PRODUCT CHARACTERISTICS AND REPUTATION EFFECTS IN THE WINE MARKET

By

MARCO COSTANIGRO

A dissertation submitted in partial fulfillment
of the requirements of the degree of

DOCTOR OF PHILOSOPHY

WASHINGTON STATE UNIVERSITY
School of Economic Sciences

MAY 2007

To the Faculty of Washington State University:

The members of the Committee appointed to examine the dissertation of MARCO COSTANIGRO find it satisfactory and recommend that it be accepted.

Chair

ACKNOWLEDGEMENTS

I would like to thank my committee members and the faculty of the School of Economic Sciences for guiding me in this learning process. Thanks to their outstanding academic preparation, careful mentoring and enriching human interaction, I leave Washington State University as a better person than I was five years ago. Dr. Jill McCluskey has been the best chair I could wish for, and that is said without exaggeration. My unlimited gratitude goes also to Dr. Ron Mittelhammer and Dr. Jonathan Yoder, for keeping their doors open for me at all times.

A big “grazie” to my family in Italy who always supported me and encouraged me to do the best I could, even when my decisions took me far away from home. To Lisa I owe more than gratitude. Thank you for your patience and understanding. I am excited to continue this journey with you at my side.

PRODUCT CHARACTERISTICS AND REPUTATION EFFECTS
IN THE WINE MARKET

Abstract

by Marco Costanigro, Ph.D.
Washington State University
May 2007

Chair: Jill J. McCluskey

This dissertation analyzes the relationships between wine attributes and prices, with a focus on reputation effects. Contributions are made in the fields of industrial organization and econometrics, developing a model of firm behavior in the presence of collective and individual reputation incentives, and a technique broadly applicable to the task of estimating class-specific parametric models in the presence of class uncertainty. Data from California and Washington wines are analyzed.

In a dynamic optimization framework, a theoretical model analyzes the firm's choice in maximizing the present value of its profits in a market in which the return of investing in quality is two-fold: collective (associated with the region of production) and firm reputation (associated with the brand or label). The results indicate that markets with fewer firms with both collective *and* firm reputation are conducive to the highest levels of quality.

The empirical part of the dissertation analyzes the effect of wine attributes on prices using hedonic models, while taking account of extreme product heterogeneity. It is hypothesized that multiple product classes exist. To identify and estimate class-specific hedonic models, two approaches are taken. The first approach uses price to segment the wine market, while the second uses all information to segment the market. In the price-segmented model, accounting for multiple wine classes results in a greater ability to explain the variability in the data and produces more accurate and interpretable results regarding the implicit prices of the attributes.

For the latter application, an innovative econometric technique is developed. First, a hedonic model for wine is estimated nonparametrically via local polynomial regression. Differences in the hedonic function across neighborhoods of data reflect changes in the underlying supply and demand functions. Data are then aggregated into groups of observations that share functionally similar estimates of the (local) hedonic functions. In this way, wine segments are endogenously determined on the basis of similarities in market equilibria. Using this methodology, four differentiated wine markets are identified: commercial, semi-premium, premium, and ultra-premium. Finally, parametric hedonic functions specific to each wine class are estimated, revealing significant differences in implicit prices of the attributes across classes.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

To my grandfather Francesco, who is not anymore with me,

to Lisa, who will always be with me,

and to my child, who will come meet us soon

A mio nonno Francesco, che non è più con me,

a Lisa, che sará sempre con me,

e a mio figlio o figlia, che sará presto con noi

CHAPTER ONE

INTRODUCTION

The international wine market is changing rapidly. Historical producers from Europe are facing increasing competition from “new world” countries such as the United States, Chile, and Australia (Yue *et. al.*, 2006). Consumer preferences for wine have been changing in the latest decades, and sales of premium wines have increased at the expenses of the lower-quality wines. The rising popularity among wine drinkers of specialized magazines such as *The Wine Spectator* or *Decanter* likely contributed to the recent changes in consumer preferences. Wine quality, producer reputation and the ability to meet consumer’s demands are the likely grounds on which producer’s successes or failures will be determined.

This dissertation focuses on two general questions that are particularly relevant to wine markets: First, what is the role of producer’s (brand or label) reputation relative to collective (region of production) reputation in determining product quality? Second, which wines compete with each other? Beyond a certain level of differentiation, two wines are no longer substitutes, and consumers consider them to be different products.   The answers to these questions will provide insights to interpreting the current trends in the wine market and offer producers information relevant to determining their strategies.

Quality and Reputation

Several studies have investigated the issues related to the establishment of producers' reputations for quality when consumers have imperfect information. For many products, referred to as experience goods, quality cannot be assessed until after consumption. Therefore, the producer's reputation plays a significant role in directing consumer choices. Reputation can be associated with the individual producer (firm reputation) or a group of producers or region (collective reputation). The characteristics of the market for a given good determine which kind of reputation is developed. In many cases, as in the wine market, both collective and firm reputations play a relevant role: wine producers are known for their label and their region of production.

The literature that focuses on the effects of reputation on quality choices is relevant. Shapiro (1982) examined the quality choice of a monopolist when consumer cannot observe all the relevant attributes of a product before purchase. He found that under such circumstances, the quality choice of the monopolist is lower than under the perfect information setting. Tirole (1996) modeled collective reputation as a function of the quality output of the individual producers and showed that low-reputation can be long lasting and bad reputation-steady states can be difficult to move away from, even when new producers attempt to increase quality.

Winfree and McCluskey (2005) investigated the public good aspect of collective reputation with an application to agricultural products. Using a dynamic optimization framework, they showed that with positive collective reputation and no traceability, there is an incentive to extract rents by producing at lower quality levels. Furthermore, they showed that the sustainable level of collective reputation decreases as the number of firms in the production district grows larger and proposed the implementation of minimum quality standards to sustain collective reputation. Carriquiry and Babcock (2007) further elaborated on the use of quality assurance systems and their effects on the equilibrium quality level under different market structure scenarios. They concluded that monopolists are more likely to invest heavily on quality, as they can capitalize the full return from investing in reputation.

Hedonic Valuation of Wines

The literature seeking to identify the determinants of wine prices using hedonic techniques is well established. A considerable amount of work has been done to determine which wine attributes affect wine prices: Combris *et al*., (1997, 2000) showed that when regressing objective and sensory characteristics on wine prices, the objective cues (such as expert rating score and vintage) are significant, while sensory variables (such as tannins content and other measurable chemicals) are not. Much of the literature (Oczkowski, 1994; Landon and Smith, 1997; Schamel and Anderson, 2003, Angulo *et al*., 2000) indicates that ratings by specialized magazines are significant and should be included in modeling wine prices. Possible explanations for the insignificance of sensory cues are the difficulty of isolating the effect of each chemical on the final flavor and smell and that only a small percentage of wine purchasers are connoisseurs. Therefore, expert ratings act as a signal to the consumer. It is uncertain whether expert ratings influence prices because they are good proxies for quality of the wine or because of their marketing effect. The region of production, capturing production costs differentials and the effects of the collective reputation of the district, and the vintage are often reported as significant variables (Angulo *et al.*, 2000; Schamel and Anderson, 2003).

Dissertation Format and Content

The format of this dissertation is three related but stand-alone articles. The first article (Chapter Two) develops a theory of producer's quality choice when there are returns to both collective and firm reputation and is broadly applicable to experience goods. While previous studies considered either collective or firm reputation, the scope of this research is broader. The model is at first derived under the setting of producers benefiting from collective (associated with the region of production) *and* firm's reputation (associated with the brand or label) and compared to the cases of markets in which firms develop only collective *or* firm's reputation. A quality response function to changes in reputation for a representative firm is then derived, and the equilibrium quality in the industry is obtained for the case of a duopoly.

The second article (Chapter Three) contributes to the hedonic literature for wine and highlights the extreme heterogeneity of the product. Historically, economists have been estimating a single hedonic function for any red or white wine. However, estimating a single hedonic price function imposes the assumption that the implicit prices of the attributes are the same for any red or white wine. In this Chapter, I investigate the hypothesis that multiple wine markets might exist and that disregarding this heterogeneity might yield to aggregation bias in the estimated implicit prices.

The third article (Chapter Four) builds on the findings of Chapter Three and is aimed at improving the econometric approach by including all information is determining product class membership. In this final article, I develop and illustrate an innovative technique designed to estimate class-specific parametric models when the class membership of each data point is uncertain.

Summary of Findings

Chapter Two shows that, if collective and firm reputation are additive, markets characterized by both kinds of reputation are conducive to higher level of quality than markets with own or collective reputation only. Regarding the dynamics of quality choices, I find that 1) more visible firms have higher incentives to invest in quality; 2) the higher is the number of firms producing under a given common appellation, the lower is the resulting quality level; and 3) when collective reputation is present, there is a positive externality to invest in quality.

In Chapter Three, the relationship between wine prices and attributes is estimated using a hedonic model. By estimating hedonic functions specific to price ranges, I show that the wine market is segmented into several product classes or market segments: commercial wines, semi premium wines, premium and ultra premium. Since the implicit prices of the attributes are different across wine classes, the segmented model shows greater ability to explain the variability in the data and produces more accurate and interpretable results. Conversely, models that do not account for the existence of classes are shown to be biased and imprecise.

The contribution of Chapter Four is both methodological and empirical in nature. With the purpose of estimating class-specific models in which wine classes are not identified solely on the basis of price ranges, a broadly applicable technique is developed. The procedure, which I call *local polynomial regression clustering,* first estimates a hedonic model nonparametrically via local polynomial regression and then aggregates data into data clusters that share functionally similar estimates of the (local) hedonic functions, identifying product classes based on similarities in market equilibria. Finally, parametric hedonic functions specific to each product class are estimated. This procedure allows identifying and characterizing wine classes on the basis of multiple attributes, in addition to producing wine-class specific estimates of the implicit prices (which are qualitatively similar to the ones obtained in Chapter Three).

References

Angulo, A.M., J.M. Gil, A. Gracia and M. Sanchez. (2000). "Hedonic Prices for Spanish Red Quality Wine," *British Food Journal,* 102(7):481-493.

Carriquiry, M. and B. Babcock. (2007). "Reputations, Market Structure, and the Choice of Quality Assurance Systems in the Food Industry," *American Journal of Agricultural Economics*, 89(1):12-23.

Combris, P. and S. Lecocq and M. Visser. (2000). "Estimation of a Hedonic Price Equation for Burgundy Wine," *Applied Economics* 32:961-967.

Combris, P., Sebastien L. and Visser, M. (1997). "Estimation of a Hedonic Price Equation for Bordeaux Wine: Does Quality Matter?" *Economic Journal,* 107(441):390-403.

Landon, S. and C.E. Smith. (1997). "The Use of Quality and Reputation Indicators by the Consumers: The Case of Bordeaux Wine," *Journal of Consumer Policy,* 20:289-323.

Oczkowski, E. (1994). "Hedonic Wine Price Function for Australian Premium Table Wine," *Australian Journal of Agricultural Economics,* 38:93-110.

Schamel, G. and K. Anderson. (2003). “Wine Quality and Varietal, Regional and Winery Reputations: Hedonic prices for Australia and New Zealand,” *Economic Record* 79:246.

Shapiro, C. (1982). “Consume Information, Product Quality, and Seller Reputation” *The Bell Journal of Economics*, 13(1):20-35

Tirole J. (1996). “A Theory of Collective Reputations (with Applications to the Persistence of Corruption and to Firm Quality),” *Review of Economic Studies*, 63(1):1-22

Winfree, J. A. and J. McCluskey. (2005). “Collective Reputation and Quality,” *American Journal of Agricultural Economics*, 87:206-213

Yue, C., S. Marette and J. Beghin. “How to Promote Quality Perception in Wine Markets: Brand advertising or Geographical Indication? *Working paper* 06-WP 426. August 2006.

CHAPTER TWO

A THEORY ON QUALITY, COLLECTIVE AND FIRM REPUTATION

Chapter Abstract:

When product quality is uncertain, producer's reputation plays a relevant role in directing consumer choices. Depending on the characteristics of a market, firm reputation, collective reputation, or a combination of both develops. This article presents a theoretical model of firm's incentives when the returns from investing in quality are two-fold: collective reputation and firm reputation. The general model is then modified to consider the special cases of collective reputation only or firm reputation only. Under each scenario, a quality response function to changes in reputation is derived for a representative firm, and the industry equilibrium quality is solved for. I find that markets with a small number of highly visible firms which develop collective and firm reputation are conducive to the highest equilibrium levels of quality.

Key Words: Quality, Collective Reputation, Firm Reputation

Introduction and Literature Review

A significant body of literature has investigated the issues related to the establishment of producer's reputation for quality when consumers have imperfect information. For many products, referred to as experience goods, quality cannot be assessed until after consumption. Therefore, the producer's reputation for quality plays a significant role in directing consumer choices. Reputation can be associated with the individual producer (firm reputation) or a group of producers or region (collective reputation). The characteristics of the market for a given good determine which kind of reputation is developed.

When producers are not traceable, collective reputations develop, generally associated with the region of production (referred to as production district hereinafter). This is often the case with agricultural products, two relevant examples being Washington state apples and Idaho potatoes. For these goods, consumers are willing to pay price premia, as these production districts have a reputation for high quality. We often observe the exclusive development of firm reputation in markets in which the location of production is irrelevant, like is the case of firms delivering services. In many cases, both collective and firm reputations play a relevant role. A classic example is wine: producers are known for their label, and their region of production. Another example is the automotive sector: Japanese car producers have developed firm-specific reputations, associated with particular brands, and a general recognition for reliability, which is collective and associated with the country of origin.

A substantial difference between these two examples deserves to be highlighted. While Beringers Cellars, a major wine producer in the California Napa Valley, is likely to recognize the reputation of its production districts as an asset, and will probably take actions to preserve it, it seems unlikely that Honda will place considerable weight on the reputation of the Japanese car industry as a whole when making decisions about the quality of their products.

A stream of literature that focuses on the effects of reputation on quality choices is relevant to the subject matter just introduced. Shapiro (1982) examined the quality choice of a monopolist when consumer cannot observe all the relevant attributes of a product before purchase. He finds that under such circumstances, the quality choice of the monopolist is lower than under the perfect information setting. Tirole (1996) models collective reputation as a function of the quality output of the individual producers, and shows that low reputation can be long lasting and bad reputation-steady states can be difficult to move away from, even when new producers attempt to increase quality.

Winfree and McCluskey (2005) investigated the public good aspect of collective reputation, with an application to agricultural products. Using a dynamic optimization framework, they show that with positive collective reputation and no traceability, there is an incentive to extract rents by producing at lower quality levels. Furthermore, they show that the sustainable level of collective reputation decreases as the number of firms in the production district grows larger, and propose the implementation of minimum quality standards to sustain collective reputation. Carriquiry and Babcock

(2007) further elaborate on the use of quality assurance systems, and their effects on the equilibrium quality level under different market structure scenarios. They conclude that monopolists are more likely to invest heavily on quality, as they can capitalize the full return from investing in reputation.

A related literature focuses on the advertisement efforts of firms that can invest in brand or generic advertisement, or both. Assuming that (costly) advertisement can change consumer believes on quality, these studies analyze the trade-offs between the two investments. Examples include Crespi and Marette (2002), Bass *et al.*, (2005) and Marette and Crespi (2003). Yue *et al.*, (2006) consider the use of brand advertisement and geographical indication in the wine market, extending their analysis to the case of firms that produce at different level of quality. Using a two-stage model with two firms, and assuming that producers can decide whether to direct their marketing efforts towards the development of collective reputations, or invest in brand advertising, they show that geographical indications are preferable for producers that decide not to invest in quality improvements, while quality improving producers will prefer the brand advertisement instrument.

To the knowledge of the author, no study on the simultaneous development of collective *and* firm reputation has been so far accomplished, as the existing literature considers either collective *or* firm reputation. This article is a first attempt to develop a more general theory, having the collective reputation only and firm reputation only as special cases. The model contributes to the literature in which reputation is the result of quality differentials, which are observed imperfectly by consumers.

Using a dynamic optimization framework similar to Shapiro (1982) and Winfree and McCluskey (2005), I analyze the choice of a firm maximizing the present value of its profits in a market in which the return of investing in quality is two-fold: collective and firm reputation. Consumer's uncertainty regarding quality is modeled as a time lag between the changes in product quality and the resulting adjustment in collective and firm reputation. While the lag time is unique for the collective reputation process, I assume that individual firms have different "visibility", and therefore the speed of the firm reputation updating process is specific to each producer.

First, I develop a "myopic" model in which firms do not behave strategically, and do not consider the effect of their own quality choice on the quality choice of other firms in the production district. The model is then expanded using a Cournot-style framework in which firms consider such effects when choosing their profit-maximizing quality level. Profit-maximizing quality choices are derived first for the case of two firms, and then generalized to the case of *N* players under both scenarios. Adoption of specific (potentially estimable) cost and price functions enables to solve for the long-run equilibrium quality level in the case of two firms. Finally, the markets with either collective reputation or firm reputation only are presented as special cases of these general models. Conclusions are made regarding the equilibrium quality level under the different scenarios, and findings are used to develop a theory of quality and reputation.

