Quality Estimation and Wine Industry Competitiveness

Angelo Zago
University of Verona - Italy

Davis, August 9, 2007

Work in progress
1. Motivation

Measuring and evaluating the quality attributes of raw materials, commodities, and intermediate products is a common problem in many sectors.

Food industries, wine and grapes, milk and cheese, fruits and juices, beets or canes and sugar, beans and coffee, etc..

Other industries, chips and computers, ores and steel, steel and construction works, crude and refined oil.
2. Objectives and plan of the paper

We propose a methodology to measure the characteristics and composition of intermediate products, i.e., grapes for wine production, and we pursue two objectives:

1. we address the issue of how to measure quality attributes for intermediate goods using a general representation of the technology, i.e., using directional distance functions.

2. we evaluate how quality attributes interact with quantity. Policy reasons, e.g., self-regulation in Europe, like appellation contrôlée in wine.
We propose an indicator based on **directional distance functions**.

We compare firms based on the distance from the frontier along a **pre-assigned direction** which reflects the preference and needs of the buyer or downstream firm with respect to the quality attributes.

It may be the case that to be valuable to a downstream firm, the composition of the raw material has to be close to an “**ideal**” bundle of attributes preferred by the buyer.
3. Literature Review

Hedonic pricing literature;

Fixler and Zieschang (1992): using radial distance functions, the quality adjusted indexes are the product of two indexes, a quality index and a CCD-type Tornqvist productivity index;

Färe et al. (1992): input-oriented Malmquist productivity change index as the geometric mean of two Malmquist indexes as defined by CCD computed using linear programming;

Färe, Grosskopf and Roos (1995): extends this productivity index by incorporating attributes together with ratios of distance functions to measure the service quality of each pharmacy. They decompose
the Malmquist productivity change index into three components, quality change, technical change and efficiency change;

Jaenicke and Lengnick (1999): merge the soil science literature on soil-quality indexes with the literature on efficiency and total factor productivity indexes, and isolate a theoretically preferred soil-quality index;

Färe et al. (1989) extend efficiency measurement when some outputs are undesirable; notion of weak disposability of outputs.
4. Methodology

Here directional distance function which allows to measure firms’ distance from the frontier moving along a pre-assigned direction.

We propose an indicator: an output aggregator expressed in difference forms, rather than in ratio forms like in the case of the Malmquist productivity index.

From the property of translation invariance in outputs, to contrast with the property of homogeneity of degree zero in outputs of the Malmquist index coming from the linear homogeneity of the output distance function à la Shephard.
Following Chambers, Chung, and Färe (1996, 1998), and Chambers (2002), we can define the directional technology distance function as:

\[
\overline{D}_T(x, y, s; g_x, g_y, g_s) = \max\{\beta \in \mathbb{R} : (x - \beta g_x, y - \beta, s + \beta g_s) \in T\},
\]

\[
g_x \in \mathbb{R}^N, g_y \in \mathbb{R}^M, g_s \in \mathbb{R}_+^N, (g_x, g_y, g_s) \neq (0^N, 0, 0^M),
\]

if \((x - \beta g_x, y + \beta g_y, s + \beta g_s) \in T\) for some \(\beta\) and \(dT(y, s, g_y, g_s) = \inf\{\delta \in \mathbb{R} : (y + \delta g_y, s + \delta g_s) \in \mathbb{R}_+^M\}\) otherwise.

Note that \((g_x, g_y, g_s)\) is a reference vector of inputs and outputs which determines the direction over which the distance function is determined. \(\overline{D}_T(x, y; g_x, g_y, g_s)\) represents the maximal translation of the input and output vector in the direction of \((g_x, g_y, g_s)\) that keeps the translated input and output vector inside \(T\).
As shown by Chambers, Chung, and Färe (1996), all known (radial) distance and directional distance functions can be depicted as special cases of the directional technology distance function.

