"Actual" and Perceptual Effects of Category Spanning

Giacomo Negro
Goizueta Business School, Emory University, Atlanta, Georgia 30322, giacomo_negro@bus.emory.edu

Ming D. Leung
Haas School of Business, University of California, Berkeley, Berkeley, California 94720, mingdleung@haas.berkeley.edu

Literature to date has demonstrated that producers and products spanning multiple categories have inferior market performance. However, two related but distinct explanations exist as to the source of such a discount. One explanation suggests that "actual" skills are degraded when producers attempt to engage across diverse categories. Another explanation involves perceptual fit to category representations held by an audience as the cause. These two explanations tend to be confounded in archival studies because external observers, responsible for the evaluation of market performance, are often aware of both the identity of producers and the underlying characteristics of their products. This leaves researchers unable to empirically separate effects. We present an analysis conducted in a setting in which it was possible to distinguish the two mechanisms: critics’ ratings of the same wines through “blind” and “nonblind” tastings. The findings indicate that after controlling for the value of ratings assigned blindly, the wines made by wineries spanning styles continue to receive lower ratings in the nonblind situation.

Key words: organization theory; organizational ecology; market categories
History: Published online in Articles in Advance September 10, 2012.

1. Introduction
A growing number of studies in economic sociology and organization theory demonstrate how social categories affect individual and firm outcomes in a variety of market situations (Zuckerman 1999, Hannan et al. 2007, Hannan 2010). As collective construals that partition individuals, products, and organizations, categories have been shown to influence diverse phenomena including the divestment of incoherent lines of business (Zuckerman 2000), the coherence of stock price movement (Zuckerman, 2004), critical appeal (Hsu 2006, Negro et al. 2010a, Kovács and Hannan 2010), the change of production methods (Negro et al. 2011), the assessment of producers’ legitimacy and value (Rosa et al. 1999, Rao et al. 2005), creditworthiness of businesses and the equity of their ratings (Fleischer 2009, Ruef and Patterson 2009), the competitive positioning of producers (Porac et al. 1995, Waguespack and Sorenson 2011), managerial compensation (Porac et al. 1999), and vital rates of organizational populations (Bogaert et al. 2010, Carroll et al. 2010).

The typical finding to date suggests that producers focusing their efforts on one or few market categories perform better than those who attempt to span multiple ones (Hannan 2010). Two distinct arguments are offered for this observed regularity—a so-called producer-side mechanism and an audience-side mechanism (Hsu et al. 2009). The producer-side argument suggests that there are operational difficulties that producers face when attempting to span disparate categories. Because the organization is targeting several categorical niches, it risks offering products of lower intrinsic quality across any one (Freeman and Hannan 1983, Dobrev et al. 2001). The audience-side argument suggests that the discount occurs because the identities of category spanners are unclear, and therefore such products and producers are perceived as more difficult to situate, understand, and/or evaluate (Zuckerman 1999, Rao et al. 2005, Hsu 2006).

A challenge remains in disentangling these two processes because although conceptually distinct, they are related and empirically co-occurring. For example, specialized critics are involved in assignments of genre labels to films, but studios and distributors claim membership of projects in categories by using genres in various stages of the organizing process, from researching spectator preferences, to hiring technical crews and artists whose work is limited to certain types of films, to promoting films to exhibitors. Accordingly, the category-spanning discount can combine effects driven by audience and producer mechanisms simultaneously. The evaluation of underlying content is confounded by provenance. This study contributes to the literature on market categories by producing evidence that perceptual, and not merely actual, effects underlie the observed link between category specialization and greater market success.

We examine whether and to what extent the audience-side mechanism linked to category-spanning affects evaluation. Our main argument is that when the identity of the producer is directly observable, the audience can
connect the membership of the producer in a category to the collective representation of that category, or schema. Given this, we expect that increasing nonconformity by a producer, as realized by category spanning, will result in worse fit to this recognized schema. Its products should therefore be perceived as less appealing.

Some studies have been able to empirically separate mechanisms that similarly co-occur by leveraging the research design advantages of certain settings. Waguespack and Sorenson (2011) were able to demonstrate how the status of a film studio—in particular, its powerful position within the industry—affected parental guidance classification of films. They show that despite underlying similarities in potentially objectionable content, films from major Hollywood studios receive more lenient ratings than films from independent production companies. Roberts et al. (2011) isolated the symbolic effect of a winemaker’s prominence on the prices of wines net of their actual contribution to the quality of the wine by exploiting the timing of wine releases. Hiring winemakers from prominent wineries increases the retail price of wines, even of those that were harvested and aging before the winemaker joined the organization.