The Model

Firms produce one unit of output each production cycle, and adjust their quality level over time to maximize their stream of profits. The quality level set by firm *i*, q_i, determines the cost of production according to the function $c(q_i)$. Market price is related to the collective reputation of the firms in the district, *R*, and the individual firm reputation r_i, via $P(R, r_i)$. Both collective and firm reputations are in quality units. Assuming the standard structural cost form, $c'(q) > 0$, $c''(q) > 0$. Also, I assume that $P_R'(R,r) > 0$, $P_r'(R,r) > 0$ and $P_{RR}''(R,r) > 0$, $P_{rr}''(R,r) > 0$. The assumption of increasing returns on reputation derives from a mere observation of the wine market, and how expensive can be the wines of the most famous producers. The condition that $c''(q) > P_{RR}''(R,r) + P_{rr}''(R,r)$ ensures that the quality choice is bounded. Whether individual or collective reputation is more effective at influencing prices is inherently an empirical question, and in this model, I impose $\left.\left(\frac{\partial P}{\partial R} = \frac{\partial P}{\partial r}\right)\right|_{R=r}$ and $\left.\left(\frac{\partial^2 P}{\partial R^2} = \frac{\partial^2 P}{\partial r^2}\right)\right|_{R=r}$. While this might seem to be a strong assumption, it only implies that at parity of reputation, the market values equally collective and firm reputation at the margin. For simplicity, I also assume that $P_{Rr}''(R,r) = 0$. As in Winfree and McCluskey (2005), reputation evolves as a Markovian process of past reputation and present quality. If there are *N* firms in the district, each firm solves the following maximization problem:

(1) $$\max_{q_i \geq 0} \int_0^\infty e^{-\delta t} \left[p_i(R, r_i) - c(q_i) \right] dt$$

subject to:

(2) $$\dot{R} = \gamma \left(\sum_{j=1}^{N} \left(\frac{q_j}{N} \right) - R \right), \text{ with } R(0) \geq 0$$

and:

(3) $$\dot{r} = \gamma \beta_i (q_i - r_i), \text{ with } r_i(0) \geq 0.$$

Where t indexes time, and δ is the discount rate. The parameter $\gamma \in (0,1)$ simulates the lag between the realization of a quality level and the learning process of consumers (or "speed of consumer learning" as in Shapiro, 1982), as consumers might not buy the product continuously. The parameter $\beta_i \in (0,1)$ is a producer-specific parameter that captures the visibility of a firm. Therefore, all firms are identical, but for the value of their visibility parameter. The rationale for such difference is that, due to factors such as size, market share or distribution system, certain firms might have a faster updating process in their reputation than others. The genesis of the visibility parameter is exogenous to the model, so that there is no contradiction with the normalization of the per-firm quantity produced. Also, it should be noticed that collective reputation will also have an associated visibility parameter, which I normalized to one. The present-valued Hamiltonian for firm *i* can be represented as:

(4) $$H_i = [p_i(R, r_i) - c(q_i)] + \lambda_i \gamma \left(\frac{q_i}{N} + \left(\frac{N-1}{N} \right) \varphi(R) - R \right) + \mu_i k_i (q_i - r_i)$$

Where $k_i = \gamma \beta_i$ and $\varphi(R)$ is firm's *i* representation of the strategy adopted by players $j \neq i$.

Scenario I: the Myopic Model

In this section, I examine the case in which firms take the quality choice of other firms in the district as given when making their own decision on quality. Let us first consider the maximization problem limiting the total number of firms to two. I will later generalize the results to the case of *N* firms. The current-value Hamiltonian for firm 1 under this duopoly scenario is:

(5) $$H_1 = [p_1(R, r_1) - c(q_1)] + \lambda_1 \gamma \left(\frac{q_1 + q_2}{2} - R \right) + \mu_1 k_1 (q_1 - r_1).$$

Where λ and μ represent the shadow prices of collective and firm reputation. The first-order conditions for of the current-valued Hamiltonian of this game are:

(6) $$\frac{\partial H_i}{\partial q_i} = 0$$

(7) $$\dot{\lambda}_1 - \delta\lambda_1 = -\frac{\partial H_1}{\partial R}$$

(8) $$\dot{\mu}_1 - \delta\mu_1 = -\frac{\partial H_1}{\partial r_1}.$$

Which respectively imply:

(9) $$c'(q_1) = \frac{1}{2}\lambda_1\gamma + k_1\mu_1$$

(10) $$\dot{\lambda}_1 = \lambda_1(\delta + \gamma) - P_R$$

(11) $$\dot{\mu}_1 = \mu_1(\delta + k_1) - P_{r1}.$$

Solving for the isoclines, I derive:

(12) $$\lambda_1 = \frac{P_R}{\delta + \gamma}$$

(13) $$\mu_1 = \frac{\mathrm{P}_{r1}}{\delta + k_1},$$

which signify that a shorter lag time in the updating process will increase the shadow value of collective and firm reputation. Substituting (12) and (13) into equation (9):

(14) $$c'(q_1) = \frac{1}{2}\gamma_\delta P_R + k_{\delta 1}\,\mathrm{P}_{r1}$$

where $\gamma_\delta = \frac{\gamma}{\delta + \gamma}$; $\kappa_{\delta 1} = \frac{k_1}{\delta + k_1}$ and $0<\kappa_{\delta 1}< \gamma_\delta <1$. Equation (14) equates the marginal cost of investing in quality to the sum of the marginal returns from collective and firm reputation and could be called the "economic equilibrium" for firm 1. The parameters γ_δ and $\kappa_{\delta 1}$ embed the discounting effect due to the fact that an investment in quality now realizes its effects on collective and firm reputation in the future, as consumers become aware of the change in quality.

While equation (14) defines an economic equilibrium, it does not directly identify the final quality equilibrium of the dynamic system for any value of *q*, *R* and *r*. To solve for it, I specify the cost and prices equations as quadratic functional forms. Therefore, $c(q_i) = c_0 + c_1 q_i + c_2 q_i^2$ and $p(q_i) = a_0 + a_1 R + a_2 R^2 + a_1 r_i + a_2 r_i^2$. The previous general structural assumptions regarding the first and second order derivatives of the cost and price functions are retained, so that we can sign $a_1 + 2a_2 R > 0$; $a_1 + 2a_2 r_i > 0$ and $c_2 > 2a_2 > 0$. Substituting the functional forms into equation (14) yields:

$$c_1 + 2c_2 q_1 = \frac{1}{2}\gamma_\delta \left(a_1 + 2a_2 R\right) + \kappa_{\delta 1}\left(a_1 + 2a_2 r_1\right), \tag{15}$$

an explicit relationship between, q, R and r for firm 1can therefore be obtained:

$$q_1(R, r_1) = \frac{1}{2c_2}\left[-c_1 + \left(\frac{1}{2}\gamma_\delta + \kappa_{\delta 1}\right)a_1\right] + \frac{1}{2}\frac{a_2}{c_2}\gamma_\delta R + \frac{a_2}{c_2}\kappa_{\delta 1} r_1, \tag{16}$$

which is a an explicit representation of the quality choice of firm one under the economic equilibrium rule of equation (14).

Quality Equilibrium in the Duopoly Case

A sufficient condition for an equilibrium to exist is that quality choices do not change through time. This condition sets $q_1 \equiv r_1$ in (16). Solving for $q_1(R)$:

$$q_1(R) = \frac{-c_1 + \left(\frac{1}{2}\gamma_\delta + k_{\delta 1}\right)a_1}{2\left(c_2 - k_{\delta 1}a_2\right)} + \frac{1}{2}\frac{\gamma_\delta a_2}{\left(c_2 - k_{\delta 1}a_2\right)} R, \tag{17}$$

a relationship that linearly links the firm quality decision to the collective reputation of the district. By symmetry:

$$q_2(R) = \frac{-c_1 + \left(\frac{1}{2}\gamma_\delta + k_{\delta 2}\right)a_1}{2\left(c_2 - k_{\delta 2}a_2\right)} + \frac{1}{2}\frac{\gamma_\delta a_2}{\left(c_2 - k_{\delta 2}a_2\right)}R \,. \tag{18}$$

I confine the equilibrium analysis to the more interesting case in which firms produce at some positive level of quality when the collective reputation is zero, i.e. $\left(\frac{1}{N}\gamma_\delta + k_{\delta i}\right)a_1 > c_1$. It should be emphasized that $q_i(R)$ is the same for both firms, with the only exception of the firm-specific visibility parameter k_i. The slope of equations (17 and 18) is $\frac{dq_i}{dR} = \frac{1}{N}\frac{\gamma_\delta a_2}{\left(c_2 - k_{i\delta}a_2\right)} > 0$, which is unambiguously less than one under the last assumption and increasing in k_i. Furthermore, the intercept is also increasing in k_i. Figure 1 represent equations (17 and 18), for the case of two firms with visibility parameters $k_1 > k_2$. As it appears clear in the graph, points to the left of point A cannot be an equilibrium, since both firms are producing above the existing level of collective reputation (identified by the 45 degree line), and therefore the reputation of the district must be increasing. A similar argument goes for points to the right of B, as both firms are free riding and diminishing the collective reputation. Clearly, an equilibrium will be reached were $q_1(R)$ and $q_2(R)$ are equidistant the $q = R$ line, that is, at point C.

The distance between the $q_i(R)$ lines and $q = R$ can be found as $q_i(R) - R$ and point C is therefore defined by the relationship $q_1(R) - R = -\left[q_2(R) - R\right]$. Solving for R we get $\overline{R} = -\dfrac{\mathrm{B}_0 + \Gamma_0}{\mathrm{B}_1 + \Gamma_1 - 2}$, where $\mathrm{B}_0, \mathrm{B}_1, \Gamma_0, \Gamma_1$ are the intercepts and slopes of the $q_1(R)$ and $q_2(R)$ lines respectively. Substituting in the parameter values from equations (17) and (18) and simplifying terms yields the equilibrium average quality in the production district as a function of the parameters of the model:

$$(19)\ \overline{Q}_{m(R,r)} = \frac{\left\{\left[-c_1 + \left(\frac{1}{2}\gamma_\delta + k_{\delta 1}\right)a_1\right]\left(c_2 - k_{\delta 2}a_2\right)\right\} + \left\{\left[-c_1 + \left(\frac{1}{2}\gamma_\delta + k_{\delta 2}\right)a_1\right]\left(c_2 - k_{\delta 1}a_2\right)\right\}}{4\left(c_2 - k_{\delta 1}a_2\right)\left(c_2 - k_{\delta 2}a_2\right) - \left[\left(c_2 - k_{\delta 1}a_2\right) + \left(c_2 - k_{\delta 2}a_2\right)\right]\gamma_\delta a_2},$$

where I use the fact that at equilibrium, average quality $\overline{Q} \equiv \overline{R}$ and the subscript $m(R,r)$ indicate the myopic model with both collective and firm reputation. Clearly, $\overline{Q}_{m(R,r)}$ will be a positive quantity under the provision that the first term in the denominator is greater than the second. Therefore, the first finding is that, when firms have different visibility, it is possible to find an equilibrium in which one firm (the more visible) produces above average quality, and the other is to some extent free riding. Also, observing equation (19) shows that the discounting effects due to speed of consumer learning and firm visibility are long lasting and persist even at equilibrium. This model also provides insight on the dynamics of quality and reputation, showing that when collective reputation is below a certain critical level (point A in figure 1), firms find it profitable to produce at higher quality levels, increasing the reputation of the district. Conversely, when collective reputation is above a certain level (point B in figure 1), it is economically convenient for both firms to erode it.

Generalization to *N* Firms and Comparative Statics

I now evaluate how changes in the parameters of the model affect the firm's quality choice. Before examining the comparative statics, let us first generalize equation (0.16) to the case of *N* firms:

$$q_i(R,r_i)=\frac{1}{2c_2}\left[-c_1+\left(\frac{1}{N}\gamma_\delta+\kappa_{\delta i}\right)a_1\right]+\frac{1}{N}\frac{a_2}{c_2}\gamma_\delta R+\frac{a_2}{c_2}k_{\delta i}r_i\,; \tag{20}$$

From (20) it can be seen that $\frac{\partial q_i}{\partial N}=-\frac{\left(a_1+2a_2R\right)\gamma_\delta}{2c_2N^2}<0$, which means that for any given level of collective reputation, all firms will lower quality as a response to an increase in the number of firms in the production district. Conversely, all firms increase quality in response to an increase in their visibility, according to $\frac{\partial q_i}{\partial k_i}=\frac{\left(a_1+2a_2r_i\right)}{2c_2}\left(\frac{\partial\gamma_\delta}{\partial k_i}\right)>0$, where $\frac{\partial\gamma_\delta}{\partial k_i}=\frac{\delta}{\left(\delta+k_i\right)^2}$. Both of these effects are increased in magnitude by the fact that $\frac{dq_i}{dR}=\frac{1}{N}\frac{\gamma_\delta a_2}{c_2}>0$, and therefore as collective reputation decreases or increases in response to changes in N or $k_{i,}$ firms further adjust their quality level in response to the ongoing change in collective reputation.

Scenario II: the Cournot Model

In this scenario, I consider the decision process of a firm that takes into account the effects of its quality level on the choice of other producers in the district when maximizing its own stream of profits. I model this using a Cournot-style model in which firm *i* considers the quality adjustment due to changes in R of firm $j \neq i$, and incorporates it into the maximization problem. In contrast with the classical Cournot models, this modeling framework does not imply that firms are making a once-and-for-all decision on quality, as any firm will still be able to adjust their quality choice in response to exogenous shocks to the level of collective reputation. For the case of a duopoly, firm 2 Hamiltonian is:

(21) $$H_2 = [p_2(R, r_2) - c(q_2)] + \lambda_2 \gamma \left(\frac{q_1(R, r_1) + q_2}{2} - R \right) + \mu_2 \kappa_2 (q_2 - r_2),$$

where $q_1(R, r_1)$ represents equation (16) from the myopic model. Solving the first order conditions yields the set of equations analogous to (9 to 11):

(22) $$c'(q_2) = \frac{1}{2}\lambda_2 \gamma + k_2 \mu_2$$

(23) $$\dot{\lambda}_2 = \lambda_2 \left\{ \delta - \gamma \left[\frac{1}{2} q_{1R} \,'(R, r_1) - 1 \right] \right\} - P_R$$

(24) $$\dot{\mu}_2 = \mu_2 (\delta + k_2) - P_{r2} \,.$$

Solving for the isoclines we obtain:

(25) $$\lambda_2 = \frac{P_R}{\left\{ \delta + \gamma \left[1 - \frac{1}{2} q_{1R} \,'(R, r_1) \right] \right\}}$$

(26) $$\mu_2 = \frac{\mathrm{P}_{r2}}{\delta + k_2} \,,$$

which substituted in equation (22) yields the economic equilibrium rule for firm two:

(27) $$c\,'(q_2) = \frac{1}{2} \frac{\gamma}{\left\{ \delta + \gamma \left[1 - \frac{1}{2} q_{1R} \,'(R, r_1) \right] \right\}} P_R + k_{\delta 2} \, \mathrm{P}_{r2}$$

Quality Equilibrium in the Duopoly Case

Using the same functional forms as in scenario I, and recalling that $q_{1R}'(R,r_1)=\frac{1}{2}\frac{\gamma_\delta a_2}{c_2}$, we obtain:

(28) $$c_1+2c_2q_2=\frac{1}{2}\Delta_2[a_1+2a_2R]+\kappa_{\delta 2}[a_1+2a_2r_2],$$

where $\Delta_2=\frac{\gamma}{\left\{\delta+\gamma\left[1-\frac{1}{4}\gamma_\delta\frac{a_2}{c_2}\right]\right\}}$. Solving for q_2 yields the analogous to (0.16):

(29) $$q_2(R,r_2)=\frac{1}{2c_2}\left[-c_1+\left(\frac{1}{2}\Delta_2+k_{\delta 2}\right)a_1\right]+\frac{1}{2}\frac{a_2}{c_2}\Delta_2R+\frac{a_2}{c_2}k_{\delta 2}r_2.$$

The average quality in the production district, $\overline{Q}_{c(R,r)}$, where c indicates the Cournot model, can be obtained following the same steps used in the myopic model. As equation (29) differs from (16) only for the fact that Δ_2 in (29) replaces γ_δ in (16), the equilibrium quality can be easily found substituting the terms in (19), which yields:

(30)

$$\overline{Q}_{c(R,r)}=\frac{\left\{\left[-c_1+\left(\frac{1}{2}\Delta_2+k_{\delta 1}\right)a_1\right]\left(c_2-k_{\delta 2}a_2\right)\right\}+\left\{\left[-c_1+\left(\frac{1}{2}\Delta_2+k_{\delta 2}\right)a_1\right]\left(c_2-k_{\delta 1}a_2\right)\right\}}{4\left(c_2-k_{\delta 1}a_2\right)\left(c_2-k_{\delta 2}a_2\right)-\left[\left(c_2-k_{\delta 1}a_2\right)+\left(c_2-k_{\delta 2}a_2\right)\right]\Delta_2a_2}.$$

Examining (30) it can be seen that the average quality in the district at equilibrium is an increasing function in Δ_2. Since it is easy to show that $\Delta_2 > \gamma_\delta$, it follows that $\overline{Q}_{c(R,r)} > \overline{Q}_{m(R,r)}$: the equilibrium average quality under the Cournot model is larger than under myopic one. This last finding deserves to be commented further. The rationale for the quality increase from the myopic to the Cournot model lies in the fact that firms under this scenario realize that their quality choice will affect the collective reputation of the district, and that the other firms will respond to an increase in R according to $\frac{\partial q_i}{\partial R} > 0$. That is, when producers benefit from a collective reputation, there is a positive externality in investing in quality.

Generalization to *N* Firms and Comparative Statics

Generalizing (29) to the case of *N* firms we obtain:

(31) $$q_i = \frac{1}{2c_2}\left[-c_1 + \left(\frac{1}{N}\Delta_N + k_{\delta 2}\right)a_1\right] + \frac{1}{N}\frac{a_2}{c_2}\Delta_N R + \frac{a_2}{c_2}k_{\delta 2}r_i\,,$$

where $\Delta_N = \dfrac{\gamma}{\left\{\delta + \gamma\left[1 - \frac{1}{N^2}\gamma_\delta \frac{a_2}{c_2}\right]\right\}}$.

Taking the derivative (0.31) with respect to N yields:

$$\frac{\partial q_i}{\partial N} = -\frac{(a_1 + 2a_2 R)\left(\Delta_N - N\frac{\partial \Delta_N}{\partial N}\right)}{2c_2 N^2} < 0 \text{, since } \frac{\partial \Delta_N}{\partial N} = -\frac{2a_2\gamma^2\gamma_\delta}{c_2\left[\delta + \gamma\left(1 - \frac{a_2\gamma_\delta}{c_2 N}\right)\right]^2 N^3} < 0 .$$

Comparing this result with the one obtained in the myopic scenario, it can be easily realized that the decrease in quality response due to the increase of firms is much larger in the Cournot model, as $\Delta_N > \gamma_\delta$ and $-N\frac{\partial \Delta_N}{\partial N}$ is a negative quantity. Taking the derivative of equation (31) with respect to k_i, returns $\frac{\partial q_i}{\partial k_i} = \frac{(a_1 + 2a_2 r_i)}{2c_2}\left(\frac{\partial \gamma_\delta}{\partial k_i}\right) > 0$, which is the same as in the myopic model. Nevertheless, the feedback effect due to $\frac{\partial q_i}{\partial R}$, which will increase both $\frac{\partial q_i}{\partial N}$ and $\frac{\partial q_i}{\partial k_i}$, is larger under the assumptions of this scenario as $\frac{\partial q_i}{\partial R} = \frac{1}{N}\frac{a_2}{c_2}\Delta_N > 0$ is strictly greater than $\frac{dq_i}{dR} = \frac{1}{N}\frac{a_2}{c_2}\gamma_\delta$ from the myopic scenario. A very important thing to notice is that $\lim_{N\to\infty}\Delta_N = \gamma_\delta$. That is, as the number of firms grows larger, the Cournot model converges in results to the myopic model and equations (31) and (20) become equivalent. It is straightforward now to realize the reason behind it: as the number of firm increases, the positive externality on collective reputation of investing in quality approaches zero. Therefore, firms belonging to a production district with a large number of firms behave myopically and take the quality choice of the other firms as given.