Here, **directional quality distance function** with the following:

\[ \overline{D_Q}(x,y,s;0^N,0,g_s) = \max\{\beta \in \mathbb{R} : (x, y, s + \beta g_s) \in T\}, \]

\[ g_s \in \mathbb{R}_+^M, g_s \neq 0^M. \]

In this paper we are interested in constructing an **indicator** of quality attributes of the output. Its general purpose is that it can create a summary measure of inputs or outputs that can be used to evaluate how these aggregate quantities vary across firms (or time).
Adapting the indicators suggested by Chambers (2002), we can define the **1-technology Luenberger quality indicator** for \((x^1, y^1, s^1, s^0)\) by the following:

\[
Q^1(s^0, s^1, y^1, x^1) = \overrightarrow{D}_{Q}(x^1, y^1, s^0; 0^N, 0, g_s) - \overrightarrow{D}_{Q}(x^1, y^1, s^1; 0^N, 0, g_s). \tag{2}
\]

\(Q^1(s^0, s^1, y^1, x^1)\) represents the difference between the amount that it is possible to translate \(s^0\) and \(s^1\) into the direction \(g_s\) and still keep both quality bundles in the output set of firm 1, i.e., we are referring to firm’s 1 technology or input-output bundle \((x^1, y^1)\).

See figure 1.
Figure 1. Directional distance function
The **0-technology Luenberger quality indicator** for \((x^0, y^0, s^1, s^0)\) is defined by the following:

\[
Q^0(s^0, s^1, y^0, x^0) = \overline{D}_Q(x^0, y^0, s^0; 0^N, 0, g_s) - \overline{D}_Q(x^0, y^0, s^1; 0^N, 0, g_s). \tag{3}
\]

Here we are computing the indicator from firm 0’s perspective, since we consider its input-output bundle \((x^0, y^0)\). If \(Q^0(s^0, s^1, y^0, x^0) > 0\), the quality is higher for firm 1 than firm 0, using as a reference firm 0’s technology or input-output bundle \((x^0, y^0)\).

The choice of the technology can affect the results. Thus the **Luenberger quality indicator** is the average of \(Q^1(s^0, s^1, y^1, x^1)\) and \(Q^0(s^0, s^1, y^0, x^0)\):

\[
Q(s^0, s^1, y^0, y^1, x^0, x^1) = \frac{1}{2}(Q^1(s^0, s^1, y^1, x^1) + Q^0(s^0, s^1, y^0, x^0)). \tag{4}
\]
5. Activity analysis and empirical implementation
We use DEA to envelop the technology associated with a cross-section sample of firms. Referring to a VRS technology, the linear program problem to solve to compute the **directional quality distance function** for each observation \( k' \), is the following:

\[
D_O(x', y', s'; 0^N, g_y, g_s) = \max \beta: \quad \sum_{k=1}^{K} z_k y'_k \geq y'_k, \quad (5)
\]

\[
\sum_{k=1}^{K} z_k s'_{km} \geq s'_{km} + \beta g_s, \quad m = 1, ..., M,
\]

\[
\sum_{k=1}^{K} z_k x'_{kn} \leq x'_{kn}, \quad n = 1, ..., N,
\]

\[
\sum_{k=1}^{K} z_k = 1, \quad k = 1, ..., K,
\]
where $g_y$ and $g_s$ are the direction vectors for output and quality attributes respectively.

1. The quality indicators

To compute the quality indicator proposed, we need to use and compute four different quality directional distance functions.

The direction vector $g_s$ has to be specified. One possibility is to consider the average attributes content of the grapes for the whole sample of firms, i.e., $g_s = \bar{s}_m$, where $\bar{s}_m = \sum_{k=1}^{K} \frac{s_{km}}{K}$ and $m = 1,..,M$.