To test our arguments, we investigate a setting that permits us to separate the audience-side from producer-side mechanisms: critical ratings of wines by two critics, both of which are highly respected but differ in their approach to how they rate their wines. One conducts blind tastings; i.e., all information regarding the producer is hidden during the tasting. The other does not. More specifically, we utilize ratings assigned through blind tastings as an assessment of the underlying characteristics of the wines and examine whether, after accounting for these ratings, the spanning of winemaking styles by producers affects the ratings of their wines assigned through nonblind tastings. Our findings indicate that, controlling for the underlying quality characteristics of their wines, wineries increasingly spanning winemaking styles receive lower ratings. In short, the rating discount from multiple category membership remains significant.

The rest of this paper is organized as follows. The second section reviews the literature on multiple categorical memberships and the category-spanning disadvantage. We identify two hypotheses reflecting our focus on two separate mechanisms. The empirical setting is then detailed in §3, along with attention to describing the two wine critics and their different approaches to tasting and rating wines. We then present results from our empirical analysis in §4 and conclude with remarks discussing the contribution of the study in §5.

2. Market Categories, Producer Identity, and Category Spanning
Categories are sociocognitive partitions that can be applied in any social domain including markets, grouping objects considered similar along certain dimensions (Zerubavel 1997). For example, films are clustered into genres such as “Western” or “horror,” food styles are clustered into cuisines such as “sushi” or “contemporary,” and stocks are rated into groups such as “market outperformer” or “market underperformer.” Categories, in this sense, reflect shared understandings between market participants as to how similar producers and/or products can be clustered together and labeled.

Note that not all objects fall neatly into a single category, and categories do not always have clearly defined boundaries. For example, some films are labeled horror and sci-fi because they combine plotlines and characters typical of both genres. Similarly, certain music styles, such as “ambient,” do not have a clearly demarcated form; works using either acoustic or electronic instruments, or no instruments at all, can be labeled as ambient. In this sense, category memberships are partial, and categories are said to be fuzzy.

Recently, a number of studies demonstrated that social objects that combine features from multiple categories are discounted; that is, they experience lower market performance and are deemed as less appealing by an audience (for a review, see Hannan 2010, Negro et al. 2010b). In general, these studies adopt the view that markets are organized as interfaces in which (i) producers present offerings to appeal to (ii) an audience that evaluates them (Zuckerman 1999, Hannan et al. 2007). As sketched above, these are the two distinct but related arguments that explain the source of the market discount from multiple category membership—an audience-side and a producer-side mechanism.

The first argument relates to the perceptions that audiences hold of the features of a market agent and its offerings relative to sets of expectations. Zuckerman (1999) identified how stock analysts were confused by companies that straddle established industry classes. Stock analysts specialize by industry, and their expertise in evaluating companies is limited to these industry groups. Companies whose operations cross industry boundaries extend beyond what the more specialized analysts understand and are more difficult to evaluate. They are more likely to be ignored by these analysts. If evaluated, however, producers spanning categories are directly judged as less appealing. In the study by Rao et al. (2005) of French haute cuisine, traditional and nouvelle cuisines are seen as having incompatible identities, and the chefs who combine dishes from the different styles in their menus receive lower ratings by the influential Michelin Guide.

Hannan et al. (2007) and Hsu et al. (2009) formalized this argument in terms of niche theory’s principle of allocation (see also Freeman and Hannan 1983, Dobrev et al. 2001). A niche defines the social space in which a producer can persist and thrive or, in the language of the theory, has positive fitness. The niche space
is fixed, so occupying a wider niche and, by implication, spanning multiple categories in a diverse market, reduces an actor’s fitness in any one category. Spanning categories is linked to poor correspondence between the features of the producer and the representation of the categories, or schemas, in the sense that the producer matches some feature values of one schema and others of other schemas. In a study of the U.S. brewing industry, Carroll and Swaminathan (2000) outlined key features that define the schema for “microbrewers,” including small-scale production, handcrafted methods, and natural ingredients. Engaged beer consumers often reject beverages made by contract brewers, which miss some of these key features and display features typical of the production system of mass brewers, despite the fact that they could produce beer of equal quality to that of the microbrewers.