Special Cases: Collective Reputation or Firm Reputation Only

It is useful to compare the results obtained so far to the ones that can be derived from models considering exclusively returns on collective reputation (as in Winfree and McCluskey, 2005) or firm reputation only (as Shapiro,1982). Such results can be easily derived under the same general assumptions adopted in this article as special cases of the economic equilibrium rules represented by equations (14) and (27).

For the case of markets in which firm reputation only exists, the myopic and Cournot equilibria will coincide, as both equations simplify to: $c'(q_i) = k_{\delta i}\,\mathrm{P}_{ri}$.

Substituting in the cost and price functional forms yields: $q_i(R, r_i) = \dfrac{-c + \left(a_1 + 2a_2 r_i\right)\kappa_{\delta i}}{2c_2}$

and using the equilibrium condition $q_i = r_i$ we have: $q_i = \dfrac{-c_1 + a_1}{2\left(c_2 - a_2 k_{\delta i}\right)}$. The duopoly equilibrium average quality is easily derived as

(32) $$\bar{Q}_r = \frac{q_1 + q_2}{2} = \frac{\left(-c_1 + a_1\right)\left[\left(c_2 - a_2 k_{\delta 1}\right) + \left(c_2 - a_2 k_{\delta 2}\right)\right]}{4\left(c_2 - a_2 k_{\delta 1}\right)\left(c_2 - a_2 k_{\delta 2}\right)}.$$

For the case of collective reputation only, equation (14) reduces to $c'(q_1) = \frac{1}{2}\gamma_\delta P_R$. Therefore, in the absence of firm-specific reputation all firms have the same quality response function $q_i(R) = \dfrac{1}{2c_2}\left[-c_1 + \dfrac{1}{2}\gamma_\delta\left(a_1 + 2a_2\right)R\right]$ and the equilibrium quality can be found as the intersection of the $q_i(R)$ line with the 45 degree line in figure 1. Solving for $\overline{R}$ from $q_i(R) = R$ we have: $\overline{Q}_{m(R)} = \dfrac{-c_1 + \frac{1}{2}\gamma_\delta a_1}{2(c_2 - a_2\gamma_\delta)}$. Similarly, we can find the equilibrium level for the Cournot model derived in scenario II as $\overline{Q}_{c(R)} = \dfrac{-c_1 + \frac{1}{2}\Delta_2 a_1}{2(c_2 - a_2\Delta_2)}$. Generalization of these models to the case of *N* firm and the comparative statics are analogous to the ones obtained so far (see table 1), with the obvious proviso that $\frac{\partial q_i}{\partial N} = 0$ and $\frac{\partial q_i}{\partial R} = 0$ for the firm-reputation-only model and $\frac{\partial q_i}{\partial ki}$ is irrelevant for the collective-reputation-only model.

Comparing these results to the ones obtained so far, I find that the following inequalities hold: $\overline{Q}_{c(R,r)} > \overline{Q}_{m(R,r)} > \overline{Q}_r > \overline{Q}_{c(R)} > \overline{Q}_{m(R)}$. That is, the highest sustainable level of quality for a duopoly is achieved in the Cournot scenario with collective and firm reputation, followed by its myopic analogous. The equilibrium quality level progressively decreases as we consider markets with own reputation only and the Cournot model with collective reputation only. The lowest quality level is achieved in markets with collective reputation only and firms behaving myopically. An additional results is that, when N is large, $\overline{Q}_{m(R,r)} = \overline{Q}_{c(R,r)} > \overline{Q}_r > \overline{Q}_{m(R)} = \overline{Q}_{c(R)}$. That is, the quality levels of the Cournot models collapse to their myopic counterparts.

Discussion and Conclusions

The model developed in this article provides a broad framework under which the relationship between quality, collective reputation and firm reputation in markets for experience goods can be analyzed. The case of markets with collective reputation only or firm reputation only can be easily derived from this general model. A summary of the resulting findings is presented in table 1, which I will discuss under the assumption that higher levels of quality are more desirable than lower.

Under the assumptions of this model, three general rules regarding the dynamics of quality can be derived: 1) quality is increasing in the visibility of the single firm, and 2) quality is decreasing in the number of firms in the production district and, 3) when the number of firm is large enough, firms behave myopically, taking the quality of other firms in the district as given. Regarding the static equilibria it can summarized: 1) the speed of consumer learning and the visibility of the individual firms have long lasting effects on the quality decision of each firm, which persist at equilibrium 2) given a set of parameter values, the equilibrium quality will be highest in market with both collective and firm reputation, intermediate for the case of own reputation only, and markets with collective reputation only will yield the lowest equilibrium quality levels. Furthermore, the model provides insight regarding the conditions under which collective reputation is increased or eroded and shows that it is possible to achieve equilibria in which certain firms produce above the average quality of the district, and other firms are free riding by producing below average quality.

Several real-world phenomena are interpretable at the light of these conclusions. For example, the common good problem pointed out by Winfree and McCluskey (2005) for the case of the collective reputation of Washington apples is consistent with these findings. Furthermore, Winfree and McCluskey (2005) and Carriquiry and Babcock (2007) argue that having traceable firms and the developing minimum quality standards could be a solution to the common good problem of collective reputation. According to these results, it can be argued that, if firms are traceable and consumers can recognize them, producers will automatically increase their quality and minimum quality standards might be unnecessary. On the other hand, when information regarding the identity of the individual producer is impossible or difficult to deliver to consumers, quality assurance systems might be a necessity.

The theory outlined in this article also sheds some light on the recent changes affecting the wine industry. Yue *et al.*, (2006) present evidence that the European wine producers are losing market share to the wineries in California, Chile and Australia. According to their article, wineries from the "old world" relied extensively on the use of geographical indications to market their wines, while the new entrants seemed to have focused their marketing efforts in brand advertisement. The authors suggest that the problems affecting collective reputation might be one of the reasons for the decline of European wineries, and argue that the small average firm size in the old world might prevent the implementation of costly quality-improving practices. Steiner's (2004) findings also point out a decline in the valuation of French wines.

According to this theory on quality, it can be argued that the cost of improving quality is likely not the only reason for the decline of European wines: small firms, inherently less recognizable by the consumers, will have a smaller incentive to invest in their own reputation, which will result in a lower quality output. Conversely, examples of successful wine regions such as Champagne in France or Napa Valley in California convincingly fit the description of the recipe for high quality products: a production district with few, highly recognizable producer.

References

Bass F.M., A. Krishnamoorthy and A.P. Suresh Sethi. (2005). "Generic and Brand Advertisement in a Dynamic Duopoly," *Marketing Science*, 24(4):556-568.

Carriquiry, M. and B. Babcock. (2007). "Reputations, Market Structure, and the Choice of Quality Assurance Systems in the Food Industry," *American Journal of Agricultural Economics*, 89(1):12-23.

Crespi M. J. and S. Marette. (2001). "Generic Advertisement and Product Differentiation," *American Journal of Agricultural Economics*, 84(3):691-701.

Marette, S. and J.M. Crespi. (2003). "Can Quality Certification Lead to Stable Cartel?" *Review of Industrial Organization*, 23(1):43-64.

Shapiro, C. (1982). "Consumer Information, Product Quality, and Seller Reputation" *The Bell Journal of Economics*, 13(1):20-35.

Steiner, B. (2004). "French Wines on the Decline? Econometric Evidence from Britain," *Journal of Agricultural Economics,* 55(2):267-88.

Tirole J. (1996). “A Theory of Collective Reputations (with Applications to the Persistence of Corruption and to Firm Quality),” *Review of Economic Studies*, 63(1):1-22.

Winfree, J. A. and J. McCluskey. (2005). “Collective Reputation and Quality” *American Journal of Agricultural Economics*, 87:206-213.

Yue, C., S. Marette and J. Beghin, “How to Promote Quality Perception in Wine Markets: Brand advertising or Geographical Indication?” *Working paper* 06-WP 426. August 2006.

Table 2.1. Summary of Findings

Model	$\frac{\partial q_i}{\partial N} < 0$	$\frac{\partial q_i}{\partial k_i} > 0$	[1]Two-firm $\bar{Q}$
Cournot(R,r)	$-\frac{(a_1+2a_2R)\left(\Delta_N - N\frac{\partial \Delta_N}{\partial N}\right)}{2c_2N^2}-\left[\frac{1}{N}\frac{a_2}{c_2}\Delta_N\right]$	$\frac{(a_1+2a_2r_i)\left[\frac{\partial k_\delta}{\partial k_i}\right]}{2c_2}+\left[\frac{1}{N}\frac{a_2}{c_2}\Delta_N\right]$	$\frac{\left\{\left[-c_1+\left(\frac{1}{2}\Delta_2+k_{\delta 1}\right)a_1\right](c_2-k_{\delta 2}a_2)\right\}+\left\{\left[-c_1+\left(\frac{1}{2}\Delta_2+k_{\delta 2}\right)a_1\right](c_2-k_{\delta 1}a_2)\right\}}{4(c_2-k_{\delta 1}a_2)(c_2-k_{\delta 2}a_2)-\left[(c_2-k_{\delta 1}a_2)+(c_2-k_{\delta 2}a_2)\right]\Delta_2 a_2}$
Miopic(R,r)	$-\frac{(a_1+2a_2R)\gamma_\delta}{2c_2N^2}-\left[\frac{1}{N}\frac{a_2}{c_2}\gamma_\delta\right]$	$\frac{(a_1+2a_2r_i)\left[\frac{\partial k_\delta}{\partial k_i}\right]}{2c_2}+\left[\frac{1}{N}\frac{a_2}{c_2}\gamma_\delta\right]$	$\frac{\left\{\left[-c_1+\left(\frac{1}{2}\gamma_\delta+k_{\delta 1}\right)a_1\right](c_2-k_{\delta 2}a_2)\right\}+\left\{\left[-c_1+\left(\frac{1}{2}\gamma_\delta+k_{\delta 2}\right)a_1\right](c_2-k_{\delta 1}a_2)\right\}}{4(c_2-k_{\delta 1}a_2)(c_2-k_{\delta 2}a_2)-\left[(c_2-k_{\delta 1}a_2)+(c_2-k_{\delta 2}a_2)\right]\gamma_\delta a_2}$
Miopic(r) Cournot(r)	-	$\frac{(a_1+2a_2r_i)\left[\frac{\partial k_\delta}{\partial k_i}\right]}{2c_2}$	$\frac{(-c_1+a_1)\left[(c_2-a_2k_{\delta 1})+(c_2-a_2k_{\delta 2})\right]}{4(c_2-a_2k_{\delta 1})(c_2-a_2k_{\delta 2})}$
Cournot(R)	$-\frac{(a_1+2a_2R)\left(\Delta_N - N\frac{\partial \Delta_N}{\partial N}\right)}{2c_2N^2}-\left[\frac{1}{N}\frac{a_2}{c_2}\Delta_N\right]$	-	$\frac{-c_1+\frac{1}{2}\Delta_2 a_1}{2(c_2-a_2\Delta_2)}$
Miopic(R)	- $-\frac{(a_1+2a_2R)\gamma_\delta}{2c_2N^2}-\left[\frac{1}{N}\frac{a_2}{c_2}\gamma_\delta\right]$	-	$\frac{-c_1+\frac{1}{2}\gamma_\delta a_1}{2(c_2-a_2\gamma_\delta)}$

1: equilibrium quality is decreasing along the column from top to bottom

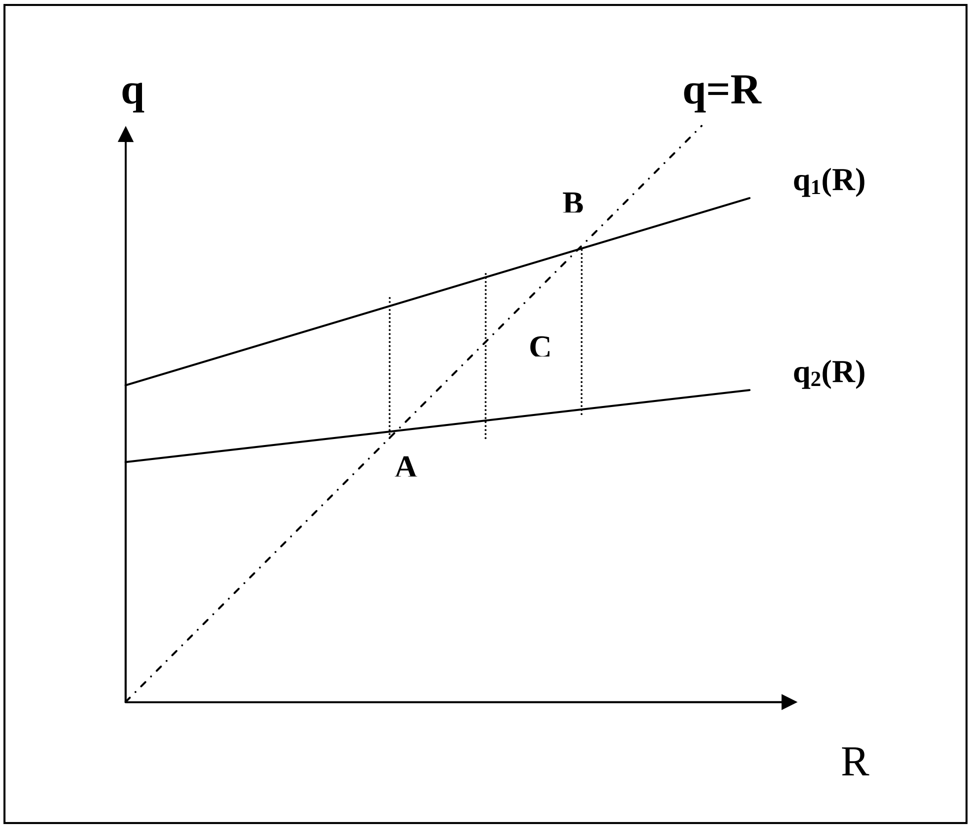

Figure 2.1: Optimal Response of Firm 1 and 2 to Changes in Collective Reputation, with a Graphical Solution for the Equilibrium Level of Collective Reputation. Case of $\beta_1 > \beta_2$.

CHAPTER THREE

SEGMENTING THE WINE MARKET BASED ON PRICE: HEDONIC REGRESSION WHEN DIFFERENT PRICES MEAN DIFFERENT PRODUCTS

Chapter Abstract: Many economists have estimated hedonic price functions for red and white wine. However, estimating a single hedonic price function imposes the assumption that the implicit prices of the attributes are the same for any red or white wine. I argue that even within these two categories, wines are differentiated, and disregarding this heterogeneity causes an aggregation bias in the estimated implicit prices. By estimating hedonic functions specific to price ranges, I show that the wine market is segmented into several product classes or market segments. I find that a model accounting for the existence of wine classes has greater ability to explain the variability in the data and produces more accurate and interpretable results regarding the implicit prices of the attributes.

Key Words: Hedonic regression, wine, market segmentation, structural breaks

Introduction

Numerous authors have estimated hedonic functions for wine. While part of this body of research focuses attention on a single type of wine such as Bordeaux, many articles estimate hedonic functions broadly applied to either red or white wines and in some cases to both. The implicit assumption in these studies is that the hedonic relationship between prices and attributes is unique and little variation in that hedonic relationship exists between or within the red and white wine classification. Using data on red wine, I investigate the possibility that multiple wine classes or market segments exist and that separate hedonic functions should be estimated for each wine class.

The article proceeds as follows: first, a brief review of the existing hedonic literature is presented, and the case for the existence of differentiated wine classes is made. Then, the hedonic model is introduced as the theoretical basis of the analysis, the data set is presented, and the empirical specification discussed. A methodology to identify market segments is then developed and applied, and class-specific hedonic models are estimated. Finally, the results are presented, and their implications discussed in light of a comparison between the class-specific (or segmented) modeling approach versus the traditional pooled approach.

Literature Review

The literature seeking to identify the determinants of wine prices using hedonic techniques is well established. A considerable amount of work has been done to determine which wine attributes significantly affect wine prices: Combris *et al*., (1997, 2000) showed that when regressing objective and sensory characteristics on wine price, the objective cues (such as expert rating score and vintage) are significant, while sensory variables (such as tannins content and other measurable chemicals) are not. Much of the literature (Oczkowski, 1994; Landon and Smith, 1997; Schamel and Anderson, 2003; Angulo *et al*., 2000) indicates that ratings by specialized magazines are significant and should be included in modeling wine prices. Possible explanations for the insignificance of sensory cues are the difficulty of isolating the effect of each chemical on the final flavor and smell and that only a small percentage of wine purchasers are connoisseurs. Therefore, expert ratings act as a signal to the consumer. It is uncertain whether expert ratings influence prices because they are good proxies for quality of the wine or because of their marketing effect. Oczkowski (2001) finds that tasting scores are only proxies for quality, and uses two-stage ordinary least squares and factor analysis to correct for measurement error in the independent variables. On the other hand, Schamel and Anderson (2003) find no evidence of such problem in their sample. The region of production, capturing production costs differentials and the effects of the collective reputation of the district, and the vintage are often reported as significant variables (Angulo *et al.*, 2000; Schamel and Anderson, 2003). Steiner (2004) finds a declining valuation of French wines with geographical appellation in the British market.

Market Segmentation

While various concerns regarding econometric pathologies endemic to the hedonic wine literature appear to be legitimate (see Oczkowski, 2001 regarding measurement error and Unwin, 1999 on heteroskedasticity, collinearity and an extensive critique of the field), I argue that the extreme heterogeneity of wine as a product class is a *prima facie* reason why fundamental model specification issues should be of principal concern to market analysts. Related to this issue, Thrane (2004) pointed out that it is unlikely that the same hedonic function will apply for red and white wines. That is, whether wine attributes affect red and white wines in the same way is a matter that should be tested empirically. In my opinion, the question can be posed on *a priori* grounds. For example, since the aging potential of red and white wines differs considerably, it seems reasonable to expect the implicit price of aging will also be distinct.

Although estimating separate hedonic functions for red and white wines appears advisable at a minimum, the approach might not go far enough in reducing sample heterogeneity and may not produce the most meaningful or accurate comparisons of attribute values. The question of how else the market is segmented remains. Consider an example from the real estate literature to illustrate this issue. Researchers routinely estimate separate hedonic functions for single family houses, duplexes, apartments and commercial buildings. It is recognized that the same attributes (say, squared footage) can affect prices in substantially different ways across property types. On the other hand, these product classes are believed to include properties that are, although heterogeneous, relatively similar. Therefore, consumers see them as variations on the same basic product. However, Straszheim (1974) argued further that market segmentation is present in the housing market even within such categorizations. He showed that by estimating separate hedonic price functions for different geographic areas of the San Francisco Bay area, the sum of squared errors in predicting prices across the entire sample was significantly reduced. Freeman (1993) has analyzed the conditions under which undifferentiated products are traded in segmented markets.

