Hedonic Wine Price Functions and Measurement Error

EDWARD OCZKOWSKI
School of Management, Charles Sturt University

Accumulated theoretical and empirical evidence suggests that wine prices depend on quality, reputation and objective characteristics. Unlike previous studies, we recognize that quality and reputation are latent constructs and therefore employ factor analysis and 2SLS techniques to consistently estimate hedonic prices in the presence of attributes measured with error. The application to Australian premium wines points to significant reputation effects but insignificant quality effects. It is also illustrated that inappropriately using standard OLS procedures can seriously distort the statistical significance of attributes, the implicit marginal attribute prices, and the predictions of ‘average’ prices for a given set of characteristics.

1 Introduction

A hedonic price function relates the price of a product to its various characteristics, the respective partial derivatives of the function represent implicit marginal attribute prices, Rosen (1974). There appears to be considerable interest in estimating hedonic price functions for wine. Econometric models, regressing price or log(price), against a set of wine characteristics have been estimated for wine sold in various countries including Australia, France, Spain, Sweden and the USA. Wine lends itself particularly well to hedonic price estimation given its highly differentiated nature; for example, in Australia alone, approximately 1200 individual wineries exist, producing over 10,000 different wine products (Winetitles 2000).

A peculiar feature of wine products, not shared by other differentiated products such as houses and automobiles, is that the quality of a wine is difficult to evaluate objectively. The quality of a wine is generally recognized to depend on subjective sensory evaluations and, therefore, cannot be easily or precisely measured. Most previous hedonic wine studies have used sensory wine evaluations of some type as measures of quality in their estimated price functions; see, for example, Combris et al. (1997). Some studies have also employed lagged measures of sensory quality evaluations to reflect a wine’s reputation in the hedonic price function (Landon and Smith 1998).

It is postulated that these previous studies are somewhat deficient, as they employ single measures of quality and reputation and use ordinary least squares (OLS) to estimate price functions. Specifically, these studies fail to recognize the problems caused by measurement error: OLS produces inconsistent estimates given the presence of regressors which contain measurement error. Our paper overcomes this deficiency by treating quality and reputation as latent constructs, which can only be reflected by multiple observed indicators. Confirmatory factor analysis (CFA) can be employed to assess empirically the ability of observed indicators to reflect latent constructs adequately. The results from a CFA can then supplement the consistent estimation of a hedonic price function which recognizes the existence of
latent constructs for quality and reputation. Bollen’s (1996) two-stage least squares (2SLS) estimator for latent variable models is employed to produce consistent estimates of implicit marginal attribute prices.

The rest of the paper is organized as follows. Section II briefly reviews previous hedonic wine price functions and examines the latent constructs of quality and reputation for Australian premium table wine using CFA. Section III describes the 2SLS methodology, and then presents and discusses the estimates for the hedonic price function. Section IV concludes.

II Hedonic Wine Price Functions, Quality and Reputation

In general, four types of characteristics have been employed in hedonic wine functions: sensory, chemical, objective and climatic. Sensory attributes are subjectively assessed measures such as a wine’s aroma, body and firmness (Combris et al. 1997). Chemical attributes represent technical measures such as a wine’s sugar and acid level, (Nerlove 1995). Objective attributes are wine traits which are easily recognizable by the consumer and appear on the label, these include the wine’s year of vintage, region from which the grapes are sourced and the grape variety (Oczkowski 1994). Climatic attributes are those which measure the effects of weather on the grapes used in wine production (Byron and Aschenfelter 1995).

