

Ithaca 2018 Abstract Submission

Title

Wine Hedonic Pricing using the Two Tier Stochastic Model

I want to submit an abstract for:

Conference Presentation

Corresponding Author

Loren Tauer

E-Mail

lwt1@cornell.edu

Affiliation

Cornell University

Co-Author/s

Name	E-Mail	Affiliation
Harold Fried	friedh@union.edu	Union College

Keywords

Hedonic wine pricing, two tier stochastic frontier, Riesling

Research Question

Determine whether a wine is overpriced or overpriced given the tasting score and characteristics of the wine.

Methods

The two tier stochastic frontier model is used to arrive at estimates of the expected price of U.S. Rieslings from 2000-2016 as well as estimates of overpricing and underpricing.

Results

The expected price with a tasting score of 90 is estimated to be \$19.03. The mean of underpricing is 8.82 percent and the mean of overpricing is 16.30 percent.

Abstract

Stochastic frontier models estimate technical or allocative efficiency. The method econometrically estimates a frontier function, using the specification of a two-component error structure, with one error component representing an inefficiency estimate, either below the frontier in the case of a production function, or above the frontier in the case of a cost function, and the second error component representing a normal random error term. The two tier stochastic frontier model has been specified with three error components, one error above the frontier, one error below the frontier, and a normal random error. Initial applications were to wage rates, with the estimation of a reduced form equation, with the positive component representing the bargaining power of the wage earners, and the negative component representing the power of firms. Other novel applications are beginning to appear.

We use the two tier model to estimate a hedonic wine price model, another novel application. Wine prices are regressed on wine attributes using a maximum likelihood estimator with three error components defined and estimated. A normal random symmetric error represents a random error. Two separate exponential distributions

are further defined and estimated, one above the estimated expected function, and one below the estimated expected function. The exponential distribution above the expected function represents overpricing of wine given the attributes used to determine the expected price of wine. The exponential distribution below the expected function represents underpricing of wine given the attributes used to determine the expected price of wine.

The wine we investigate is U.S. Rieslings with vintages between the years 2000 and 2016, with data collected from the Wine Spectator web site. We drop observations of 375 ml bottles, which mostly represents ice wines, as well as observations on 1.5 liter bottles (of which they were very few), resulting in 1951 observations. Wine prices were converted into 2016 prices using the CPI for food and beverages, and converted into natural log form for estimation, which allowed interpreting the errors as percentages. The dependent variable is the release price of the wine. The independent variable is the score ranking of the wine between 0 and 100, as determined by tastings conducted by Wine Spectator.

As stated by Wine Spectator, the price reported in their data is the release price of the wine and is set before the wine is rated. Thus, we determine not only the relationship between the expected price as a function of the tasting score, but also whether the release price is greater or less than the expected price given the tasting score using the two tier components. Other wine attributes can be added to the expected price model, including vintage year, the American Viticultural Area (AVA) or the U.S. state of production. Variables determining overpricing and underpricing can also be included in the exponential distribution terms to explain overpricing or underpricing.

The estimates by maximum likelihood produces the result: $\ln(\text{price}) = 0.3808 + 0.0285 \cdot \text{score}$, with the coefficient on score having a z value of 12.23, statistically different from zero at the probability level of 0.00. The constant term has a z value of 1.90, which is statically different from zero at the 0.05 level. The result is that the expected 2016 price of a U.S. Riesling with a tasting score of 90 should be \$19.03; a wine with a tasting score of 80 should be \$14.31, which may be viewed as high, but the only variable currently included in the expected price is the tasting score, and other attributes are undoubtedly important.

Both the above and below errors were estimated as exponential distributions, and their means where statistically different from zero. The mean (and standard deviation) of underpricing was 0.0882, and the mean (standard deviation) of overpricing was 0.1630. Thus for the population of wines investigated and the model estimated, it appears that as a group the wines appear to be more overpriced than underpriced. However, individual wines themselves were either overpriced or underpriced.

An estimate of the overpricing or underpricing of each individual wine can be derived from the estimates of the exponential functions estimated assuming the expected equation is deterministic, but because this equation was estimated with a residual normal distribution error term, an estimate using the normal random error as done in the technical efficiency literature, would be more appropriate and was used. The percentage overprice averaged 18.3 percent with a standard deviation of 13.68, and ranged from a low estimate of 6.43 percent to a high estimate of 192.32 percent. The percentage of underprice averaged 9.26 percent with a standard deviation of 3.09, with a low estimate of 5.89 to a high estimate of 31.56 percent.

Interpreting the price equation as random results in both an overprice and an underprice estimate for each wine. An example is the year 2000 "Blue Teal" wine from New Mexico (wine observation #1 in the data), that had an estimate of underpricing of 17.81 percent and an estimate of overpricing of 7.63 percent. That wine was released at a price of \$6 (\$8.82 in year 2016 prices) with a score of 82. Blue Teal was a low priced wine with a low rating but could have been more aggressively priced. In contrast, the year 2016 "Keuka Lake Vineyard Dry 20 Rows" from the Finger Lakes Region of New York (the last wine observation #1951), had an underprice estimate of 6.73 percent and an overprice estimate of 25.73 percent. The Keuka wine was priced at \$30 and had a high score of 92. Obviously a very good wine, but maybe overpriced.

The greatest underpriced wine is appropriately named, "For a Song", 2014, from Columbia Valley Caliche Lake Vinery with a release price of \$8 (\$8.17 in year 2016 prices), with a score of 90. This wine was estimated to be underpriced by 31.56 percent, but only overpriced by 6.43 percent. The greatest overpriced wine is "Navarro White Riesling Anderson Valley Late Harvest Cluster Select Sweet", 2006, from California, with a release price of \$59 (\$74.68 in year 2016 prices) and a score of 88. It was estimated to be overpriced by 192.32 percent but

underpriced by 5.89 percent. By its name, this wine appears to be a desert wine and might be placed into a different classification from the table Rieslings.

Further work will investigate whether vintage year or region may be factors in either determining the expected wine price or in underpricing or overpricing. We may also find that some wineries consistently overprice or underprice their wines, or that there may be learning where an underpriced or overpriced wine in any year may be followed with an adjustment in price for future vintages. A more ambitious exercise would be to quantify the tasting notes for each wine and include these variables into the hedonic regression.