The objective of this paper is to investigate whether the wine market should be segmented into differentiated markets or product classes. A major empirical challenge is the identification of such classes, since there are many possible ways to segment the wine market. Both the wine industry (Ernst and Young, 1999) and typical wine consumers (Hall *et al*., 2001) use price categories to define product-class categories, which provides a rationale for one approach to identification of product classes: segment the wine market by price. The consulting report by Ernst and Young (1999), which aimed at providing tailored marketing strategies for the Australian wine industry, divided wines into commercial, semi-premium, premium and ultra-premium categories on the basis of retail price ranges[1] and analyzed each category. The wine industry tends to self-select into product-class categories. Scott Morton and Podolny (2002) analyzed the motivations of California winery owners and their effects on price and quality. They

[1]The authors specify that the price ranges for commercial, semi-premium, premium and ultra-premium are, respectively: less then A$8 or A$10, between A$10 and A$15 or A$20, between A$20 and A$30 or A$50, and more than A$30 or A$50

found that the owners who derive strong non-financial returns charge more for their product and placed themselves on the high quality end of the spectrum. The owners whose focus was mostly financial were less likely to produce high quality wines.

On the consumer side, Hall *et al*. (2001) found that price is used as a quality cue and that consumers look for different attributes, or value the same attributes differently, depending on the occasion the wine is meant to be consumed. One could argue that most consumers have a price range in mind before purchase, which depends on the occasion of consumption, and also look for different attributes depending on the occasion: a wine that works well enough for a casual dinner with friends might not be perceived as appropriate for a gala anniversary celebration. In other words, not all wines are fungible across occasions of consumption, and when deciding which wine to buy, the typical consumer compares alternatives within the chosen price range. Consequently, the same attributes can have different relevance across price categories and therefore different estimated coefficients.

Following the Ernst and Young (1999) analysis, I assume that four market segments are also applicable to the U.S. red wine market but leave the identification of the classes' boundaries to empirical estimation. Hedonic regressions specific to each wine class are then estimated. Although the literature on wine valuation is extensive, to the knowledge of the author, segmentation of the U.S. red wine market has not been investigated and an analysis of the potentially different effects of product attributes across price segments has not been accomplished.

Theoretical Context

Following the standard hedonic price model (Rosen, 1974), the price of wine, *P*, is assumed to be described by a hedonic price function, $P = P(z)$, where *z* is a vector of attributes. The hedonic price of an additional unit of a particular attribute is determined as the partial derivative of the hedonic price function with respect to that particular attribute. Each consumer chooses an optimal bundle of attributes and all other goods in order to maximize utility subject to a budget constraint. For continuously varying attributes, the chosen bundle will place the consumer so that his or her indifference curve is tangent to the price gradient, $\partial P / \partial z_j$, for each attribute. Therefore, the marginal willingness to pay for a change in a wine attribute is equal to the derivative of the hedonic price function with respect to that attribute. Finite differences $\Delta P / \Delta z_j$ represent marginal willingness to pay for discretely varying attributes. Given that the market is segmented by price categories, the hedonic analysis is then represented by a set of hedonic price functions of the general form $P = P_m(Z) \ \mathit{for} \ P \in (\ell_m, h_m],\ m = 1, ..., s$, where s denotes the number of segments, and ℓ_m and h_m denote the lower and upper price boundaries of market segment *m*, respectively, with corresponding marginal willingness to pay for attributes given by $\partial P_m / \partial z_j$ or $\Delta P_m / \Delta z_j$ for market segment *m*.

Data

The data set is composed of 13,024 observations derived from ten years (1991-2000) of ratings scores reported in the *Wine Spectator* magazine (online version) for California and Washington red wines. Four of the variables are non-binary: price of the wine adjusted to 2000 values by a consumer price index (CPI) for alcohol, score obtained in expert sensory evaluation by the *Wine Spectator*'*s* experts, the number of cases produced, and the years of aging before commercialization. Descriptive statistics for these variables are reported in table 1. Note that wine prices have a positively skewed distribution, but the majority of the observations fall in the $10 to $50 range. California has more wines in the "expensive" category than Washington, with few Washington wines exceeding $100.

Indicator variables were used to denote regions of production, wine varieties, and the presence of label information. The regions of production for California wines include Napa Valley, Bay Area, Sonoma, South Coast, Carneros, Sierra-Foothills and Mendocino, while Washington wines were not separated by regions. These geographical partitions are those adopted by the *Wine Spectator* to categorize the wines, often pooling several American Viticultural Areas (AVAs) in the same region. Varieties include Zinfandel, Pinot Noir, Cabernet Sauvignon, Merlot, and Syrah grapes, as well as wines made from blending of different varieties (non-varietals). The vintage year is available for each wine along with other label information such as "reserve" and "estate produced."

Table 2 reports all variables and abbreviations used throughout the paper together with a brief description.

Specification

Economic theory often suggests the expected sign of the partial derivatives of price with respect to specific attributes but does not restrict functional form. Nevertheless, the choice of the functional form of the hedonic model is fundamental since it determines how the marginal prices will be functionally related to the attributes. Triplett (2004) argued that model specification is ultimately an empirical matter. Given the uncertainty surrounding the correct specification, a flexible functional form is arguably a prudent empirical modeling strategy. A series of possible transformations of the dependent variable were considered and evaluated on the basis of variance stabilization, normality of the residuals and misspecification[2]. As in Landon and Smith (1997), I find that the inverse square root is the best performing transformation. The final specification of the independent variables was also determined by screening possible transformations of the non-binary variables and examining excluded variable residual plots. Furthermore, intercept and slope shifters are used to allow separate regression

[2] The Goldfeldt-Quandt test was used to detect heteroskedasticity proportional to predicted values and the RESET test for misspecification. For the normality of the residuals, we employed three different tests: Anderson-Darling, Komolgorov-Smirnoff and Ryan-Joiner.

functions for Washington and California wines. The functional form ultimately selected is the following:

$$
\begin{aligned}
(1)\qquad Price^{-0.5} &= \beta_0 + \beta_0^w + \left(\beta_1 + \beta_1^w WA\right)(Score) + \left(\beta_2 + \beta_2^w WA\right)(Score)^2 + \left(\beta_3 + \beta_3^w WA\right)(Age) + \\
&\left(\beta_4 + \beta_4^w WA\right)(Age)^2 + \left(\beta_5 + \beta_5^w WA\right) LN(Cases) + \sum_{i=1}^{5}\left(\beta_{5+i} + \beta_{5+i}^w WA\right)(Variety_i) + \\
&\sum_{i=1}^{9}\left(\beta_{10+i} + \beta_{10+i}^w WA\right)(Vintage_i) + \sum_{i=1}^{3}\left(\beta_{19+i} + \beta_{19+i}^w WA\right)(Label_i) + \sum_{i=1}^{7}\beta_{22+i}(Region_i) + \varepsilon_i
\end{aligned}
$$

where WA denotes an indicator variable for Washington State.

The model in (1) was estimated via OLS. Formal testing still detected a moderate degree of heteroskedasticity[3], but the possible gains in estimation efficiency that might be achieved by adjusting the estimator for an appropriate heteroskedastic process are muted by the consistency of the OLS estimator and the large sample size on which the estimators are based[4]. Nevertheless, as a matter of caution the covariance matrix of the parameters was estimated using White's consistent heteroskedasticity-robust estimator.

---

[3] It should be noticed that the power of a test is increasing in sample size. In the limit, if the sample is large enough, a formal test will reject virtually *any* hypothesis stated in the form of strict equality.

[4] There is also the mitigating issue of the need to discover the correct heteroskedastic structure of the error process.

Structural Breaks in Prices

Conceptually the problem of partitioning the data by price is one of locating a set of *n* breakpoints that represent the price ranges that demarcate *n* + 1 market segments. I assume that four differentiated market segment exist, therefore setting the number of structural breaks, *n*, to three. To estimate the optimal location of the structural breaks, the criterion of maximizing goodness of fit to the data was adopted. In particular, the set of breakpoints were chosen that minimized the sum of square errors across the four models (one for each price segment) over all the possible different market partitions. The combinatorial nature of the search problem is clear: the number of alternative possible market segmentations is large, and for each of them four vectors of ordinary least square (OLS) coefficients are needed to calculate the test statistics. In order to reduce the number of calculations, a total of thirty-six possible breakpoints located over the range from $10 to $70 were set. The grid commenced with increments of $1 in the lower range of prices, from $10 to $35, where most of the data lies; then with steps of $2 in the range from $35 to $45, but $40 was also included; and finally with steps of $5 from $45 to $70. An algorithm was written in GAUSS to estimate the statistics for all combinations of three breakpoints yielding calculable parameter estimates (*i.e.*, for nonsingular explanatory variable matrices). Once the optimal price breakpoints were located, price range-specific hedonic regressions were estimated via OLS.

Results and Discussion

The price breakpoints minimizing the sum of squared errors (SSE) identified four price categories corresponding to the four hypothesized market segments: commercial wines (price less than $13), semi-premium (between $13 and $21), premium (between $21 and $40) and ultra-premium (greater than or equal to $40)[5]. The corresponding sample sizes associated with these market segments are 1635, 4,114, 4,809 and 2,475 observations, respectively.

Estimated coefficients of the model (1) for the pooled (estimating a single hedonic function) and the segmented models are reported in tables 3 and 4. Coefficients relative to the Washington slope shifters were mostly insignificant in the segmented model and are not included in the tables. Comparing the pooled to the segmented approach, the value of adjusted R^2 increases from 0.67 to 0.91. As in Straszheim (1974), the greater flexibility of the segmented approach allows us to capture the specifics of each wine class, resulting in substantially greater explanatory power.

The hypothesis that OLS regression coefficients are equal across the price categories was tested via a Wald statistic. The test statistic was framed analogous to a Chow-type test, whereby parameters associated with like variables across each of the price-segmented models were hypothesized to be equal. White's heteroskedasticity

---

[5] A reasonable concern is that the price breakpoints might not be the same for California and Washington. To investigate this hypothesis, we run the GAUSS algorithm on the two separated dataset. Interestingly, we find that the price categories minimizing SSE are the same for WA, CA and the pooled data set.

robust estimator was used in the test to represent the covariance matrix of the parameter estimates, and a value of the Wald statistic was then calculated to test the equality of all coefficients across classes. All test results strongly reject the null hypothesis (see table 5).

In interpreting the results in table 3 and 4, it is important to note that owing to the transformation of the dependent variable, coefficients with a negative sign signify a positive impact of the wine attribute on price, and vice-versa. Empirical results conformed to *a priori* expectations: for all estimated models, price is increasing in aging and rating score over the range of the data and decreasing in the number of cases produced. Confirming previously published results, regional appellations command price premia relative to a generic California wine, with "Napa Valley" bringing the largest premium. The coefficients associated with the variety variables capture the difference in price relative to Zinfandel grapes and the coefficients for vintages refer to price differences relative to the excluded year 2000. Interestingly, all price impacts are negative and show a very clear pattern: the 1991 and 1992 vintages were the largest in magnitude, and then slowly decreased year by year. This suggests that these indicator variables may not only be representing a vintage effect (e.g. good or bad climatic conditions that can affect wine production) but may be confounded by a temporal trend of the prices not accounted for by the CPI scaling.[6]

Examination of additional estimated hedonic function coefficients and corresponding implicit prices serves to further characterize each wine class and provide for a further contrast between the approach I propose in this article and the traditional pooled approach. I emphasize that, owing to the transformation of the dependent variable, implicit prices are functions of both estimated coefficients and prices. Average implicit prices for each of the attributes were calculated using market-segment specific price averages and results and figures refer to the appropriate ranges of the data.

The derivative of price with respect to the number of cases produced is strictly negative for all market segments and approaches zero as the number of cases increases. As for the quantitative difference across market segments, increasing total production decreases the market price of wines only slightly in the commercial market segment. The decrease is more pronounced in the two middle segments and is quite substantial in the ultra-premium wine segment (more than five times the estimate relative to the commercial segment). By estimating separated models, I am able to segregate wines that have a "collectible" or "cult wine" value from the "consumption" type wines. The value of an additional point in the *Wine Spectator* tasting score shows an analogous effect:

[6] Several authors (Pakes, 2003 and Triplett, 2004) suggested the use hedonic models to calculate CPI indexes as an alternative to the currently used matched models. The model specification of this research fits the "time dummy variable" method described by Triplett (2004, p. 48) to calculate CPI indexes. The fact that a time trend is still present despite the fact that prices had already been CPI adjusted suggests that, as many authors observed, the two methods yield considerably different results.

better scores in the tasting review increase the price of the wine significantly. This effect is increasingly important in order of the commercial, to semi-premium and premium, market segments, and becomes highly relevant for the ultra-premium wines.

Differences across market segments regarding the impact of cellaring on wine price are even more pronounced (figure 1). As expected, wine aging for the commercial, semi-premium and premium classes exhibits decreasing marginal returns over time. In contrast, ultra-premium wines show different pricing dynamics: the implicit price of aging increases over the full range of the data. The pooled regression approach does not account for qualitative differences (different signs or slopes across price segments), as only one coefficient (or, for the case of polynomials, one set) is estimated for each attribute. On the other hand, marginal prices are weighted by price, so that quantitative differences are embedded in the regression even in the pooled approach.

Examples of wine class-specific peculiarities in the estimated implicit prices are multiple. Washington wines sell for a discount in the premium and ultra-premium classes, but are no different from California wines in the commercial and semi-premium markets. Blended wines sell for a high premium in the fine wines segment, while they are not different from Zinfandel wines in the inexpensive price segment[7]. Among the varietals, Merlots have the highest associated price premium in the commercial segment,

[7] Blended wines are a heterogeneous category. They range from “table wines” made from several grape varieties mixed in unknown percentages to high-quality, finely balanced wines, such as Meritage. In this instance, the segmented approach allows differentiating between these attributes that share a common denomination, but are inherently different.

while Cabernets and Pinot Noirs are the most expensive ultra premium varietal wines. In general, I find that the segmented model produces a much richer and detailed amount of information relating to the character of wine markets.

Lastly, I emphasize that estimated price premia from the pooled approach are consistently higher than those from the segmented approach. This can be explained in the context of the different interpretation of the estimates: the price premia associated with the pooled data refer to the mean value of the excluded variable for the *entire* price range, while the segmented price premia refer to the mean value of the excluded variable *within* the price category. The difference is not merely semantic. If wines in different classes are actually different products, this effect can result in false significance of the explanatory variables

Conclusions

This article provides empirical evidence that the wine market is differentiated into multiple segments or wine classes. I find that a model considering market segmentation has greater ability to explain the variability of the data and, just as importantly, produces more defensible and informative estimates of the hedonic relationship between prices and wine attributes. The analysis identifies wine classes based on price ranges as well as out-of-sample information relating to the existence of different wine segments. By specifying hedonic functions for different product-class categories, I find evidence that consumers value the same wine attributes differently across categories.

There are many possible ways to segment the wine market. While the current approach produced reasonable results, the matter of how to best identify segments in the wine market remains an open question. Research is ongoing to develop methods that identify wine classes using information in addition to price and that endogenously determine the number of price segments (see Chapter Four).

References

Angulo, A. M., J.M. Gil, A. Gracia and M. Sanchez. (2000). "Hedonic Prices for Spanish Red Quality Wine," *British Food Journal*, 102(7):481-493.

Combris, P., L. Sebastien L and M. Visser. (1997). "Estimation of a Hedonic Price Equation for Bordeaux Wine: Does Quality Matter?" *Economic Journal* 107(441):390-403.

Combris, P. and S. Lecocq and M. Visser. (2000). "Estimation of a Hedonic Price Equation for Burgundy Wine," *Applied Economics*, 32:961-967.

Ernst and Young Entrepreneurs (1999). Etude des Filieres et des Strategies de Développement des Pays Producteurs de Vins dans le Monde: Analyse de la Filiere Viticole Australienne, *ONIVINS (Office National Interprofessionnel des Vins).*

Freeman, A.M. (1993). *The Measurement of Environmental and Resource Values: Theory and Methods,* Resources for the Future, Washington, DC.

Hall, J., L. Lockshin and G.B. O'Mahony. (2001). "Exploring the Links Between the Choice and Dining Occasion: Factors of Influence." *International Journal of Wine Marketing*, 13(1):36.

Landon, S. and C.E. Smith. (1997). "The Use of Quality and Reputation Indicators by the Consumers: The Case of Bordeaux Wine," *Journal of Consumer Policy,* 20:289-323.

Nerlove, M. (1995). Hedonic Price Functions and the Measurement of Preferences: the Case of Swedish Wine Consumers, *European Economic Review* 39:1967-1716.

Oczkowski, E. (1994). "Hedonic Wine Price Function for Australian Premium Table Wine," *Australian Journal of Agricultural Economics,* 38:93-110.

Oczkowski, E. (2001). "Hedonic Wine Price Functions and Measurement Error," *Economic Record,* 77:374–82.

Pakes, A. (2003). A Reconsideration of Hedonic Price Indexes with an Application to PC's, *The American Economic Review* 93(5):1578-1596.

Rosen, S. (1974). "Hedonic Prices and Implicit Markets: Product differentiation in Pure Competition," *Journal of Political Economy,* 82:34-55.

Schamel, G. and K. Anderson. (2003). “Wine Quality and Varietal, Regional and Winery Reputations: Hedonic prices for Australia and New Zealand,” *Economic Record,* 79:246.

Scott Morton, F.M. and J.M. Podolny. (2002). “Love or Money? The Effects of Owner Motivation in the California Wine Industry.” *Journal of Industrial Economics,* 40:431-456.

Spawton, T. (1991). “Marketing Planning for Wine,” *European Journal of Marketing,* 25(3):2-47.

Steiner, B. (2004). “French Wines on the Decline? Econometric Evidence from Britain,” *Journal of Agricultural Economics,* 55(2):267-88.

Straszheim, M. (1974). “Hedonic Estimation of Housing Market Prices: A Further Comment,” *Review of Economics and Statistics,* 56(3):404-406.

Thrane C. (2004). “In Defense of the Price Hedonic Model in Wine Research,” *Journal of Wine Research,* 15(2):123-134.

Triplett, J. (2004). Handbook on Hedonic Indexes and Quality Adjustments in Price Indexes: Special Application to Information Technology Products. OECD Science, Technology and Industry Working Papers, 2004/9, OECD Publishing.

Unwin, T. (1999). Hedonic Price Indexes and the Qualities of Wine. *Journal of Wine Research* 10:95-104.