Oczkowski (1994) and Landon and Smith (1997, 1998) argue that a key assumption which should guide the choice of employed characteristics is the perfect information assumption of the Rosen (1974) framework. The extent to which consumer preferences can help to determine market prices depends on their knowledge of the wine’s characteristics. This implies that individual sensory, chemical and climatic characteristics are unattractive candidates for inclusion within hedonic price functions and, instead, objective easily observable characteristics should be preferred. This argument does not, however, preclude the use of overall sensory ‘quality’ score measures from widely accessible published wine guides. These guides improve the information flow to consumers about a wine’s quality and, therefore, they may serve to reflect demand side influences on prices. However, Landon and Smith (1997, 1998) and Wade (1999) argue that, because of the contemporaneous nature of published wine guides and price determination, consumers may not possess information on current wine quality before market price determination. In response, it is proposed that consumer decisions are based on a wine’s longer-term reputation and that these reputation preferences influence price determination.

In summary, these are the main empirical findings from hedonic wine price functions:

1 When both individual sensory and objective characteristics are examined, the former tends to be insignificant and the latter significant (Combris et al. 1997, 2000).
2 When wine-guide overall sensory current quality scores are employed with objective characteristics, quality scores and objective traits are significant (Golan and Shalit 1993; Oczkowski 1994; Landon and Smith 1997; Schamel et al. 1998; Wade 1999; Angulo et al. 2000; Combris et al. 2000).
3 When reputation (lagged quality) sensory measures appear with objective characteristics, reputation scores and objective traits are significant (Landon and Smith 1997; Wade 1999).
4 When both reputation and current quality appear with objective characteristics, reputation, quality and objective traits are significant but reputation is far more economically important than quality (Landon and Smith 1998).

This accumulated theoretical and empirical evidence suggests the following general hedonic price function for wine:

\[ \text{Wine price} = f(\text{quality, reputation, objective characteristics}) \]  

A questionable feature of the previous research, however, relates to the use of single summary measures of quality and reputation in the standard regression framework. There is substantial evidence to suggest that expert wine tasters’ evaluations do differ (Cliff and King 1996) and, therefore, these concepts are better treated as latent constructs. This line of argument invalidates the use of OLS and single measures of quality and reputation, these being contaminated by measurement error. Measurement error in re-

---

1 For a comprehensive critique of some estimated hedonic wine price functions, see Unwin (1999).
Regressors leads to correlation between regressors and the error term and, therefore, inconsistent OLS estimates; see, for example, Davidson and Mackinnon (1993, p. 211). The analysis of latent variables can be usefully conducted through CFA; see, for example, Bollen (1989, ch. 7). Unlike unrestricted exploratory factor analysis, CFA permits the explicit specification of an a priori restricted factor structure for the latent constructs.

In Australia, three guides – the Penguin guide (Kyte-Powell and Hooke 2000), Oliver (2000) and Halliday (2000) – provide quality scores based on the international show judging system which allocates 50 per cent of the score for palate, 35 per cent for nose and 15 per cent for colour. Rather than employ short-term lagged quality scores as measures of reputation as in Landon and Smith (1997, 1998) and Wade (1999), we consider long-term quality scores on wines that are independent of annual vintage variations. The wine guides of Oliver (2000) and Bradley (2000) provide single score measures of the reputation for each wine over numerous vintages. For more details on the data, see the data appendix.

Conceptually, one expects a high correlation between quality and reputation, and this is permitted in the CFA. Factor cross loadings and correlated measurement errors are ruled out a priori. The specification of the CFA model is

\[
Q_i = \lambda_0 + \delta_{Q_i} \quad i = 1, \ldots, 3
\]

\[
R_i = \lambda_0 + \delta_{R_i} \quad i = 1, 2
\]

where, \( Q \) and \( R \) are the latent constructs of quality and reputation respectively; \( Q_i \) and \( R_i \) are the observed (wine guide) indicators of \( Q \) and \( R \) respectively; \( \lambda_i \) are estimable factor loadings; and \( \delta_i \) are independently and identically distributed (i.i.d.) measurement error terms with zero mean and constant variance.

The CFA was performed by EQS software (Bentler 1995). Descriptive statistics of the measures are provided in the data appendix; correlations between the five indicators range from 0.34 to 0.65. Maximum likelihood estimates robust to non-normality are reported given that some of the measures exhibited excessive kurtosis. CFA results are reported in Table 1.