As a result of poor correspondence to schemas, a candidate spanning multiple categories is regarded as less appealing. This prediction is consistent with research in cognitive psychology and consumer behavior suggesting that congruity of an object to a schema leads to favorable responses because, other things being equal, people like objects that conform to their expectations, allow predictability, and require less extensive cognitive elaboration (Rosch 1978, Garbarino and Edell 1997). The brewery example points to another important aspect of the theory. In markets, the audience can often evaluate the fit to a schema when it can connect the identity of the producer to visible features. In other words, a beer enthusiast expresses dislike for contract breweries and their products when their production methods and ingredients are visible to them. When such visibility is absent, or limited because of successful deception (Phillips and Kim 2009), the loss of appeal is more difficult to observe. This argument does not imply an actual lack of skill or ability on the part of the candidate but rather points to risks inherent in not adhering to a recognized schema-based identity.

A second argument has been proposed for the existence of a market discount from multiple category membership that implicates skill development or learning—that there is an actual functional reason for the discount. This discount is hypothesized to originate from the producer side. Hannan et al. (2007) and Hsu et al. (2009) argued that producers have limited attention, time, and material resources to devote to maintaining their fitness with a particular niche—what they term engagement. These producers, or candidates, have to choose how to allocate their limited engagement across multiple categorical niches. Spreading one’s effort, or engagement, across several categories means that one will be less well prepared to address the needs of any single one. As Hsu et al. (2009) demonstrated, those eBay sellers who listed items across several different product categories were less facile at engaging with their audience. Specifically, these generalist sellers were less likely to use category-specific acronyms in their listings, thereby demonstrating their reduced expertise in any particular category.

Empirically, the difficulty in teasing the “actual” versus “perceptual” mechanisms apart lies with the fact that the audience is often aware of the underlying characteristics as well as the provenance of the producer and product. It therefore becomes challenging to determine whether audience-side or producer-side effects drive market evaluation. On one hand, the superior underlying features can be the reason for a favorable judgment. On the other hand, the knowledge of the identity of the producer can sway the audience’s opinions as well. Waguespack and Sorenson (2011) highlighted this issue in their study of how U.S. films are rated by the Motion Picture Association of America (MPAA) industry association and showed that, controlling for the underlying content of a movie, the identity of the movie studio affects the leniency with which a film is rated by the MPAA as suitable for general versus adult viewers. They find that the major, and more powerful, movie studios rated favorably in the past are more likely to receive more lenient ratings than independent studios that had more restrictive ratings in the past. Another example is the analysis by Roberts et al. (2011), which identified the effect of hiring a prominent winemaker on a winery’s wine prices. The mechanism that leads to observations of increased prices is confounded here as well because a “better” winemaker should improve the quality of wines as well as the prominence of those wines. Roberts et al. (2011) disentangled these effects by demonstrating that wines that were already aging, and therefore untouched by the more prominent winemaker upon their being hired, also realized an increase in price, a “symbolic” effect of their appointment.

One solution to address this challenge is to identify a setting where the information regarding the identity of the producer whose product is being evaluated can be obscured from the audience (see Simcoe and Waguespack 2011 for an application to organizational status). For example, academic journals often utilize a blind review process where the author is supposedly unknown to the reviewer. One reason for using this method is to ensure that the underlying quality of the manuscript, and not confounding factors such as reputation or identity of the author, matters for the decision to publish or reject a manuscript. Similarly, symphonies can use “blind” auditions, where a screen obscures the identity of the musician (Goldin and Rouse 2000), thereby reducing the likelihood of discrimination based on the gender of the performer. In this vein, we identified a research setting where critical ratings of the same products were conducted by two different evaluators—one with knowledge of the producer’s identity and the other without.
If a producer’s identity is unknown to the audience, then evaluations made by the audience cannot take this factor into account. The evaluation should therefore reflect the underlying characteristics of the product as determined by the abilities of the producer. Accordingly, our expectation is that those producers who span multiple categories will be disadvantaged because their products will be of actual lower quality. Therefore, we predict the following in our empirical setting of critical ratings of Italian wines.

**Hypothesis 1.** *Wines made by producers spanning across winemaking styles receive lower ratings in blind tastings.*

On the other hand, if the producer’s identity is known, the audience can connect the identity to the product. In this instance, perception of the fit to the category schema will also affect the evaluation. Our contention is that spanning categories will lead to lower appeal from this alone. However, this evaluation is confounded with the fact that it can still reflect the underlying quality characteristics of the products. To determine whether mere knowledge that a producer is spanning categories can lead to poorer critical ratings, we need to partial out the value of the underlying characteristics of the product under consideration. In doing so, we would predict that the remaining effect of category spanning is attributable to perceptual factors. In our empirical setting, critics’ ratings of Italian wines, the argument above leads to the following hypothesis.