Table 3.1. Descriptive Statistics of Quantitative Explanatory Variables

State	Variable	Mean	Median	Minimum	Maximum
California[a]	Price[c]	31.1	22	3	2,000
	Cases	6,719	1,467	16	950,000
	Score	86.1	87	60	99
	Age	2.8	3	1	9
Washington[b]	Price[c]	23.3	20	5	144
	Cases	6,720	1,000	45	550,000
	Score	86.8	87	67	96
	Age	2.8	3	1	7

[a] 11,774 observations.
[b] 1,250 observations.
[c] Adjusted to year 2000 by a CPI index for alcohol.

Table 3.2. Descriptions of the Abbreviation Used for the Explanatory Variables

Predictor	Short Description
Score	Rating Score from the Wine Spectator
Scscore	Score Centered by Subtracting its Mean
Scscore2	Scscore Squared
Age	Years of Aging Before Commercialization
Agesc	Age Centered by Subtracting its Mean
Agesc2	Agesc Squared
Cases	Number of Cases Produced
Lncas	Natural Log of Hundreds of Cases Produced
Napa	
Bay area	
Sonoma	
South coast	Region of Production
Carneros	
Sierra foothills	
Mendocino	
Washington	
Nonvarietal	
Pinot noir	
Cabernet	Grape Variety
Merlot	
Syrah	
Reserve	"Reserve" was Reported on the Label
Vineyard	Specific Name of the Vineyard on the Label
Estate	"Estate" Produced Wine
91, …, 99	Vintage
Wa	Washington State wines

Table 3.3. OLS Estimates for Pooled and Segmented Hedonic Functions (First Set)

	Pooled	Segmented			
Adjusted R^2	0.67	0.91*			
		Commercial	Semi-Premium	Premium	Ultra-Premium
		0.29	0.21	0.19	0.33
N	13,024	1,635	4,114	4,809	2,475
Covariate	Coefficient*10^2 (*t*-ratio)				
Constant	21.999	29.233	23.943	18.794	14.543
	(104.7)	(56.66)	(140.3)	(116.6)	(50.2)
Scscore	-0.620	-0.275	-0.158	-0.145	-0.168
	(-61.29)	(-6.36)	(-15.86)	(-18.69)	(-10.79)
Scscore2	-0.022	-0.005	-0.008	-0.007	-0.027
	(-16.47)	(-1.25)	(-6.61)	(-6.01)	(-10.35)
Agesc	-1.302	-0.572	-0.185	-0.185	-0.175
	(-23.69)	(-4.62)	(-4.62)	(-4.5)	(-1.83)
Agesc2	0.108	0.334	0.078	0.038	-0.109
	(2.63)	(2.71)	(2.06)	(1.24)	(-2.01)
Lncas	1.004	0.486	0.292	0.290	0.185
	(41.75)	(7.82)	(15.82)	(16.28)	(5.51)
Napa[a]	-5.483	-1.602	-1.406	-0.478	0.188
	(-36.77)	(-5.78)	(-13.64)	(-3.51)	(0.83)
Bay Area[a]	-3.437	-1.266	-0.746	-0.135	0.364
	(-17.34)	(-3.67)	(-5.47)	(-0.84)	(1.41)
Sonoma[a]	-4.053	-2.251	-1.034	-0.110	0.573
	(-28.31)	(-10.38)	(-11.18)	(-0.82)	(2.55)
South Coast[a]	-3.222	-2.491	-0.809	0.147	1.246
	(-20.46)	(-8.81)	(-7.47)	(1.02)	(4.89)
Carneros[a]	-4.291	-2.899	-1.537	-0.086	0.423
	(-23.74)	(-6.28)	(-11.76)	(-0.56)	(1.69)
Sierra Foothills[a]	-2.327	-1.634	-0.296	-0.037	1.440
	(-10.3)	(-5.26)	(-2.02)	(-0.18)	(4.61)
Mendocino[a]	-2.406	-1.555	-0.538	0.196	1.183
	(-12.64)	(-5.24)	(-4.22)	(1.18)	(3.91)
Wa[a]	-0.652	1.178	-0.904	1.001	1.817
	(-0.95)	(1.44)	(-1.38)	(3.91)	(3.74)

[a] Omitted variable: generic California
[*] Calculated stacking the segmented datasets in a single (block diagonal) design matrix and estimating the segmented hedonic model all at once, with a single constant.

Table 3.4. OLS Estimates for Pooled and Segmented Hedonic Functions

Covariate	Pooled	Segmented: Commercial	Semi-Premium	Premium	Ultra-Premium
	Coefficient*10^2 (*t*-ratio)				
Nonvarietal[a]	-4.319	-0.910	-0.926	-1.207	-2.096
	(-27.29)	(-1.98)	(-5.93)	(-10.82)	(-9.05)
Pinot Noir[a]	-3.252	-0.588	-0.727	-1.122	-0.839
	(-32.16)	(-2.04)	(-9.13)	(-15.68)	(-4.01)
Cabernet[a]	-2.479	-0.684	-0.512	-1.060	-1.333
	(-24.4)	(-3.3)	(-6.7)	(-13.34)	(-6.33)
Merlot[a]	-2.200	-1.069	-0.759	-0.770	-0.653
	(-21.79)	(-4.77)	(-10.06)	(-9.751)	(-2.95)
Syrah[a]	-0.582	0.003	-0.392	-0.186	-0.029
	(-4.26)	(0.01)	(-3.8)	(-1.932)	(-0.11)
Reserve[b]	-1.105	0.690	-0.143	-0.512	0.438
	(-10.89)	(2.36)	(-1.5)	(-7.56)	(4.30)
Vineyard[b]	-0.858	-1.407	-0.151	-0.281	-0.171
	(-11.46)	(-5.99)	(-2.14)	(-5.25)	(-1.77)
Estate[b]	-0.601	-2.483	0.171	-0.261	-0.022
	(-2.88)	(-4.72)	(1.01)	(-1.62)	(-0.1)
91[c]	5.353	1.590	1.195	1.231	-
	(31.78)	(3.57)	(8.2)	(9.27)	-
92[c]	5.339	1.977	1.236	1.190	-0.021
	(31.38)	(4.37)	(8.6)	(9.38)	(-0.06)
93[c]	4.372	1.679	1.096	1.156	0.307
	(27.24)	(3.62)	(7.61)	(9.84)	(1.13)
94[c]	4.097	0.915	0.824	1.176	0.498
	(27.8)	(2.01)	(5.82)	(10.82)	(2.23)
95[c]	3.311	0.333	0.477	0.926	0.657
	(23.24)	(0.75)	(3.41)	(9)	(3.48)
96[c]	2.397	-0.001	0.236	0.699	0.646
	(17.75)	(-0.003)	(1.7)	(7.06)	(3.52)
97[c]	1.749	-0.209	0.225	0.696	0.383
	(13.07)	(-0.45)	(1.58)	(7.3)	(2.38)
98[c]	0.393	0.090	0.171	0.408	-0.385
	(2.77)	(0.181)	(1.12)	(4.03)	(-2.4)
99[c]	0.668	-0.346	0.170	0.294	0.050
	(5.02)	(-0.72)	(1.1)	(3.08)	(0.34)

[a] Omitted variable: Zinfandel
[b] Omitted variable: no additional label information
c Omitted variable: year 2000
- Variable not present in market segment.

Table 3.5. Wald Statistics (p-values) Testing the Hypothesis of Parameters Equality Across Market Segments.

	Semi-Premium	Premium	Ultra-Premium
Commercial	7,866	19,112	11,118
	(0.000)	(0.000)	(0.000)
Semi-Premium		15,702	20,600
		(0.000)	(0.000)
Premium			9,838
			(0.000)

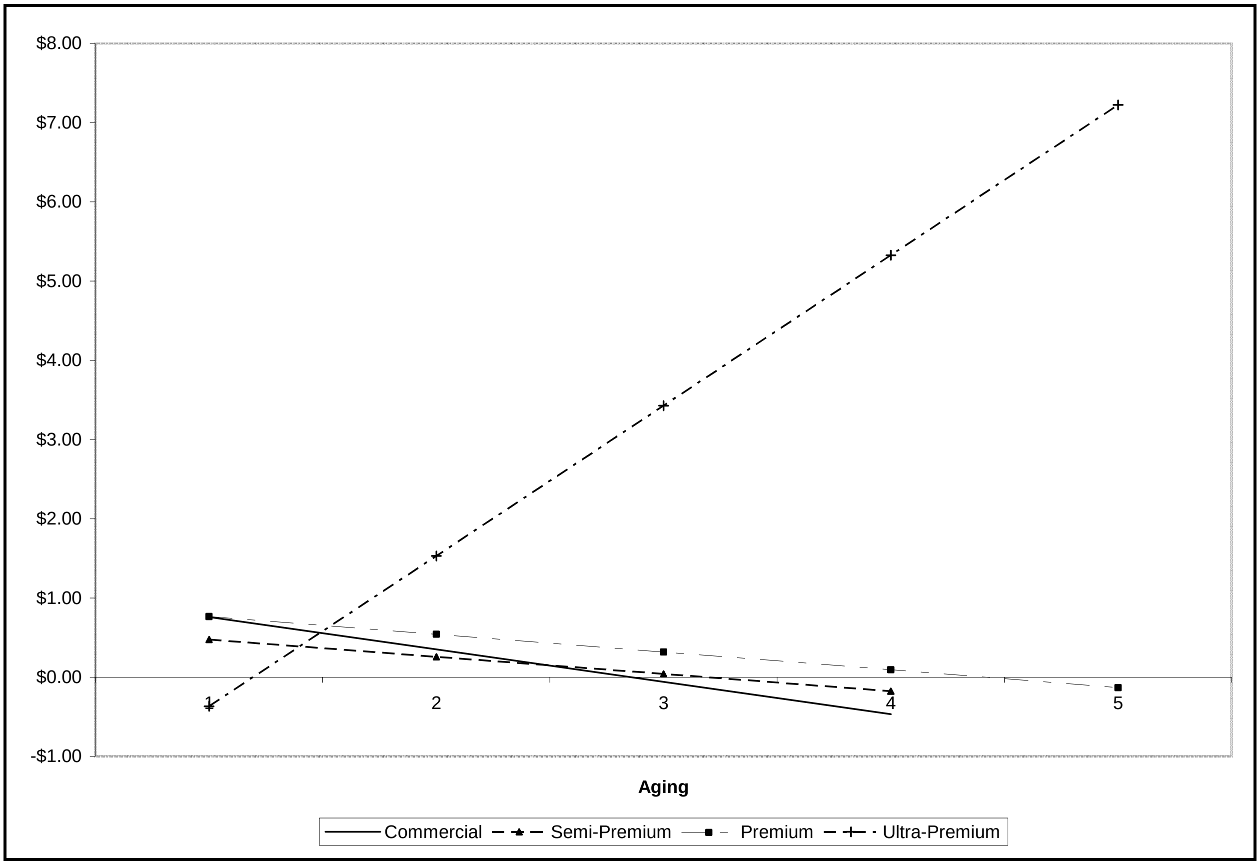

Figure 3.1. Implicit Price of Aging for Commercial, Semi-Premium Premium and Ultra-Premium Wines. Calculated Using Estimates From the Segmented Model and Class-Specific Price Averages

CHAPTER FOUR

LET THE MARKET BE YOUR GUIDE: ESTIMATING EQUILIBRIA IN DIFFERENTIATED PRODUCT MARKETS WITH CLASS-MEMBERSHIP UNCERTAINTY

Chapter Abstract: a method of identifying and analyzing differentiated product classes when product membership is uncertain is presented. The procedure, *local polynomial regression clustering*, first estimates a hedonic model nonparametrically via local polynomial regression and then aggregates data into data clusters that share functionally similar estimates of the (local) hedonic functions, identifying product classes based on similarities in market equilibria. Finally, parametric hedonic functions specific to each product class are estimated, revealing market-specific differences in implicit prices of the attributes across classes. The methodology is presented and applied in the illustrative context of markets for wines, resulting in the identification and analysis of four distinct product classes.

Keywords: local polynomial regression clustering, hedonic regression, wine

Introduction

Economists have long been interested in markets for differentiated products. Examples include wine, automobiles, and real estate. When differentiated products are located farther apart in product space, they no longer compete against each other (Hotelling, 1929). Based on these examples, one can hypothesize that product space is likely multi-dimensional and the boundaries of product classes are uncertain. The hedonic approach (Rosen, 1974) was developed to estimate implicit prices of product attributes *within* a given product class. It is assumed that the goods in question are somewhat differentiated, but similar enough that consumers consider them as variations of the same product. However, as two products diverge, the market valuation of the attributes included in them will diverge as well.

Acknowledging this process in modeling is challenging since a clear-cut classification rule generally does not exist, but the cost of ignoring the problem is potentially faulty estimates of implicit attribute prices derived from the hedonic price function. On the other hand, deriving accurate information on the marginal contribution of each attribute can help firms in their pricing and production decisions. It is with this motivation that I develop a method of identifying and analyzing differentiated product classes when product membership is uncertain. The methodology is presented and applied in the illustrative context of markets for wines, resulting in the identification and analysis of four distinct product classes.

The characteristics that differentiate one wine from another and how the market values such differences are intriguing research questions. Major factors influencing the purchase of wine include previous experience and knowledge of the product, objective cues such as production region, brand, and label, the occasion in which the wine will be consumed, and the price itself (Spawton, 1991). Several authors utilize the hedonic approach to investigate the determinants of wine prices with substantial agreement on the important characteristics. Combris *et al.,* (1997, 2000) show that when regressing objective characteristics and sensory characteristics on wine price, the objective cues (such as vintage) are significant, while sensory variables such as tannins content and other measurable chemicals are not. Nevertheless, evidence indicates that ratings by specialized magazines are significant and should be included in the hedonic function when modeling wine prices (Oczkowski, 1994; Landon and Smith, 1997; Shamel and Anderson, 2003, Angulo *et al*., 2000). In addition to expert ratings, the region of production, and the vintage are often reported as significant variables (Angulo *et al*., 2000; Schamel and Anderson, 2003).

The hedonic approach specifies a single function in which the product's price depends on its characteristics. Regressing price on the attributes provides estimates of their implicit prices, embedding supply and demand factors, cost, and willingness to pay (Nerlove, 1995). Previous hedonic studies of wine commonly estimate a single price function for all wines, implicitly assuming that there is limited variation over product characteristics, or, if the variation is substantial, consumers' reactions to it are either limited and/or they are somewhat unable to appreciate it. A strong argument can be made against the assumption of limited variation (real or perceived) in wine. The general product category of wine is a composite product class, which includes a number of product classes. “Two-Buck Chuck” and a bottle of Penfolds Grange have little in common and will be consumed on radically different occasions. Between these two extremes, a continuum of product differentiation exists. Wines sharing similar characteristics compete with each other as substitute goods. However, beyond a certain level of differentiation, two wines are no longer substitutes, and consumers might consider them to be different products. Hall *et al.,* (2001) supports one aspect of this phenomenon with his findings that consumers value the same attributes differently, depending on the occasion in which the wine is meant to be consumed.

The real estate hedonic literature offers a clarifying example of composite product classes. Researchers routinely estimate separate hedonic functions for type or use of structure (e.g. single family houses, multi-family units, and office buildings) or geographic area (Straszheim, 1974). They expect the implicit prices of the attributes to differ significantly across classes, and they therefore model product classes separately. A similar argument can be made for wine: certain wines are so differentiated from others that they belong to separate product classes. Unfortunately, identifying classes for wine is not as straightforward as in the real estate example, but classifications such as “premium” and “collectible” are widely used to separate these products from less expensive and refined wines.

A use-based classification of products is helpful in better formalizing how composite product classes originate. Intuitively, as the vectors of attribute levels characterizing two goods increasingly diverge, it is more likely that consumers purchase the goods for different purposes. That is, the more two products are differentiated, the less they are fungible in the same or similar use. Further, the costs associated with assembling a given bundle of attributes in the same product will change as the vector of attributes changes. This transition process can be either gradual, perhaps generating hybrid products that can be employed for multiple purposes, or the presence of an attribute might unambiguously determine the membership of a good to a given product class.

Identification of product class membership is important because firms must understand who their competitors are in each market to develop successful production and marketing strategies. A study commissioned by Australian wine producers and conducted by Ernst & Young Consulting (1999) recognized that "multiple wine markets" exist and attempted to identify such markets. Their classification divides wines into commercial, semi-premium, premium and ultra-premium, mostly on the basis of retail price ranges.[8]

The current study takes a different approach to addressing the problem of identifying and analyzing differentiated product classes. Using nonparametric techniques, local observation-specific estimates of hedonic price functions are produced. Significant differences in the (local) hedonic functions reflect changes in the underlying

---

[8] The report specifies that the price ranges for commercial, semi-premium, premium and ultra-premium are, respectively: less than A$8 or A$10, between A$10 and A$15 or A$20, between A$20 and A$30 or A$50, more than A$30 or A$50.

supply and demand functions and represent an objective criterion for determining product classes. This approach has two fundamental benefits: 1) the classification is embedded in the estimation approach and is objectively determined from observed data, and 2) it does not rely on a detailed structural model of the process by which product classes are determined, mitigating the potential for model misspecification. Once product classes are determined, implicit prices of product attributes specific to each product class are obtained.

The methodology section that follows presents a brief survey of current literature on the estimation of class-specific parametric models under class membership uncertainty. Alternative approaches are compared and discussed, and then the local polynomial regression clustering method used in this study is presented. Next, the data is described, along with the final empirical specification of the hedonic models. Results are then reported for the application to wine markets, including the wine characteristics that are typical of each wine class and, for every product class, I present the implicit price of each wine attribute. The discussion includes an analysis of the major differences across wine classes and economic implications of the differences.

Methodology

Estimation of class-specific hedonic functions when there is class membership uncertainty is a challenging task, and the method by which the product classes are identified can influence results significantly. Partitioning by price in the spirit of Chapter

Three is one possible approach. However, the aforementioned multifaceted rationale underlying the genesis of composite product classes provides support for investigating additional classification schema. Any classification scheme should facilitate the identification of product classes whose product prices are all well-represented by class-specific implicit price functions of the respective attributes of the products. For example, assume linear hedonic price functions for m product classes, as

(1) $$p_{ij} = x_{ij}\beta_j + \varepsilon_{ij},\ j = 1,...,m$$

where p_{ij} is the price of product i belonging to product class j, x_{ij} is a row vector of associated attributes and β_j is a vector of coefficients, or equivalently implicit prices in this case. The objective is to identify product classes that allow one to estimate hedonic price functions in which the vectors of parameters can vary across product classes, but are stable within classes.