In general, the goodness-of-fit measures indicate acceptable overall fit with both the CFI and GFI exceeding 0.95. The GFI test rejects the null hypothesis of a valid specification, however; as has been well documented (Hair et al. 1998, p. 655), the test too frequently rejects models for sample sizes exceeding 200. All factor loadings are acceptably high and statistically significant. As indicated by the factor loadings and reliability measures, the quality construct is less precisely measured than the reputation construct. In other words, the tasters' evaluations of quality are less similar than those of reputation. The highest standardized factor loadings

<p>| TABLE 1 |</p>
<table>
<thead>
<tr>
<th>Confirmatory Factor Analysis Estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quality</td>
</tr>
<tr>
<td>---</td>
</tr>
<tr>
<td>Penguin</td>
</tr>
<tr>
<td>Halliday</td>
</tr>
<tr>
<td>Oliver</td>
</tr>
<tr>
<td>Quality</td>
</tr>
<tr>
<td>Variance extracted</td>
</tr>
<tr>
<td>Composite reliability</td>
</tr>
<tr>
<td>Variance extracted</td>
</tr>
<tr>
<td>Composite reliability</td>
</tr>
<tr>
<td>GFI Test ( \chi^2(4) )</td>
</tr>
<tr>
<td>CFI</td>
</tr>
</tbody>
</table>

Notes: \( N = 276 \). Standardized estimates are presented. Robust \( t \)-ratios are presented in parentheses. Robust estimates are presented for the GFI (goodness-of-fit test) and CFI comparative fit index. GFI is the goodness-of-fit index. A discussion of all statistics and concepts can be found in Hair et al. (1998).
occur with the Oliver measures for both quality
and reputation.

III Estimation Methods and Results

Section II established reasonable support for
the observed indicators of the latent constructs,
quality and reputation. To consistently estimate
the relationship specified in (1), we make use of
Bollen’s (1996) 2SLS estimator for latent variable
models. The technique is surprisingly simple and
permits the use of econometric developments in
2SLS estimation and testing, such as hetero-
scedasticity correction and non-nested model and
specification testing.

Bollen (1996) proposes that, for each latent
construct, a single observed indicator be chosen
as the reference or scaling variable; this serves
to set the metric for the construct. A scaling vari-
able must be specified to ensure identification
and operationalize the latent construct. The latent
variable has the same units of measurement
(on average) as the scaling variable, (Bollen
1989, pp. 239–40). The scaling variable enters
as a regressor and the remaining indicators
serve as instruments in a 2SLS regression.
The non-scaling indicators are valid instru-
mants as they are expected to be highly correlated
with the scaling variable given acceptable
composite reliability, and by construction they
are not theoretically correlated with the error
term if measurement errors of the indicators
are uncorrelated. The latent variable model for
(1) is

$$\log(\text{Price}) = \beta_0 + \beta_1 Q + \beta_2 R + \beta_3 Z + u$$

(4)

where $Z$ are objective characteristics measured
without error, $u$ is a classical error term which
is assumed to be uncorrelated with the measurement
errors in (2), $\text{cov}(u, \delta_i) = 0$, and all other terms
are defined as for (2) and (3). Assume that $Q$, and
$R_i$ are identified as the scaling variables (i.e.,
$\lambda_{\delta_1} = \lambda_{\delta_2} = 1$) then (2) and (4) can be rewritten in
observable variables as

$$\log(\text{Price}) = \beta_0 + \beta_1 Q + \beta_2 R + \beta_3 Z + e$$

(5)

where $e = u - \beta_1 \delta_1 - \beta_2 \delta_2$. For consistent
estimation, OLS cannot be applied to (5) because
the error term is correlated with the regressors, e.g.,
$\text{cov}(e, Q) \neq 0$. Consistent estimates for (5) can be
obtained by applying 2SLS using as instru-
ments $(Q, Q, R, Z)$.