Another choice is the ideal composition of the good.
The choice of the **reference observation (the base)**, to have the 0-technology, allows for different options. The *average* of the observations defined by:

\[
\begin{align*}
    s^0 &= \sum_{k=1}^{K} \frac{S_{km}}{K}, \quad m = 1, \ldots, M, \\
    x^0 &= \sum_{k=1}^{K} \frac{x_{kn}}{K}, \quad n = 1, \ldots, N, \\
    y^0 &= \sum_{k=1}^{K} \frac{y_k}{K}.
\end{align*}
\]

Possible drawback, it may lead to an unrealistic artificial technology, i.e., to a not feasible input/output combination.

Another possibility could be the *minimum quality* composition required by the law or by industry standards, the one that all firms should provide as a minimum requirement.
Once we obtain the results, we will look at the **distribution of the indicators** smoothed using standard normal kernel function and optimal bandwidth:

\[
    f(x) = \frac{1}{Kh} \sum_{j=1}^{K} k\left(\frac{x_j - x}{h}\right).
\]  

(7)

A more formal statistical analysis based on the kernels is the Li’s **nonparametric testing of closeness** between two unknown distribution functions. It tests the null that two distributions, e.g., \(f(x)\) and \(g(x)\), are identical, i.e., \(H_0 : f(x) = g(x)\) for all \(x\), against the alternative that they are different, that is \(H_1 : f(x) \neq g(x)\) for all \(x\).

The Li’s T-statistic is asymptotically distributed as a standard normal with a critical value of 2.33 (1% significance level).
2. The quality-quantity trade off

To evaluate the trade-off between output quantity and aggregate quality, it is a natural choice to look at the relationship between the quality indicators and the yields.

We consider the different options used for the direction vector $g_s$, and show the relationship via a graphical representation using lowess smoothing, a scatterplot smoothing technique based on a locally weighted regression of a $y$ variable - the quality indicator in our case - on an $x$ variable, i.e., the yields.

In lowess, the regression is weighted to give the highest weight to the point and less to those that are farther away. Lowess thus tends to follow the data, because of its locality.
6. The data

We use data provided by the “Istituto Agrario di San Michele all’Adige”, located near Trento, in the Northern Italian Alps.

2 varieties: Chardonnay and Merlot.

On the input side, we have the following
- altimetry,
- number of vines per hectare,
- the depth of the roots, a categorical variable (1-3),
- the water reservoir, in the range 1-4,
- total calcium, 1-5,
- skeleton, a categorical variable (1-4),
- internal drainage, a categorical variable (1-5),
- external drainage, a categorical variable (1-3).

Then variables more “in the control of the producers”, such as:
- the number of buds per branch, a result of the pruning intensity,
- irrigation, a dummy for the presence of irrigation,
- cultivated, a dummy for the presence of grass.

For the grapes obtained in the different fields, we have information on
- production per hectare,
- sugar content (measured in degree Brix),
- tartaric acid,
- malic acid,
- potassium,
- pH,
- total acidity.
<table>
<thead>
<tr>
<th>Variable</th>
<th>Unit of measure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Altimetry</td>
<td>mt.</td>
</tr>
<tr>
<td>Vines per hectare</td>
<td>no.</td>
</tr>
<tr>
<td>Buds per branch</td>
<td>no.</td>
</tr>
<tr>
<td>Water holding capacity</td>
<td>1-4</td>
</tr>
<tr>
<td>Total calcium</td>
<td>1-5</td>
</tr>
<tr>
<td>Grapes production per ha</td>
<td>0.1 t./ha</td>
</tr>
<tr>
<td>Sugar content</td>
<td>°Brix</td>
</tr>
<tr>
<td>Total acidity</td>
<td>gr./l.</td>
</tr>
<tr>
<td>pH</td>
<td>1-14</td>
</tr>
<tr>
<td>Tartaric acidity</td>
<td>gr./l.</td>
</tr>
<tr>
<td>Malic acidity</td>
<td>gr./l.</td>
</tr>
<tr>
<td>Potassium content</td>
<td>gr./l.</td>
</tr>
</tbody>
</table>
7. Results