Existent Class-Specific Methodology

A problem in estimating hedonic models like equation (1) is that it lends itself to multiple methods of implementation and several alternative approaches can be found in the literature. While the methods differ widely in technical aspects, they can be categorized into two alternative, but related general ideas. One approach consists of first partitioning the data into classes according to some data similarity criterion, and then estimating class-specific regression models within each data partition. The second method is based on specifying more complex models that simultaneously estimate class membership of the observations and associated class-specific vectors of parameters.

Straszheim (1974) provides a simple example of the data partition approach. In order to identify market segments in the San Francisco Bay Area real estate market, he divides the dataset *a priori* into geographical areas and estimates group-specific hedonic models. His results show that the within sample predictive ability of the segmented model is superior to a model ignoring market segments. This generated a subsequent stream of studies using multivariate analysis to identify data partitions that were relatively homogeneous under some distance criterion measuring data similarity. Data clustering algorithms (Ward and K-means are frequent choices) are often applied either to the original data or to factor and principal component scores relating to the data. Dale-Johnson (1982) and Watkins (1999) provide examples of the application of factor analysis, and Bourassa *et al.*, (1999 and 2003) apply clustering algorithms to principal component factor scores. Wilhelmsson (2004) adopts a slightly different method, clustering on the residuals from a regression performed using the entire dataset, ignoring submarkets.

An alternative to these two-step procedures consists of specifying a likelihood function that embeds an explicit model for determining class membership, thereby accounting for the existence of product classes. Maximization of the likelihood function involves simultaneous estimation of class membership of the sample observations and the parameter estimates of the associated class-specific regression functions, which must usually be done numerically using iterative computer-intensive function-maximization algorithms. The resulting models are often complicated to interpret, and the estimated coefficients are often not straightforwardly usable, *per se*, for deriving model implications such as implicit prices. General examples are random parameter models (see Allenby and Rossi, 1999), latent class models (see Greene, 2001 for a survey of literature) and mixture approaches that are used to model unexplained heterogeneity (for example, Arcidiacono and Jones, 2003). Goodman and Thibodeau (1998) provide an application in the hedonic framework, estimating a two-stage hierarchical model of market segmentation in real estate.

Both of the preceding approaches have strengths and weaknesses. Data clustering procedures are relatively simple to implement and yield models that are both generally tractable to estimate and readily interpretable. The main concern with data clustering procedures is their effectiveness in identifying data partitions whose observations correspond to the appropriate product class relationships. Bourassa (2003) finds that the performance in out-of-sample prediction of models that identify market segments using multivariate data clustering techniques is poor when compared to models that partition data *a priori* based only on location. The reasons underlying poor performance can be multiple, but at least one appears prominent. In particular, there is no general theoretical rationale for expecting that partitioning data on the basis of similarity in the values of covariates identified through clustering algorithms will result in data partitions that have associated regression functions that are functionally stable in terms of their parametric representation.

The likelihood-based approaches have the attractive feature of embedding the existence of separate product classes within a general nested model framework. However, as is the case in hierarchical models, researchers are forced to formally specify the model depicting the process that leads to changes in parameter values. Even though statistical software is available for estimating random parameter and latent class models, the presence of multiple classes and the need to calculate parameter vectors for each of the classes can lead to convergence issues, exacerbates the risk of local minima, and there is the general problem of finding good starting values for the numerical algorithms. Moreover, the potential for model specification errors are aggravated by the requirement that the process determining class membership must be an explicit part of the overall specification of the regression model.

Given the challenges and problems of existing approaches, I propose and implement *Local Polynomial Regression Clustering* (LPRC). The approach mitigates some of the principal shortcoming of the available methodologies, and strives to estimate multiple, class-specific hedonic functions that are stable in their parametric representations.

Local Polynomial Regression Clustering

In order to motivate the general approach, consider a collection of m relationships between dependent variables and covariates, as

(2) $$\boldsymbol{y}_{ij} = g_j(\boldsymbol{x}_{ij}\,\boldsymbol{;}\,\beta_j) + \boldsymbol{\varepsilon}_{ij}\,,\ i = 1,\ldots,n_j\,,\ \ j = 1,\ldots,m\,,$$

where n_j denotes the number of sample observations associated with the j^{th} relationship. In the current general application context, this system would represents *m* different hedonic price functions associated with *m* different wine classes. Using the entire collection of sample data $(\boldsymbol{y},\boldsymbol{x})$ consisting of $n = \sum_{j=1}^{m} n_j$ observations on prices and attributes, as well as any auxiliary data $\boldsymbol{z}$ relevant to the determination of product classes, represent the set of m relationships on the data partition $\bigcup_{j=1}^{m} D_j = D$, as

(3) $$\boldsymbol{y}_{ij} = \sum_{j=1}^{m} g_j(\boldsymbol{x}_{ij}\,\boldsymbol{;}\,\beta_j)\mathrm{I}_{(D_j)}(\boldsymbol{x}_{ij}\,,\boldsymbol{z}_{ij}) + \boldsymbol{\varepsilon}_{ij}\,,$$

where $\mathrm{I}_{(D_j)}(x,z)$ is an indicator function that equals 1 when $(x,z) \in D_j$, and equals 0 elsewhere. Let

(4) $$\boldsymbol{b}_j^r(\boldsymbol{x_0}) \equiv \left[\left.\frac{\partial g_j(\boldsymbol{x;}\beta_j)}{\partial \boldsymbol{x}}\right|_{\boldsymbol{x_0}} \quad \left.\frac{\partial^2 g_j(\boldsymbol{x;}\beta_j)}{\partial \boldsymbol{x}^2}\right|_{\boldsymbol{x_0}} \quad \cdots \quad \left.\frac{\partial^r g_j(\boldsymbol{x;}\beta_j)}{\partial \boldsymbol{x}^r}\right|_{\boldsymbol{x_0}} \right],$$

be the $k \times r$ matrix of derivatives of the function $g_j(\bullet)$, up to order *r*, at an evaluation point $(\boldsymbol{x_0}, \boldsymbol{z_0}) \in D_j$.

The overall objective of LPRC is to identify a partition $\bigcup_{j=1}^{m} \hat{D}_j = D$ of the observed sample data such that the estimated model $\boldsymbol{y}_{ij} = \sum_{j=1}^{m} \hat{g}_j(\boldsymbol{x}_{ij}; \hat{\beta}_j) \mathrm{I}_{(\hat{D}_j)}(\boldsymbol{x}_{ij}, \boldsymbol{z}_{ij}) + \boldsymbol{v}_{ij}$ approximates well the relationship between y_{ij} and each of the associated elements of $\boldsymbol{x}_{ij}$ up to the r^{th} order derivative relationship. Viewed in the context of a hedonic relationship, the goal is to approximate implicit price relationships between y_{ij} and the product attributes $\boldsymbol{x}_{ij}$, which at a point of evaluation and for the ℓ^{th} attribute, are of the form

(5) $$\boldsymbol{P_{ij}}[\ell] = \left. \frac{\partial g_j(\boldsymbol{x})}{\partial x_{ij}[\ell]} \right|_{\boldsymbol{x_0}}, \text{ and } \frac{\partial^{h-1} \boldsymbol{P_{ij}}[\ell]}{\partial x_{ij}^{h-1}[\ell]} = \left. \frac{\partial^h g_j(\boldsymbol{x})}{\partial x_{ij}^h[\ell]} \right|_{\boldsymbol{x_0}} \text{ for } h = 2, 3, \ldots, r.$$

This goal is pursued by using local polynomial regression of order r to generate local nonparametric observation-specific estimates of the derivative relationships $\boldsymbol{b}_j^r(\boldsymbol{x}_0)$ as

(6) $$\hat{\boldsymbol{b}}^r(\boldsymbol{x_0}) = \left[\left. \frac{\partial \hat{g}(\boldsymbol{x}, \boldsymbol{z})}{\partial \boldsymbol{x}} \right|_{\boldsymbol{x_0}, \boldsymbol{z_0}} \quad \left. \frac{\partial^2 \hat{g}(\boldsymbol{x}, \boldsymbol{z})}{\partial \boldsymbol{x}^2} \right|_{\boldsymbol{x_0}, \boldsymbol{z_0}} \quad \cdots \quad \left. \frac{\partial^r \hat{g}(\boldsymbol{x}, \boldsymbol{z})}{\partial \boldsymbol{x}^r} \right|_{\boldsymbol{x_0}, \boldsymbol{z_0}} \right]$$

for $(\boldsymbol{x_0}, \boldsymbol{z_0}) \in D$, where $\boldsymbol{y} = \hat{g}(\boldsymbol{x,z}) + \hat{\boldsymbol{\varepsilon}}$ denotes the underlying nonparametric regression that is estimated locally at each data point, and then clustering into *m* classes on the basis of similarity in the values of $\hat{\boldsymbol{b}}^r(\boldsymbol{x_0})$. The objective of this type of clustering is the achievement of similarity in implicit price functions up to a given order of derivative relationship, as opposed to the usual clustering approach that is based on defining clusters in terms of similarity of the covariate values in the sample of data.

In the application of LPRC to wine data, presented in sections 3 through 5 ahead, partitions of the data represent the specific empirical domains of the hedonic price functions corresponding to different wine classes. In the empirical application, the first step of LPRC is implemented using local linear regressor (LLR), a nonparametric technique motivated by the well known Weierstrass polynomial approximation theorem. More specifically, a first-order local linear regression function approximation $\boldsymbol{y} = \hat{g}(\boldsymbol{x,z}) + \hat{\boldsymbol{\varepsilon}}$ is estimated for each of the *n* evaluation points $(\boldsymbol{x_0}, \boldsymbol{z_0}) \in D$ as

(7) $$\underset{\hat{a}(\boldsymbol{x_0}),\, \hat{\boldsymbol{b}}(\boldsymbol{x_0})}{\arg\min} \left\{ \sum_{i=1}^{n} [y_i - \hat{a}(\boldsymbol{x_0}) - (\boldsymbol{x}_{i.} - \boldsymbol{x}_0^{'})\hat{\boldsymbol{b}}(\boldsymbol{x_0})]^2 K((\xi_{i.} - \xi_0^{'})/h) \right\}$$

where *n* is the number of observations in the data sample, $K(\bullet)$ represents a kernel weighting function, ξ refers to the data used in defining the kernel weighting in the weighted local regressions where ξ_0 corresponds to the vector point of evaluation and $\boldsymbol{\xi}_{i.}$ represents the ith row in ξ, *h is* the bandwidth in the kernel weighting function, and

$\left.\frac{\partial \hat{y}}{\partial \boldsymbol{x}}\right|_{\boldsymbol{x_0}} = \hat{\boldsymbol{b}}(\boldsymbol{x_0}), \ \forall \boldsymbol{x_0}$. Focusing on local first-order derivative relationships, as is done in (7), implies that $r = 1$ in (4) and (6) so that hedonic functions across wine classes are differentiated on the basis of similarities to the first order of approximation, i.e., similarities in first order derivatives.

The LOESS algorithm (Cleveland *et al.*, 1988), based on the tricube weight function, is used in the minimization of the weighted sum of squares function in (7) so that the explicit definition of the kernel function is

$K((\xi_{i.} - \xi_0^{'})/h) = W\left([(\xi_{i.} - \xi_0^{'})(\xi_{i.} - \xi_0^{'})']^{1/2}/h\right)$, where $W(u) = (1 - |u|^3)^3$ for $|u| \in [0,1]$, and $W(u) = 0$ otherwise. The LOESS approach is based on the use of a nearest-neighbor type of bandwidth in which h not only affects the relative weights applied to sample observations used in the local regression estimates, but also determines the proportion of the sample that is included in the estimation procedure. Several ways of determining the optimal value of the bandwidth have been proposed, and in this application we use the least squares cross validation approach presented by Stone (1974).

The $n \times k$ matrix of estimated partial derivatives, $\hat{\boldsymbol{B}} = \left[\hat{\boldsymbol{b}}(\boldsymbol{x_0})', \boldsymbol{x_0} \in D\right]$, is the object of the second stage analysis in the LPRC procedure. In this clustering step, the Ward algorithm is used to group estimates on the basis of similarity of the $\hat{\boldsymbol{b}}(\boldsymbol{x_0})$ values, thereby identifying domains of data with similar hedonic implicit price functions and exhibiting similar implicit wine attribute prices. The Ward algorithm minimizes, through the choice of the clusters $\hat{D}_j,\ j=1,\ldots,m$ and the associated data partition $\bigcup_{j=1}^{m}\hat{D}_j = D$, the sum of squared deviations (SSD) from the cluster centroids $\hat{\bar{\boldsymbol{b}}}_j$, j= 1,…, *m*, i.e., Ward's algorithm solves the minimization problem:

$$(8) \qquad \min_{\hat{D}_j,\ j=1,\ldots,m} \sum_{j=1}^{m} \sum_{\boldsymbol{x_0}\in\hat{D}_j} (\hat{\boldsymbol{b}}(\boldsymbol{x_0}) - \hat{\bar{\boldsymbol{b}}}_j)'(\hat{\boldsymbol{b}}(\boldsymbol{x_0}) - \hat{\bar{\boldsymbol{b}}}_j).$$

The algorithm begins by treating each observation as a cluster, and then at each iteration of the algorithm, the number of clusters is reduced as all possible unions of the existing clusters are considered. The union that minimizes the SSD over all clusters is implemented, and the algorithm continues aggregating clusters until the target number of clusters is obtained. Based on the out-of-sample information discussed previously, I set $m = 4$ so that the last iteration of the algorithm partitions the dataset into four subsets, corresponding to four wine classes.

Once the partition of the data into *m* product classes is identified, the final step of the local polynomial regression clustering algorithm is the estimation of *m* class-specific hedonic functions for the m subsets of the data. These hedonic functions represent the empirical counterpart to models such as in equation (1), and by virtue of having been fit to subsets of data for which derivatives of the regression model are made to be similar via the clustering step, bias in the estimation of hedonic function derivatives is expected to be mitigated..

Summarizing the steps in the LPRC algorithm, the first step provides a nonparametric flexible hedonic model that can represent, through the nonparametric local polynomial regression approach, a wide range of implicit relationships between product characteristics and prices. The clustering step then identifies groups of sample observations that share similar estimated implicit price functions and identifies the domains over which simpler, more structured global hedonic functions apply. The resultant class-specific hedonic functions are then estimated in the final step.

Some readers may question whether the second and third steps of the procedure are worth the additional effort, since a non-parametric estimate of the hedonic function is already available from the first step. There are multiple reasons to do so. As discussed earlier, the wine industry recognizes the existence of four wine classes (e.g. Ernst & Young Consulting, 1999), while the non-parametric approach ignores this information and instead treats each observation uniquely by virtue of estimating a separate hedonic relationship at each observation value. Furthermore, it is of interest to both identify the wine classes as well as concomitantly characterize them in terms of their attributes. Finally, cluster-specific parametric regression models that are parametrically parsimonious and functionally non-complex can provide estimates of the implicit prices of attributes that can be highly efficient as well as readily interpretable.

While perhaps not obvious, LPRC has some similarities with likelihood-based approaches. In fact, Ward clustering received attention in the model-based Gaussian hierarchical clustering literature. As Fraley (1998) shows, if the process generating the multiple class dataset can be modeled using a multivariate normal distribution with class-specific means and common variances, then using the Ward algorithm to identify classes is equivalent to maximizing the classification likelihood function.[9] Therefore, under certain distributional assumptions on the observation-specific local polynomial regression coefficients (which researchers often seem willing to make when adopting likelihood based procedures), LPRC can be considered as a likelihood-based method.

An approach that is similar in spirit, but notably different in implementation and estimation objective is that of so-called "Partition Regression" introduced by Guthrey (1974). The method seeks to minimize the sum of squared errors over the choice of a fixed number of partitions, and implementation is effectively relegated to data that is intrinsically ordered in some fashion, as in a time series. Another recent approach by Christopeit and Hoderlein (2006), called "Local Partitioned Regression", is a method that significantly extends the reach of nonparametric regression models beyond additive models, and can incorporate local polynomial approximations in the process, but remains a local observation-specific method that does not lead to the identification of models differentiated by, and associated with product classes.

---

[9] More formally, modifying Fraley's (1998) model, we denote by J the total number of classes and the density of a p-dimensional observation x_i from the jth class by $f_j(x_i \mid \theta_j)$, for some vector of parameters θ_j. Given observations $x=(x_1,...,x_n)$, let $\gamma = (\gamma_1,...,\gamma_n)$ denote identifying labels for the classification, where $\gamma_i = j$ if x_i comes form the *jth* subpopulation. Then it can be shown that the Ward algorithm maximizes the classification likelihood: $L(\theta_1,...,\theta_J;\gamma \mid x) = \prod_{i=1}^{n} f_{\gamma_i}(x_i \mid \theta_{\gamma_i})$, under the assumption that $f_j(x_i \mid \theta_j)$ is the multivariate normal distribution with mean $\mu = (\mu_1,...,\mu_J)$ and variance $\Sigma_j = \sigma^2 I$.

Data

The data set is composed of 9,820 observations[10] derived from ten years (1991-2000) of tasting ratings reported in the *Wine Spectator* Magazine (online version) for California (8,848 observations) and Washington (972 observations) red wines. Table 1 presents variable names and abbreviation, along with a short data descriptions. Four of the variables in the dataset are non-binary: price of wine (adjusted by the 2000 consumer price index for alcohol), tasting score obtained in the expert sensory evaluation provided by the *Wine Spectator*, the number of cases produced, and the years of aging before commercialization. Descriptive statistics for these variables are reported in table 2. Note that wine prices have a positively skewed distribution, but most observations fall in the $10 to $50 range. Indicator variables were used to denote region of production, wine variety, and the presence of other label information. The regions of production for California wines include Napa Valley, Bay Area, Sonoma, South Coast, Carneros, Sierra-Foothills, and Mendocino. Washington wines were not separated by regions.[11] Varieties include Zinfandel, Pinot Noir, Cabernet, Merlot and Syrah grapes, as well as wines made from blending different varieties (non-varietals).

[10] This dataset is a random sample including about 70% of the larger dataset used in Chapter Three. The remaining 30% was used to validate LPRC and assess out of sample predictive performance.

[11] These geographical partitions are the ones adopted by the *Wine Spectator* to categorize the wines, often pooling several American Viticultural Areas (AVAs) in the same region.

Additional variables refer to information appearing on the label of certain bottles. Some wines are produced from the vineyards adjacent to the wineries ("estate" wines), others are classified as "reserve" by the wine maker. I denominate this kind of marketing device by the term "patronage," whereby wineries try to associate a wine to the bucolic image of the winery, or the experience of the wine maker. Finally, the vintage year is available for each wine.