**Hypothesis 2.** *Wines made by producers spanning across winemaking styles receive lower ratings in non-blind tastings, even after controlling for ratings assigned blindly.*

### 3. Methods

#### 3.1. Styles of Making Barolo and Barbaresco Wines

The empirical details contained in this section draw from fieldwork conducted between 2005 and 2009, including visits to 45 winemakers, wine critics, and oenologists in Italy. Barolo and Barbaresco wines come from the district called the Langhe, located in the southeastern corner of the Piemonte region, a short distance from the French and Swiss borders. Most wineries are family businesses and very small in scale by worldwide standards. Many of the better-known producers in this region, such Cantina Bartolo Mascarello, produce approximately 30,000 bottles (about 2,500 cases) per year. In comparison, New York-based Constellation Brands sells approximately 75 million cases annually.

The production of these wines is subject to a written legal code, the disciplinare di produzione. It states that these wines can be produced from only one grape varietal—Nebbiolo—grown in specifically delimited areas. Barolo must come from vineyards in 11 villages in the Cuneo province and must be aged for three years, or five years if designated as a riserva (reserve). Barbaresco comes from vineyards in four villages in an adjoining southern district in the same province, and it requires only two years of aging, or four years if designated as a riserva. (For the purpose of our analysis, Barolos and Barbarescos do not differ in other meaningful ways.) Of note is the fact that the disciplinare is silent on several other choices regarding the vinification process, including the kind of barrels that producers can use for the aging of their Barolos and Barbarescos.

Events beginning in the 1980s led to the partitioning of winemaking into modernist, traditional, and middle styles, resulting in strong identity claims (Negro et al. 2010a, 2011). Traditionally, casks as large as 120 hectoliters (hl) made from Slovenian oak or chestnut wood (botti) were used to age Barolo and Barbaresco wines. During the 1980s, a group of wineries came to question the validity of established methods. They maintained that the aromas and volatile acidity of traditional wines often lead to a flawed product. Lack of color, strong tannins, and off-putting aromas were frequent critiques. These wineries, categorized as modernists or international (as we follow here), introduced various changes, including reduced yields in the vineyard and a shorter maceration period.

The pivotal element, however, was the vessel used for the aging. Producers began aging wine in barriques, small barrels (usually 2.25 hl) made from aromatic French oak. Such producers aimed at wines with greater concentration, softer tannins, and darker color that were more in line with the taste of international markets. Angelo Gaja is regarded as the initiator of the use of barriques for these wines. His lead was followed by others, notably Elio Altare, who is also known for an emblematic episode: after failing to convince his father to allow production in barriques, he destroyed his family’s botte with a chainsaw in 1982. For many, this act symbolized the intensity of the controversy that erupted in the Langhe.

Modernism sparked a defense around regional typicality that sought to preserve traditional practices, notably the use of botti, as the essence of authenticity. Many of the traditional wineries we interviewed motivated their resistance to change in technical and cultural terms. Some said to us that they like their wine to “taste natural, without too much wood flavor” but also that they decided to remain close to conventional practices. One of them said, “We are deeply rooted in this region and in the simple life. We search for the authenticity of things, and live in the simplicity of the land. Would you believe that we never had a TV? We like living this way.”
Our fieldwork also suggests that greater diversity in methods and styles persists despite the opposition between tradition and modernity. The prevailing view among producers, critics, and enthusiasts acknowledges the existence of a distinct “middle” style in which wineries use both barriques and larger botti casks (typically in this sequence) to age a wine. A coeditor of the influential wine guide edited by Gambero Rosso said to us that the debate in an “either/or” proposition that “you use either barriques or casks” is “somewhat naïve; I think you can in fact use both.” The audience recognizes three styles, but some debate remains as to whether barrique-only and barrique-and-botti wines are just variations of a single international style. To make our test simpler, we focus on the categories where the contraposition is clearer. Accordingly, we identify two “pure” categorical wine styles: (i) botti only and (ii) barrique only.