A feature of Bollen’s procedure is that 2SLS esti-
mates will vary given different selected scaling
variables. As a consequence, some set of criteria
or method is needed to choose the ‘optimal’ set
of scaling variables. At least four alternative
approaches have been recommended or can be
suggested. First, it is often suggested that the
indicator which best reflects the construct theo-
retically is the appropriate choice; see, for example,
Bollen and Paxton (1998, p. 286). In our wine
example, one might argue that a particular wine
guide is superior to others given the qualifications
and experience of the expert(s). Clearly, however,
any such choice is highly subjective. Second, in
the CFA, the item with the highest standardized
factor loading best reflects the construct empiri-
cally and therefore should be used as the scaling
variable. In our case, this suggests that Oliver’s
quality and reputation measures should be chosen
(see Table 1). Third, one could regress each
observed indicator in turn against all other instru-
ments and choose as scaling variable that with the
largest $R^2$ (Bollen and Paxton 1998, p. 287). The
rational for this method is that the higher the corre-
lation between the regressors and the instru-
mants in a 2SLS regression, the more asymp-
totically efficient the estimator (Harvey 1981, pp.
79–80).

An unattractive feature of the previously sug-
gested empirical approaches for selecting a
scaling variable is that choice is made without
regard for its purpose, that is to be employed in a
hedonic price function. In other words, factor
loadings and $R^2$s on other instruments do not
employ price and hedonic model specification
information. This suggests that a fourth approach
to choosing an appropriate set of scaling variables
is to estimate all possible hedonic price models
($6 = 3$ quality $\times 2$ reputation) with alternating
scaling variables. A choice between models (and
hence scaling variables) can be made empirically
using standard criteria such as goodness-of-fit
measures, specification and over-identifying test
statistics.

We will make comment on all four ap-
proaches and therefore initially present 2SLS es-
timates for models (1)–(6) with alternating
scaling variables in Table 2. The data employed
in these regressions is described in the data
appendix. Unlike previous Australian studies
such as Oczkowski (1994) where numerous
objective characteristics also appear in regres-
sions, the relative small size of our data set does
not permit us to identify significant wine variety

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
<table>
<thead>
<tr>
<th>Model</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
</tr>
</thead>
</table>
| Quality Reputation Method | Penguin | Oliver | 2SLS | Penguin | Bradley | 2SLS | Oliver | Bradley | 2SLS | Halliday | Oliver | Bradley | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Oliver | 2SLS | Halliday | Olive...
or regional impacts. Of all the objective characteristics evaluated only the red wine, old vintage and year of marketing distinctions proved to be statistically significant in most models.

The 2SLS estimates are presented with a series of goodness-of-fit and diagnostic test measures. Consistent with some previous studies, the results of the RESET and RESET (linear) tests demonstrate that a log-linear specification is preferred to a linear specification. Some of the key features of models (1)-(6) include:

1. Acceptable goodness-of-fit values, consistent with previous studies
2. The absence of significant specification error (RESET) and heteroscedasticity (Het)
3. Reputations significant in five of the six models
4. Quality insignificant in all models
5. Red wine and old vintage significant in most models

In comparing models (1)-(6), the OIR test clearly rejects models (1) and (3) implying that these models have either inadequate specifications or instruments. Of the remaining models, models (5) and (6) have the highest $R^2$ values and therefore are probably preferred. Non-nested encompassing tests for IV regressions developed by Smith (1992) were performed to choose between models (5) and (6). When model (5) is the null $t = -1.11$ and when model (6) is the null $t = -1.48$. Therefore neither model rejects the other and hence both appear to be equally acceptable specifications. It is noteworthy that both preferred models employ the Halliday quality measure as scaling variable.