Table 4. Luenberger Quality Indicators

<table>
<thead>
<tr>
<th></th>
<th>No. obs</th>
<th>Mean</th>
<th>St. dev</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Chardonnay</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Average</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1994</td>
<td>214</td>
<td>0.004</td>
<td>0.056</td>
<td>-0.137</td>
<td>0.349</td>
</tr>
<tr>
<td>1995</td>
<td>187</td>
<td>0.005</td>
<td>0.044</td>
<td>-0.216</td>
<td>0.130</td>
</tr>
<tr>
<td>1996</td>
<td>213</td>
<td>0.004</td>
<td>0.047</td>
<td>-0.173</td>
<td>0.164</td>
</tr>
<tr>
<td><strong>Ideal</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1994</td>
<td>214</td>
<td>0.006</td>
<td>0.091</td>
<td>-0.292</td>
<td>0.636</td>
</tr>
<tr>
<td>1995</td>
<td>187</td>
<td>0.000</td>
<td>0.085</td>
<td>-0.559</td>
<td>0.217</td>
</tr>
<tr>
<td>1996</td>
<td>213</td>
<td>-0.003</td>
<td>0.079</td>
<td>-0.503</td>
<td>0.238</td>
</tr>
<tr>
<td><strong>Merlot</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Average</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1994</td>
<td>78</td>
<td>-0.011</td>
<td>0.082</td>
<td>-0.188</td>
<td>0.286</td>
</tr>
<tr>
<td>1995</td>
<td>127</td>
<td>0.008</td>
<td>0.037</td>
<td>-0.145</td>
<td>0.151</td>
</tr>
<tr>
<td>1996</td>
<td>120</td>
<td>0.014</td>
<td>0.058</td>
<td>-0.083</td>
<td>0.451</td>
</tr>
<tr>
<td><strong>Ideal</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1994</td>
<td>78</td>
<td>-0.018</td>
<td>0.181</td>
<td>-0.517</td>
<td>0.739</td>
</tr>
<tr>
<td>1995</td>
<td>127</td>
<td>0.005</td>
<td>0.088</td>
<td>-0.581</td>
<td>0.494</td>
</tr>
<tr>
<td>1996</td>
<td>120</td>
<td>0.011</td>
<td>0.108</td>
<td>-0.306</td>
<td>0.869</td>
</tr>
</tbody>
</table>
Distributions of Quality Indicators – Chardonnay
Distributions of Quality Indicators – Merlot
<table>
<thead>
<tr>
<th></th>
<th>CRS</th>
<th>Ideal vs</th>
<th>VRS</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Chardonnay</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1994</td>
<td>0.0319</td>
<td>0.1382</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1995</td>
<td>0.0068</td>
<td>0.0783</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1996</td>
<td>0.0018</td>
<td>0.0242</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>All</strong></td>
<td>0.0325</td>
<td>0.1424</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Merlot</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1994</td>
<td>--</td>
<td>0.2653</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1995</td>
<td>0.0058</td>
<td>0.0661</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1996</td>
<td>--</td>
<td>0.0167</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>All</strong></td>
<td>0.0735</td>
<td>0.2170</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Tradeoff of Yields versus Quality Indicators - Chardonnay
Tradeoff of Yields versus Quality Indicators - Merlot
Tradeoff of Yields versus Ideal Quality Indicator - By Years
Comparison of Tradeoffs: Sugar and Ideal QI versus Yields
8. Concluding Remarks

Quality is important \( \rightarrow \) need to be able to understand what market wants.

Here 2 different measures of aggregate quality, to be compared with the standard practice of using only sugar content.

They seem similar, but they show different results when compared with respect to the yields.

How to use a more limited set of variables? How to give incentives to producers? Calculate the shadow prices of yields and quality attributes.

How to take equity issues into consideration?