Empirical Model Specification and Estimation

To implement LPRC, the argument of the Kernel function needs to be chosen. Ernst & Young Consulting (1999) and the results obtained in Chapter Three we can conclude that price is heavily linked to the process of class differentiation; indicating that price is a reasonable candidate for the argument of the Kernel function. We carry out the first step of LPRC in a neighborhood of price.

To avoid simultaneity issues arising from defining $\boldsymbol{\xi} = \boldsymbol{P}$, I implement an instrumental variables-type of approach and generate predicted prices using an iterative process. First, a vector of predicted prices is estimated via least squares using the complete dataset using the model in equation (9) below[12]

$$P^{-0.5} = \alpha_0 + \beta_1(Score) + \beta_2(Age) + \beta_3(Cases) + \sum_{i=1}^{5}\beta_{4+i}(Variety) \\ + \sum_{i=1}^{9}\beta_{9+i}(Vintage) + \sum_{i=1}^{3}\beta_{18+i}(Label) + \sum_{i=1}^{7}\beta_{21+i}(Region) + \varepsilon \quad . \tag{9}$$

The specification in (9) is an unconditional (on class) regression of the reciprocal square root of price on wine attributes. While useful predictions of price can be obtained from this unconditional regression, it is understood that implicit prices of attributes derived from the regression are generally expected to be biased and inconsistent relative to class-specific values given the existence of different wine classes. The predicted prices are then used in an application of LLR, applied individually to each evaluation point in the sample of observations, by solving (7) with the regressand y_i defined as the reciprocal square root of price, the regressors, $\mathbf{x}_{i.}$, are specified as in (9),

---

[12] The inverse square root transformation of the dependent variable was chosen instead of the more traditional semi-log specification based on a Box-Cox type grid search of possible transformations, which included the semi-log specification. A battery of tests for each transformation was used to identify the best choice for variance stabilization (Goldfeld-Quandt), normality of the residuals (Anderson-Darling and Komolgorov-Smirnov), and proper specification (RESET). Also notice the inclusion of original quantity produced (*cases*) in the right-hand side of the hedonic function. While (9) is *not* an inverse demand function, we estimate β_3 to capture the effect that “rarity” has on certain cult wines.

and the argument, $\xi_{i.}$, of the kernel density tri-cube weighting function is set equal to $\left(\hat{P}_i - \hat{P}_0\right)/h$ from (9). The optimal size of the bandwidth, *h*, is determined by cross-validation. The results of the LLR are used to generate the entries of the matrix of first order derivatives $\hat{\boldsymbol{b}}(\boldsymbol{x_0})$[13].

The second step of LPRC was implemented using the Ward algorithm to identify four wine classes, as discussed in section 2. I focus the search on similarity of the elements of $\hat{\boldsymbol{b}}(\boldsymbol{x_0})$ that refer to the derivatives of continuous wine attributes, and thus focus on attributes for which derivatives are literally defined. Finally, for each wine class a parametric hedonic function is estimated, producing conditional class-specific counterparts to the relationships depicted in equation (9)[14].

---

[13] The LLR estimation was iterated to assess and achieve stability of the estimates, where predicted prices obtained from the initial LLR fit of the hedonic model were used in the second LLR application. We find that an additional iteration results in stable estimates of the parameters of the local linear regressions.

[14] The out of sample predictive performance of LPRC was tested on an additional sample of 3265 observation. The methodology was compared to a price range partition approach in the spirit of Ernst and Young (Chapter Three) and three alternative clustering approaches drawn from the real estate hedonic literature (namely: cluster on non binary variables, on principal component factor scores and on OLS residuals). The predictive performance was assessed calculating a Median Percent Error Rate for each approach. The LPRC was, by far, the best performing approach. Complete results are available from the authors upon request.

Results

Within-class descriptive statistics of the non-binary and binary variables can be used to identify and characterize wine classes (see tables 3 and 4 respectively). Based on LPRC, eleven percent of the observations can be classified as commercial, 32 percent are semi-premium, 33 percent are premium, and 22 percent are ultra-premium. The small percentage of commercial wines is likely a consequence of the nature of the sample, as the *Wine Spectator* focuses on tasting higher quality wines[15].

Parameter estimates are reported in table 5 for the hedonic function associated with each wine class. The large majority of the estimated coefficients are statistically significant. The coefficients are not particularly interesting in and of themselves because of the nonlinear transformation of the dependent variable. I calculate the implicit prices of the attributes (i.e. the derivative of total price with respect to the attributes), which are functions of both the estimated coefficients and prices. Table 6 shows, for each wine class, the estimated average implicit prices (with a 95 percent confidence interval)[16] of

[15] It is plausible that the lower quality spectrum of the commercial wine class is underrepresented in the sample. Of course, results apply generally to American wines, and not only to the population of wines evaluated by the Wine Spectator, only to the extent that the Wine Spectator observations are a representative random sample of the American wine market.

[16] A confidence interval not including zero implies that the corresponding estimated coefficient is statistically significant at the α =0.05 level. For the same attribute, non-overlapping confidence intervals across wine classes implies rejecting, at the α =0.05, the null hypothesis that two estimated coefficients are equal. Note that the probability of a type I error is inflated if we consider multiple comparisons.

each attribute included in the hedonic function, which are calculated using the within-class average price.

The first and most important result in tables 5 and 6 is that parameter estimates and implicit prices of the attributes differ significantly across wine classes. This supports the class-specific hedonic regression approach and suggests that estimating a single hedonic regression for all wines will result in misleading estimates of the implicit prices. In addition, for any given product attribute, a general trend common to all estimated implicit prices is that variances increase from the commercial to the ultra-premium wines, resulting in wider confidence intervals. This suggests that a pooled model ignoring wine classes would be heteroskedastic.[17]

For the sake of both brevity and clarity, I focus additional discussion of implicit prices on a comparison of commercial and ultra-premium wines, but significant differences can be found when comparing each pair of wine classes. Commercial wine prices are relatively insensitive to variations in the quantity produced in a given vintage: a one hundred case increase in a given vintage lowers price about a tenth of a cent. Unitary increases in tasting scores (figure 1) and years of aging are positively correlated, with *ceteris paribus* price premia of \$0.16 and \$1.20, respectively. The viticultural areas commanding the highest price premia with respect to generic California wines are, in the commercial class, Carneros and Napa Valley (figure 2). Interestingly, Washington wines

---

[17] An LM test for group-wise heteroskedasticity rejects the null hypothesis of equality of variance across the four class-specific hedonic regressions at the $\alpha = .05$ level.

also obtain positive price premia in the commercial class. Shyrah and Pinot noir varietals are the most expensive commercial wines.

The situation is quite different in the ultra-premium class. The negative impact on price of an increase in quantity produced in a given vintage is substantially larger than the commercial wines estimate. The implicit price of aging raises to almost six dollars per year[18], and each additional point in the wine spectator tasting evaluation is worth almost four dollars. These results confirm that ultra premium wines have a collectible value, directly related to the rarity of the wine, and suggest that the association between price and quality is more pronounced in the ultra premium class than in the commercial class. Furthermore, price premia associated with the region of production also increases by several orders of magnitude, with the exception of Washington wines, which are valued the same as generic California wines. In contrast with the results relative to the commercial class, the most expensive ultra premium wines are made via blending multiple grapes, followed by Pinot Noir varietals.

[18] A gross (inclusive of storage costs) yearly rate of return on wine aging can be calculated for each class using estimates of the implicit price of aging and mean prices. These rates are, from commercial to ultra-premium, 13%, 11%, 10% and 13% respectively. As these rates represent the returns at which producers find it convenient to introduce the wines into the market, economic theory suggest that they should be similar.

Summary and Conclusions

There are two main contributions of this study. First, I present and implement a new approach for estimating product class-specific parametric models with class-membership uncertainty called Local Polynomial Regression Clustering (LPRC). Second, we present empirical evidence that wine is a composite product class. Following the industry's findings of four wine classes, we are able to objectively characterize and differentiate these classes. Realizing the composite nature of wine, estimation by product classes allows the estimation of a set of differentiated and class-informative hedonic models. In particular, results show that implicit prices of the attributes are different across wine classes, supporting the notion that separate hedonic price function equilibria exist for these product classes.

This empirical application of LPRC yields implications for the wine industry in the form of class specific implicit prices of wine attributes, and also provides a method of identifying wines, and wineries, that compete in the same market segment. The approach could be applied with a greater level of generality in other empirical econometric settings. The advantages of LPRC over existing methods for identifying and analyzing differentiated product classes under class membership uncertainty are multiple: first, it allows researchers to introduce, relatively straightforwardly and generally, sample information relating to the process generating class-specific parameters values via the specification of kernel function arguments without requiring a detailed structural model specification of the process by which market observations are categorized into product classes. Second, no specific assumptions regarding the parametric family of distributions underlying the sample of data are necessary to justify the use of the approach. Finally, a unique solution, consisting of an optimal data partition and class-specific parameter estimates, always exists. The risk of local solutions and the need to find starting values in more complicated nested models of differentiated product classes is also mitigated. Future research is ongoing that is related to finding ways to reduce computational time and to making the number of classes endogenous to the LPRC approach in order to accommodate applications in cases where less or no non-sample information is available relating to the number of existent product classes.

References:

Allenby, G. and P. Rossi. (1999). "Marketing Models of Consumer Heterogeneity," *Journal of Econometrics*, 89:57-78.

Angulo, A.M., J.M. Gil, A. Gracia and M. Sanchez. (2000). "Hedonic Prices for Spanish Red Quality Wine," *British Food Journal*, 102:481-493.

Arcidiacono, P. and J.B. Jones. (2003). "Finite Mixture Distributions, sequential likelihood and the EM algorithm," *Econometrica*, 71:933-946.

Bourassa, S.C., F. Hamelink, M. Hoesli and B.D. MacGregor, (1999). "Defining Housing Submarkets," *Journal of Housing Economics*, 8:160-183.

Bourassa, S.C., M. Hoesli and V.S. Peng, (2003). "Do Housing Submarkets Really Matter?" *Journal of Housing Economics*, 12:12-28.

Christopeit, N. and S.G.N Hoderlein. 2006. "Local Partitioned Regression," *Econometrica*, 74:787-817.

Cleveland, W.S., S.J. Delvin and E. Grosse. (1988). "Regression by Local Fitting," *Journal of Econometrics*, 37:87-114.

Combris, P., S. Lecocq and M. Visser, (1997). “Estimation of a Hedonic Price Equation for Bourdeaux Wine: does quality matter?” *The Economic Journal,* 107:390-403.

Combris, P., S. Lecocq and M. Visser. (2000). “Estimation of a Hedonic Price Equation for Burgundy Wine,” *Applied Economics,* 32:961-967.

Dale-Johnson, D. (1982). “An Alternative Approach to Housing Market Segmentation Using Hedonic Price Data,” *Journal of Urban Economics,* 11:311-332.

Ernst and Young Entrepreneurs, “Etude des Filieres et des Strategies de Développement des Pays Producteurs de Vins dans le Monde: Analyse de la Filiere Viticole Australienne,” *ONIVINS (Office National Interprofessionnel des Vins)* (1999).

Fan, J. and I. Gijbels, “Local Polynomial Modeling and Its Applications,” Chapman & Hall (1996).

Fraley, C. (1998). “Algorithms for Model-Based Gaussian Hierarchical Clustering,” *SIAM Journal on Scientific Computing,* 20:270-281.

Goodman, A.C. and T.G. Thibodau. (1998). “Housing Market Segmentation”, *Journal of Housing Economics* 7:121-143.

Greene, W. (2001). “Fixed and Random Effects in Nonlinear Models,” *Working Paper EC-01-01*, Stern School of Business, Department of Economics, (2001).

Guthrey, S.B. (1974). “Partition Regression”, *Journal of the American Statistical Association*, 69:945-947.

Hall, J., L. Lockshin and G.B. O’Mahony. (2001). “Exploring the Links Between the Choice and Dining Occasion: Factors of Influence,” *International Journal of Wine Marketing,* 13:36-54.

Hotelling, H. (1929). “Stability in Competition.” *Economic Journal,* 39:41-57.

Landon, S. and C.E. Smith. (1997): “The Use of Quality and Reputation Indicators by the Consumers: The Case of Bordeaux Wine,” *Journal of Consumer Policy,* 20:289-323.

Nerlove, M. (1995). “Hedonic Price Functions and the Measurement of Preferences: the Case of Swedish Wine Consumers,” *European Economic Review* 39:1967-1716.

Oczkowski, E. (1994). “Hedonic Wine Price Function for Australian Premium Table Wine,” *Australian Journal of Agricultural Economics,* 38:93-110.

Rosen, S. (1974). "Hedonic Prices and Implicit Markets: Product differentiation in Pure Competition," *Journal of Political Economy,* 82:34-55.

Schamel, G. and K. Anderson. (2003). "Wine Quality and Varietal, Regional and Winery Reputations: Hedonic prices for Australia and New Zealand," *The Economic Record*, 79:357-369.

Spawton, T. (1991). "Marketing Planning for Wine," *European Journal of Marketing*, 25:2-47.

Stone, M. (1974). "Cross-Validatory Choices and Assessment of Statistical Prediction (with Discussion)," *Journal of the Royal Statistical Society*, 36:111-147.

Straszheim, M. (1974). "Hedonic Estimation of Housing Market Prices: A Further Comment," *Review of Economics and Statistics,* 56:404-406.

Watkins, K. (1999). "Property Valuation and the Structure of Urban Housing Markets," *Journal of Property Investment and Finance*, 17:157-175.

Wilhemsson, M. (2004). "A Method to Derive Housing Sub-Markets and Reduce Spatial Dependency," *Property Management*, 22:276-288.

Table 4.1. Short Descriptions and Abbreviations of Variables

Variable	Short Description	Binary/Nonbinary
Price	Retail Prices Suggested by the Winery	Nonbinary
Score	Rating Score from the Wine Spectator	Nonbinary
Age	Years of Aging Before Commercialization	Nonbinary
Cases	Number of Cases Produced	Nonbinary
Generic California*		
Napa Valley		
Bay Area		
Sonoma		
South Coast	Region of Production	Binary
Carneros		
Sierra/Foothills		
Mendocino		
Washington		
Zinfandel*		
Non-varietal		
Pinot Noir		
Cabernet	Grape Variety	Binary
Merlot		
Syrah		
Reserve	Patronage Information: "Reserve"	
Vineyard	Patronage Information: Specific Name of the Vineyard	Binary
Estate	Patronage Information: "Estate" Produced Wine	
91, …, 99, 00*	Vintage	Binary

* Benchmark variable (omitted) in the set of dummies variables in regression context.

Table 4.2. Descriptive Statistics of the Non Binary Variables (Whole Sample)

	Variable*			
	Price**	Cases	Score	Age
Mean	27.69	6,358	86.13	2.77
Median	22.00	1,200	87	3
SE Mean	0.20	255	0.04	0.01
Minimum	4.00	16	60	1
Maximum	150.00	900,000	99	9
First-Third Qurtile	[15.00-33.00]	[500-3,700]	[84-88]	[2-3]

* Sample size is 9820.

** Adjusted by a CPI price for alcohol.

Table 4.3. Descriptive Statistics of the Non Binary Variables by Wine Class

Variable	Decriptive Statistic	Wine Class: Commercial	Semi-Premium	Premium	Ultra-Premium
	N	1,097	3,172	3,315	2,229
	Percent	11.17%	32.32%	33.78%	22.71%
Price*	Mean	11.84	18.91	32.59	43.83
	SE Mean	0.13	0.13	0.25	0.58
	Median	11.11	17.51	30.00	35.00
	Minimum	4.52	6.00	9.00	8.32
	Maximum	54.00	89.27	150.00	150.00
	First-Third Quartile	[9.04-13.68]	[14.56-21.66]	[23.98-38.00]	[23.46-55.00]
Cases	Mean	31,334	4,805	2,256	2,388
	SE Mean	2,063	165	70	91
	Median	6,700	1,543	900	965
	Minimum	35	16	20	21
	Maximum	900,000	100,000	42,500	50,000
	First-Third Quartile	[1,600-31,200]	[600-4,500]	[400-2,075]	[450-2,235]
Score	Mean	81.43	84.38	87.62	88.70
	SE Mean	0.12	0.06	0.04	0.07
	Median	82	85	88	88
	Minimum	60	68	79	76
	Maximum	91	93	95	99
	First-Third Quartile	[79-84]	[83-87]	[86-89]	[87-91]
Age	Mean	2.31	2.65	2.86	3.04
	SE Mean	0.02	0.01	0.01	0.02
	Median	2	3	3	3
	Minimum	1	1	1	1
	Maximum	4	5	8	9
	First-Third Quartile	[2,3]	[2,3]	[2,3]	[3,3]

* Adjusted by a CPI price for alcohol.