A clear difference between these findings and those of Landon and Smith (1998) is the statistical insignificance of the quality variable. It would be important to ascertain whether this difference may be due to the technique of estimation employed. To investigate this issue, counterpart OLS models were estimated for models (1)-(6) using the identified scaling variable as the regressor. In four of the six OLS models, quality becomes statistically significant (quality remains insignificant when Oliver quality is employed) and generally greater significance was attained for all other regressors. As an example, model (7) in Table 2 is the OLS counterpart of model (5), the improved significance of quality and other variables is apparent. These results imply that the Landon and Smith (1998) finding, that both quality and reputation are statistically significant, may be due to the estimation technique and the general notion that OLS 'fits too well' when employed with regressors which contain measurement error. However, our 2SLS findings are consistent with the result that reputation is far more economically important and statistically significant than quality in hedonic wine price models.

It is clear from the CFA results that multicollinearity problems may exist in models (1)-(6) given the high correlation between quality and reputation. Models (8)-(12) in Table 2 investigate this issue by dropping the reputation and quality variables in turn. Models (8)-(10) confirm previous findings that when quality appears alone, it is statistically significant, even given the explicit recognition of measurement error. Models (11)-(12) illustrate the same finding for reputation. Comparing the $R^2$'s between models (8)-(10)

---

2 The Oczkowski (1994) study covered the years 1991 and 1992 and included 1604 wines from the Penguin guide only. Given the need to identify wines which were evaluated by all four wine guides the number of data points is substantially reduced for a two-year period to 276 wines. In addition to examining variety and regional impacts, measures of cellaring potential and the number of cases produced by the wine producer were also found to be statistically unimportant. The overall diagnostic test results, to be presented later, suggest that the omission of other objective characteristics such as variety and region have not had a severe adverse affect on the specification of the hedonic price function.

3 An augmented version of the encompassing test was employed where the testing regressor variable is also included in the instrument set. The augmented version of the encompassing test appears to have superior finite sample empirical size and power properties than the encompassing and Cox tests, see Oczkowski (2001).

4 The 1st-stage regression $R^2$'s, for the scaling variables are: model (1) quality 0.24, reputation, 0.57, model (2) (0.28, 0.48), model (3) (0.28, 0.52), model (4) (0.39, 0.47), model (5) (0.33, 0.59) and model (6) (0.33, 0.59). Even though not definitive, these values do also point to the general superior performance of models (5) and (6).

5 Note however, model (7) exhibits some specification error and the original model exhibits significant heteroscedasticity. The improved parameter significance is not due to the heteroscedastic adjustment as the unadjusted $t$-ratios also point to improved significance.
against models (11)-(12) points to the superior performance of the reputation only models. Comparing the GR's between models (11) and (12) suggests that model (11) using Oliver reputation is preferred. In terms of explanatory power, models (11)-(12) are comparable models to (5)-(6) and therefore little, if any, explanatory power is lost when omitting the insignificant quality variable. Once again, models (13)-(14), the OLS counterparts for models (11)-(12), illustrate the 'over-fitting' of OLS through the greater significance of all variables.

The marginal (dollar and percentage) effects of the characteristics on wine prices are presented in Table 3 for the two preferred reputation models (11) and (12). For comparison purposes, results for the counterpart OLS models (13) and (14) are also presented. Interestingly, the alternative reputation models using 2SLS produce different results for the objective characteristics, red wine is more economically important in the Oliver model and old vintage is more economically important in the Bradley model. The differences between the 2SLS and OLS reputation marginal impacts are large; in most cases, the 2SLS marginal reputation impacts are more than 50 per cent higher than their OLS counterparts.

One of the important uses of hedonic functions for wine consumers is their ability to identify bargains and over-priced wines. Effectively, the estimated function provides a guide to the average wine price for the stated characteristics. Actual wine prices 'below' the regression plane (negative residuals) represent bargains, that is, the retail price is less than the average price as predicted by the model. The reverse argument applies for over-priced wines. To investigate wine prediction differences due to reputation measure and estimation technique, we consider three prediction examples in Table 4. The differences due to reputation measure and estimation technique can be substantial. For example, for the cheapest wine ($9.00), the Oliver 2SLS model predicts a slight under-charging while the Oliver OLS model implies it is a significant bargain buy. For the average priced wine ($32.40), the Oliver 2SLS model suggests a substantial bargain is to be had, while the Bradley 2SLS model suggests it is slightly over-priced. All models suggest that the highest priced wine ($275) is substantially over-priced, but to varying degrees.

