	0.25094	6.30E-02	0	1
BULGARIA	314	1.93E-02	0.13756	1.89E-02	0	1
CHILE	248	1.78E-02	0.13212	1.75E-02	0	1
SPAIN	1067	6.29E-02	0.24289	5.90E-02	0	1
HUNGARY	281	1.47E-02	0.12045	1.45E-02	0	1
ITALY	1240	8.83E-02	0.2838	8.05E-02	0	1
NEW ZEALAND	502	2.18E-02	0.14614	2.14E-02	0	1
PORTUGAL	485	2.92E-02	0.16835	2.83E-02	0	1
ROMANIA	55	4.06E-03	6.36E-02	4.05E-03	0	1
SOUTH AFRICA	405	3.05E-02	0.17186	2.95E-02	0	1
FRANCE	7062	0.55838	0.49664	0.24665	0	1
CAB SAUV	841	5.00E-02	0.21797	4.75E-02	0	1
CHARDONN	1152	5.71E-02	0.23208	5.39E-02	0	1
CHENIN BLANC	73	4.57E-03	6.74E-02	4.55E-03	0	1
GEWÜRZ-TRAMINER	76	7.36E-03	8.55E-02	7.31E-03	0	1
PINOT NOIR	181	1.29E-02	0.11305	1.28E-02	0	1
RIESLING	227	2.03E-02	0.14106	1.99E-02	0	1
SANGIOVESE	14	1.52E-03	3.90E-02	1.52E-03	0	1
SEMILLON	918	5.25E-02	0.22314	4.98E-02	0	1
SAUVIGNON	712	3.43E-02	0.18193	3.31E-02	0	1
VINTAGE-83	16	2.03E-03	4.50E-02	2.03E-03	0	1
VINTAGE-85	48	4.57E-03	6.74E-02	4.55E-03	0	1
VINTAGE-86	75	9.14E-03	9.52E-02	9.06E-03	0	1
VINTAGE-87	216	1.57E-02	0.12447	1.55E-02	0	1
VINTAGE-88	304	2.79E-02	0.16476	2.71E-02	0	1
VINTAGE-89	608	5.23E-02	0.22263	4.96E-02	0	1
VINTAGE-92	2780	0.20964	0.40711	0.16574	0	1
VINTAGE-94	109	6.35E-03	7.94E-02	6.31E-03	0	1
ASDA	530	3.17E-02	0.17529	3.07E-02	0	1
CWS	118	1.17E-02	0.10743	1.15E-02	0	1
CO-OP	131	1.68E-02	0.12835	1.65E-02	0	1
M&S	216	2.28E-02	0.14942	2.23E-02	0	1

SAFEWAY	329	2.11E-02	0.14362	2.06E-02	0	1
BREEDE	5	1.27E-03	3.56E-02	1.27E-03	0	1
COONAWARA	162	6.35E-03	7.94E-02	6.31E-03	0	1
HUNTER VALLEY	172	5.33E-03	7.28E-02	5.30E-03	0	1
LA MANCHA	61	5.08E-03	7.11E-02	5.05E-03	0	1
RIOJA	301	1.95E-02	0.13844	1.92E-02	0	1
VALENCIA	239	1.40E-02	0.11734	1.38E-02	0	1
VENETO	399	2.84E-02	0.16621	2.76E-02	0	1
DOURO	99	6.35E-03	7.94E-02	6.31E-03	0	1
BORDEAUX	760	5.63E-02	0.23062	5.32E-02	0	1
LANGUEDOC	1617	0.10939	0.31217	9.74E-02	0	1
LIBOURNE	139	1.17E-02	0.10743	1.15E-02	0	1
MEDOC	348	3.32E-02	0.17931	3.22E-02	0	1
PROVENCE	101	9.14E-03	9.52E-02	9.06E-03	0	1
SAUTERNES	32	4.06E-03	6.36E-02	4.05E-03	0	1
CÔTE CHALONNAISE	56	5.84E-03	7.62E-02	5.81E-03	0	1
CÔTE DE BEAUNE	134	1.88E-02	0.13577	1.84E-02	0	1
CHABLIS	176	1.57E-02	0.12447	1.55E-02	0	1
CÔTE DE NUIT	139	1.42E-02	0.11838	1.40E-02	0	1
SONOMA VALLEY	342	5.58E-03	7.45E-02	5.55E-03	0	1

* There are 3 940 unique and hence different bottles in the sample. Corresponding descriptive statistics are in brackets. Since the same unique bottle frequently appears in different outlets, the total sample size is 14,440.

** The difference between the total sum of all observed prices after accounting for replicates [14,440] and the sum of observations for the above attributes as they remained in the final specification is therefore due to:

(1) statistically non-significant attributes

(2) the nature of the data set:

some wines are specified by less attributes than others: (a) indication of the retailer's name from whom the price was collected is only given if the retailer's name appears on the bottle label, or (b) it is due to legal restrictions (*i.e.* EU or national law do not allow to indicate the region of origin or the vintage for certain wines).

*** The sample mean applies to the observations not accounting for replicates, which explains the divergence between the proportion of non-zero's of each attribute in each category (*i.e.* the mean value) and the number of observations.

"E-02" and "E-03" mean 10^{-2} and 10^{-3}, respectively.