Table 4.4. Percent Distribution of Binary Variables by and Across Wine Classes

		Distribution By Wine Class				Distribution Across Wine Classes				
		Commercial	Semi Premium	Premium	Ultra-Premium	Commercial	Semi-Premium	Premium	Ultra-Premium	Total
Region of Production	Generic California	59.9%	9.5%	1.0%	1.1%	83.8%	13.3%	1.4%	1.5%	100%
	Napa Valley	2.8%	14.9%	36.7%	49.6%	2.7%	14.4%	35.3%	47.7%	100%
	Bay Area	2.9%	5.9%	6.6%	4.6%	14.6%	29.3%	33.2%	22.9%	100%
	Sonoma	9.2%	28.6%	27.8%	21.0%	10.6%	33.0%	32.0%	24.3%	100%
	South Coast	3.4%	11.3%	9.9%	9.4%	9.9%	33.4%	29.1%	27.6%	100%
	Carneros	0.4%	5.3%	5.5%	4.2%	2.4%	34.5%	35.6%	27.5%	100%
	Sierra / Foothills	7.6%	4.2%	0.4%	1.6%	55.0%	30.2%	3.1%	11.7%	100%
	Mendocino	6.3%	6.5%	2.6%	2.6%	35.2%	36.2%	14.3%	14.3%	100%
	Washington	7.6%	13.9%	9.6%	5.9%	20.5%	37.6%	26.0%	15.9%	100%
	Total	100%	100%	100%	100%					
Grape Variety	Nonvarietal	1.6%	4.7%	7.2%	13.2%	6.1%	17.6%	27.0%	49.3%	100%
	Pinot	9.1%	15.2%	24.9%	17.0%	13.8%	22.9%	37.7%	25.6%	100%
	Cabernet	33.1%	20.2%	30.7%	43.0%	26.0%	15.9%	24.2%	33.9%	100%
	Merlot	22.7%	22.3%	17.7%	11.2%	30.7%	30.2%	23.9%	15.2%	100%
	Shyrah	3.6%	6.4%	8.5%	6.0%	14.9%	26.0%	34.7%	24.5%	100%
	Zinfandel	29.8%	31.2%	10.9%	9.6%	36.6%	38.3%	13.4%	11.8%	100%
	Total	100%	100%	100%	100%					
Patronage Information	Reserve	5.0%	8.4%	16.1%	15.4%	11.1%	18.7%	35.9%	34.3%	100%
	Vineyard	1.7%	10.9%	31.0%	25.8%	2.5%	15.7%	44.6%	37.1%	100%
	Estate	0.6%	1.2%	3.2%	3.3%	7.7%	14.4%	38.5%	39.4%	100%
	None	92.6%	79.5%	49.7%	55.5%	33.4%	28.7%	17.9%	20.0%	100%
	Total	100%	100%	100%	100%					

Table 4.5. Hedonic Regression Estimated Coefficients by Wine Class

	Commercial			**Semi Premium**			**Premium**			**Ultra Premium**		
Variable	Coefficient	T	p-Val	Coefficient	T	p-Val	Coefficient	t	p-Val	Coefficient	t	p-Val
Constant	0.51301	16.91	0.000	0.62990	25.79	0.000	0.77894	26.19	0.000	0.84257	40.39	0.000
Cases*	0.00002	7.05	0.000	0.00012	14.45	0.000	0.00020	14.05	0.000	0.00017	11.19	0.000
Score	-0.00202	-5.76	0.000	-0.00416	-15.47	0.000	-0.00624	-19.84	0.000	-0.00668	-29.67	0.000
Age	-0.01943	-8.18	0.000	-0.01231	-10.85	0.000	-0.00863	-9.41	0.000	-0.00990	-10.28	0.000
Napa Valley	-0.05377	-6.64	0.000	-0.03592	-10.90	0.000	-0.02844	-5.10	0.000	-0.04430	-7.21	0.000
Bay Area	-0.01105	-1.38	0.168	-0.01954	-5.23	0.000	-0.01243	-2.13	0.033	-0.03649	-5.47	0.000
Sonoma	-0.04493	-9.21	0.000	-0.02609	-9.12	0.000	-0.01192	-2.16	0.031	-0.03233	-5.25	0.000
South Coast	-0.04362	-5.77	0.000	-0.02325	-7.31	0.000	-0.00293	-0.51	0.608	-0.02323	-3.68	0.000
Carneros	-0.06813	-3.10	0.002	-0.02641	-6.65	0.000	-0.01327	-2.24	0.025	-0.02854	-4.21	0.000
Sierra/Foothills	-0.03183	-5.89	0.000	-0.01089	-2.64	0.008	-0.00898	-0.91	0.360	-0.01649	-2.10	0.036
Mendocino	-0.03218	-5.49	0.000	-0.00707	-2.01	0.045	-0.00038	-0.06	0.952	-0.01488	-2.08	0.038
Washington	-0.03879	-7.28	0.000	-0.00416	-1.38	0.166	0.01084	1.91	0.056	-0.00371	-0.56	0.573
Nonvarietal	-0.01164	-1.08	0.280	-0.02266	-6.29	0.000	-0.04552	-14.93	0.000	-0.04651	-15.27	0.000
Pinot	-0.01226	-2.36	0.019	-0.02703	-10.00	0.000	-0.03921	-15.84	0.000	-0.03954	-14.22	0.000
Cabernet	0.00250	0.69	0.493	-0.01820	-7.89	0.000	-0.03386	-14.33	0.000	-0.03566	-13.45	0.000
Merlot	-0.00788	-1.97	0.049	-0.02104	-9.48	0.000	-0.02596	-10.90	0.000	-0.02938	-9.88	0.000
Shyrah	-0.02635	-3.53	0.000	-0.00803	-2.63	0.009	-0.01332	-5.09	0.000	-0.01804	-5.24	0.000
Reserve	-0.01031	-1.65	0.100	-0.01506	-6.00	0.000	-0.00881	-5.77	0.000	-0.01082	-6.06	0.000
Vineyard	-0.01845	-1.80	0.072	-0.01760	-7.79	0.000	-0.01116	-8.46	0.000	-0.01027	-6.68	0.000
Estate	0.00633	0.38	0.707	-0.01008	-1.64	0.102	-0.00415	-1.37	0.170	-0.00951	-2.70	0.007
91	0.02623	3.42	0.001	0.04621	10.21	0.000	0.04836	12.71	0.000	0.04892	12.09	0.000
92	0.02096	2.78	0.006	0.04576	10.07	0.000	0.04955	13.86	0.000	0.05294	13.03	0.000
93	0.02287	2.94	0.003	0.03551	8.10	0.000	0.04290	13.60	0.000	0.04805	12.70	0.000
94	0.01423	1.84	0.066	0.03054	7.12	0.000	0.04391	14.56	0.000	0.04140	12.38	0.000
95	0.00378	0.47	0.640	0.02232	5.26	0.000	0.03466	13.54	0.000	0.03006	9.11	0.000
96	0.01127	1.46	0.146	0.01511	3.61	0.000	0.02615	10.31	0.000	0.02414	7.83	0.000
97	0.00128	0.16	0.872	0.01005	2.39	0.017	0.02009	8.78	0.000	0.01668	5.66	0.000
98	-0.00273	-0.32	0.749	0.00535	1.21	0.225	0.00645	2.90	0.004	0.00211	0.71	0.475
99	-0.00140	-0.16	0.872	0.00536	1.17	0.243	0.00694	3.26	0.001	0.00554	1.93	0.053

* in hundreds

Table 4.6. Estimated Implicit Prices (Evaluated at Sample Mean) With 95% Confidence Interval by Wine Class*

	Commercial			Semi Premium			Premium			Ultra Premium		
Variable	Lower 95%	Estimate	Upper 95%	Lower 95%	Estimate	Upper 95%	Lower 95%	Estimate	Upper 95%	Lower 95%	Estimate	Upper 95%
Cases	-$0.002	**-$0.001**	-$0.001	-$0.023	**-$0.020**	-$0.017	-$0.086	**-$0.075**	-$0.065	-$0.115	**-$0.097**	-$0.080
Score	$0.11	**$0.16**	$0.22	$0.60	**$0.68**	$0.77	2.09	**$2.32**	$2.55	$3.62	**$3.88**	$4.13
Age	$1.20	**$1.58**	$1.96	$1.66	**$2.03**	$2.39	2.54	**$3.21**	$3.88	$4.65	**$5.74**	$6.84
Napa	$3.09	**$4.38**	$5.67	$4.85	**$5.91**	$6.97	6.51	**$10.58**	$14.65	$18.72	**$25.71**	$32.69
Bay Area	-$0.38	**$0.90**	$2.18	$2.01	**$3.22**	$4.42	0.37	**$4.63**	$8.88	$13.59	**$21.17**	$28.76
Sonoma	$2.88	**$3.66**	$4.44	$3.37	**$4.29**	$5.21	0.41	**$4.44**	$8.46	$11.76	**$18.76**	$25.76
South Coast	$2.35	**$3.55**	$4.76	$2.80	**$3.83**	$4.85	-3.10	**$1.09**	$5.28	$6.30	**$13.48**	$20.66
Carneros	$2.04	**$5.55**	$9.06	$3.06	**$4.34**	$5.63	0.62	**$4.94**	$9.26	$8.85	**$16.56**	$24.27
Sierra	$1.73	**$2.59**	$3.45	$0.46	**$1.79**	$3.12	-3.86	**$3.34**	$10.54	$0.64	**$9.57**	$18.49
Mendocino	$1.69	**$2.62**	$3.56	$0.03	**$1.16**	$2.30	-4.48	**$0.14**	$4.76	$0.50	**$8.64**	$16.77
Washington	$2.31	**$3.16**	$4.01	-$0.29	**$0.68**	$1.66	-8.17	**-$4.03**	$0.11	-$5.38	**$2.15**	$9.68
Nonvarietal	-$0.77	**$0.95**	$2.67	$2.57	**$3.73**	$4.89	14.72	**$16.94**	$19.16	$23.52	**$26.99**	$30.45
Pinot	$0.17	**$1.00**	$1.83	$3.57	**$4.45**	$5.32	12.78	**$14.59**	$16.40	$19.78	**$22.94**	$26.10
Cabernet	-$0.78	**-$0.20**	$0.37	$2.25	**$2.99**	$3.74	10.88	**$12.60**	$14.32	$17.68	**$20.69**	$23.71
Merlot	$0.00	**$0.64**	$1.28	$2.75	**$3.46**	$4.18	7.92	**$9.66**	$11.40	$13.67	**$17.05**	$20.43
Shyrah	$0.95	**$2.15**	$3.34	$0.34	**$1.32**	$2.31	3.05	**$4.96**	$6.86	$6.55	**$10.47**	$14.38
Reserve	-$0.16	**$0.84**	$1.84	$1.67	**$2.48**	$3.29	2.16	**$3.28**	$4.39	$4.25	**$6.28**	$8.31
Vineyard	-$0.13	**$1.50**	$3.14	$2.17	**$2.89**	$3.62	3.19	**$4.15**	$5.11	$4.21	**$5.96**	$7.70
Estate	-$3.17	**-$0.52**	$2.14	-$0.32	**$1.66**	$3.64	-0.67	**$1.55**	$3.76	$1.51	**$5.52**	$9.52
91	-$3.36	**-$2.14**	-$0.91	-$9.06	**-$7.60**	-$6.14	-20.77	**-$17.99**	-$15.22	-$32.99	**-$28.39**	-$23.79
92	-$2.91	**-$1.71**	-$0.50	-$8.99	**-$7.53**	-$6.06	-21.05	**-$18.44**	-$15.83	-$35.34	**-$30.72**	-$26.10
93	-$3.10	**-$1.86**	-$0.62	-$7.26	**-$5.84**	-$4.43	-18.27	**-$15.97**	-$13.66	-$32.19	**-$27.88**	-$23.58
94	-$2.39	**-$1.16**	$0.08	-$6.41	**-$5.02**	-$3.64	-18.54	**-$16.34**	-$14.14	-$27.83	**-$24.03**	-$20.22
95	-$1.59	**-$0.31**	$0.98	-$5.04	**-$3.67**	-$2.30	-14.76	**-$12.90**	-$11.03	-$21.19	**-$17.44**	-$13.69
96	-$2.15	**-$0.92**	$0.31	-$3.84	**-$2.49**	-$1.14	-11.58	**-$9.73**	-$7.88	-$17.51	**-$14.01**	-$10.50
97	-$1.38	**-$0.10**	$1.17	-$3.01	**-$1.65**	-$0.30	-9.14	**-$7.48**	-$5.81	-$13.03	**-$9.68**	-$6.33
98	-$1.14	**$0.22**	$1.59	-$2.31	**-$0.88**	$0.55	-4.02	**-$2.40**	-$0.78	-$4.61	**-$1.23**	$2.16
99	-$1.28	**$0.11**	$1.51	-$2.36	**-$0.88**	$0.59	-4.14	**-$2.58**	-$1.03	-$6.48	**-$3.21**	$0.05

* If upper and lower bounds have concord signs, the implicit price is significant at the $\alpha = .05$ level.

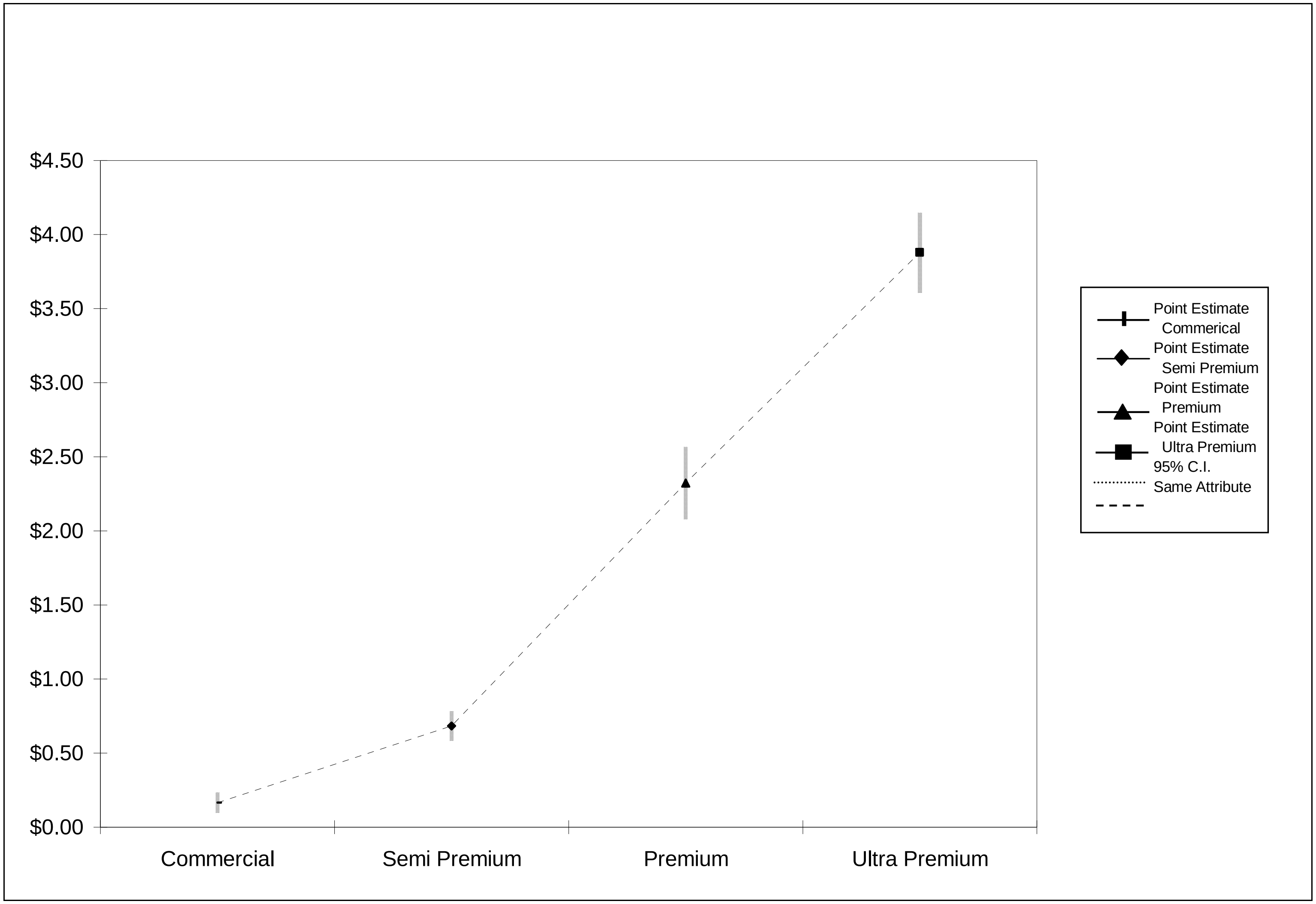

Figure 4.1. Implicit Price of Tasting Score by Wine Class

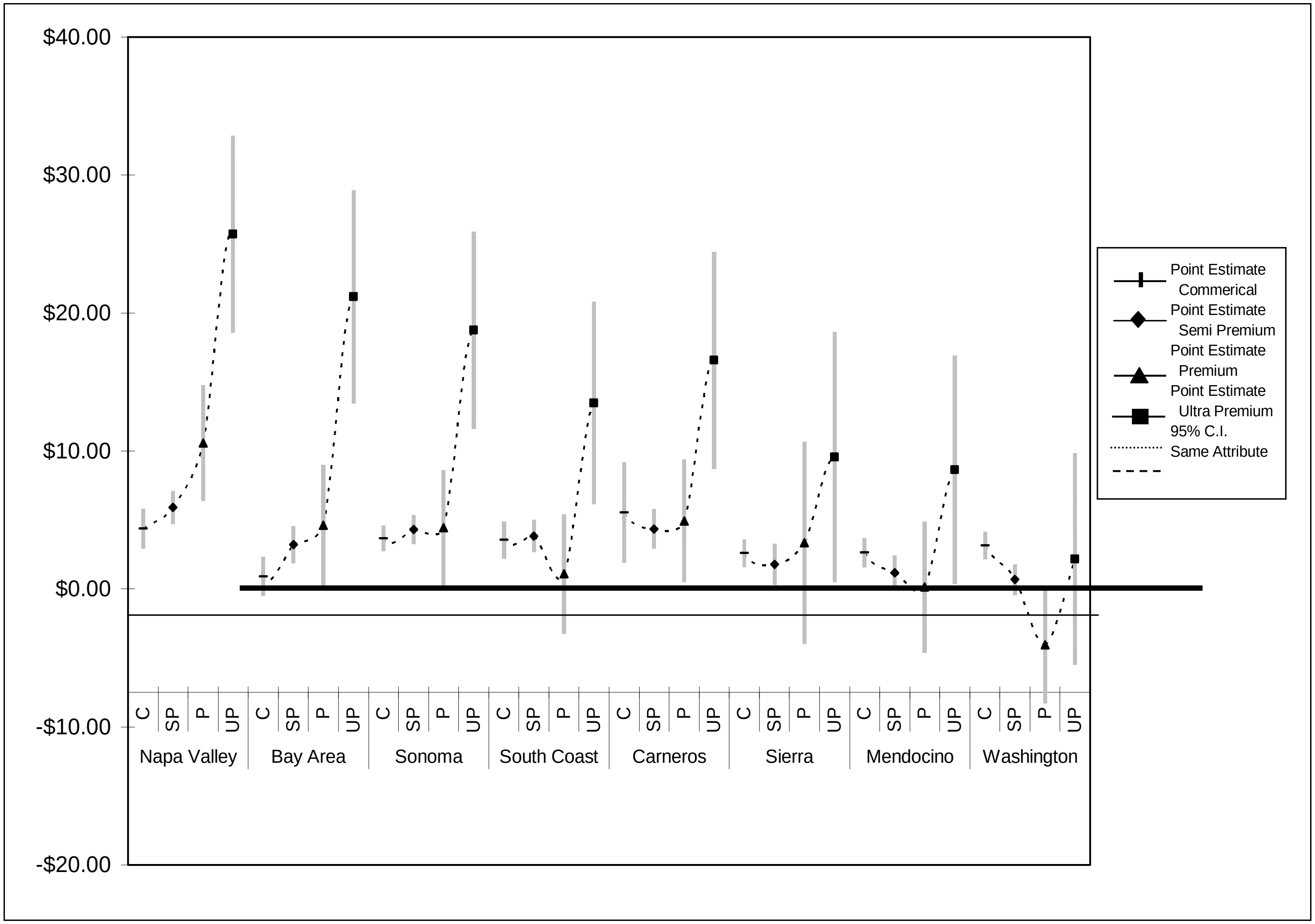

Figure 4.2. Implicit Price Premia of Regions of Production by Wine Class Relative to Generic California Wine