IV Conclusion

This paper has presented consistent hedonic price estimates despite the existence of characteristics which can only be measured with error. An important implication of the results relates to the use of OLS in the presence of imprecisely measured hedonic attributes. The impact can manifest itself through:

1. erroneously identifying statistically significant attributes (wine quality, Table 2);
2. estimating substantially different implicit marginal attribute prices (Table 3), and
3. inaccurately predicting 'average' prices for a given set of characteristics (Table 4).

These impacts cast a serious warning about employing OLS techniques for wine hedonic price models.

Various criteria were considered to determine the appropriate set of scaling variables for the 2SLS estimator. In general, the criteria recommended different models; however, on the basis of conventional regression model goodness-of-fit and diagnostic testing procedures, the reputation-only models were considered most appropriate. In choosing between reputation models, consumers may have a preference for a particular wine guide, given the experience and qualifications of the expert taster. In other words, the preference for a particular wine guide represents an example of a theoretical choice for the observed indicator which best reflects the reputation construct.

In conclusion, a useful exercise would be an investigation into whether the presented general results hold for other countries. In particular, are the insignificant quality impacts and the significant reputation impact primarily a reflection of
the relatively low composite reliability of the quality measure compared to the reputation measure? Also, to what extent are the OLS estimates of wine attribute prices for other countries affected by measurement error? Finally, the described general technique might also usefully be employed for estimating price functions for other differentiated products whose attributes are difficult to precisely measure, for example, restaurant meal prices (Gunawardana and Harvrla 1996) and recorded music (Harchaoui and Hamdad 2000).

### DATA APPENDIX

Data relates to dry red and white wines from wine guides published during: 1999 (N = 132) and 2000 (N = 144). The data is sourced from the Penguin Guides: Hooke and Kyte-Powell (1999) and Kyte-Powell and Hooke (2000), unless stated otherwise.

**Price:** Recommended retail price (mean = 32.41, std dev. = 28.2, range: 9–275).

### QUALITY:

**Penguin** (mean = 4.40, std dev. = 0.58, range: 2–5, possible range: 0–5, increments: 0.5).

**Halliday** (mean = 89.61, std dev. = 3.70, range: 81–97, possible range: 0–100, increments: 1.0). **Source:** Halliday (1999, 2000).

**Oliver** (mean = 17.38, std dev. = 1.07, range: 15–19.5, possible range: 0–20, increments: 0.1). **Source:** Oliver (1999, 2000).

**Bradley** (mean = 4.13, std dev. = 0.81, range: 2–6, possible range: 0–6, increments: 1.0). **Source:** Bradley (1999, 2000).

**Note:** 5 golden stars allocated a 6 rating. This measure ‘reflects some of the earned respect for a particular label over the years of its production’ Bradley (2000, p. iv).

**Oliver** (mean = 2.61, std dev. = 1.14, range: 0–5, possible range: 0–5, increments: 1.0). **Source:** Oliver (1999, 2000). **Note:** an unrated wine is allocated a 0 rating. Published scores have been reversed coded. This measure ‘is a reflection of each wine’s performance over the last four years’ Oliver (2000, p. 7).

### RED WINE:

- Dry red wine = 1, dry white wine = 0 (mean = 0.59, std dev. = 0.49)
- Vintage 1995 or earlier = 1, otherwise = 0 (mean = 0.04, std dev. = 0.20)

### OLD VINTAGE:

- Marketed year 2000 = 1, otherwise = 0 (mean = 0.52, std dev. = 0.50)
REFERENCES