Although the choices about maceration, yeasts, and temperature control also affect taste, these features also tend to be obscured; barriques have rather detectable effects on the wine and are identifiable. Producers, critics, and consumers consider the type of barrels used for aging as the key feature that distinguishes membership in a wine style. The aging methods thus serve what Hannan et al. (2007, p. 82) defined as a minimal test code for categorization. In other words, audience members use some critical features to make inferences about the fit to the rest of the schema on which they have difficulty observing a priori. If the audience believes that a producer “passes” the test, then it is assumed that the unperceived/nondefault values of schema-relevant features also fit the schema (Rosch and Mervis 1975, Rosch 1978).

3.2. Approaches to Wine Tasting

Our analyses use ratings and other information on wineries and wines from two specialized publications: I Vini di Veronelli, known as the Veronelli guide, and I Vini d’Italia, which is published by Gambero Rosso Editore and the Slow Food organization and known as the Gambero Rosso guide. These guides offer wide and deep coverage of Italian wineries and labels, and they are considered reliable by previous studies (Odorici and Corrado 2004; Negro et al. 2010a, 2011). The Veronelli guide reviews more than 10,000 wines by 2,000-plus wineries every year and has been published since 1991. The guide was founded and led by Luigi Veronelli, who was widely considered the most influential Italian wine intellectual and critic from the 1950s to his death in 2004. Analogously to the Veronelli guide, the Gambero Rosso guide covers a vast selection of the country’s labels, including over 2,200 wineries and 16,000 wines in recent editions. Overall, these two guides cover more wines and producers than competing domestic publications or international sources such as the Wine Advocate.

The two guides target wine connoisseurs as well as nonexperts including wine merchants, restaurant sommeliers, and consumers. The publications have earned audience respect—they are deemed the most prominent in Italy—and the list of the wines that receive their highest honors is routinely covered by the national press and is employed by other experts and commentators for their recommendations. The entries of the guide are organized by locality and producer.

Ratings from these guides matter to the commercial success of wines. In an interview with the authors, one of the (now former) editors of the Gambero Rosso guide said, “After they receive the highest ratings, some producers raise their prices a lot. . . . When the list is official, former ‘friends’ who didn’t win are not friends anymore. This gives me a great responsibility because I know that my evaluations will affect these producers and their businesses.” To determine whether ratings significantly affect the consumer market, we gathered data on the online prices of Barolo and Barbaresco labels using the website http://www.wine-searcher.com, which has retail price listings from over 7,800 merchants of wines from around the world (the prices mentioned in the guides are suggested by the producer and not actual market prices). In the appendix, we present regression estimates suggesting significant effects of ratings by these Italian critics on international retail wine prices. For example, receiving the highest accolade from Veronelli increases prices by about 25%. In this sense, the setting is not unlike many other cultural industries where domestic ratings affect the international performance of products. In film, for example, U.S. movies generate most of their revenues outside the domestic market, but U.S. critics such as Rogert Ebert or Variety influence reception internationally.

Both guides use ratings to communicate assessments of labeled wines. However, they differ in their approach to how wines are tasted and rated. The star ratings included in the Veronelli guide are supplied for individual wines after one or multiple tastings conducted by Luigi Veronelli and his three collaborators (all of whom have continued to edit the guide after Veronelli’s death). One of the guide editors, who has done most of the tastings of Barolos and Barbarescos for Veronelli, explained that he has information, albeit incomplete, about the wine—the critic at least knows the producer’s identity and other information that typically includes wine type, the area of production, and the vintage. When we interviewed him, he said that the tastings are conducted during individual visits to the wineries, often from the barrel.

In addition to the ratings, the Veronelli guide reports further details about each wine including the type of method used for the aging. Demonstrating the extent to which the barrel type has been used to demarcate category boundaries, symbols appear in the entry for each
wine to indicate whether the wine had been aged in botti or barriques (or both). The information is provided to the editors by the producers. We confirmed the coding of the data culled from the guide by conducting a telephone survey in 2006 of all producers in the data set to confirm their choices of vinification methods for each of their wines.

On the other hand, for the Gambero Rosso guide, the tastings are conducted by wine tasters divided into regional groups, with different wine labels sampled uniformly by type and vintage. Tasters do not know the price category of the wine or the identity of the winery. In a meeting in late 2005, one of the editors of the Gambero Rosso guide explained to us their approach: “[T]he guide is made with the contribution of more than 120–130 tasters who make the selections in every geographical area... activities, primarily tasting, are coordinated by commissions that are responsible for bringing in the samples from a region or a relevant winemaking area. We use blind tasting. . . Our tastings are collective, and so the decision is collegial. When we decide the Tre Bicchieri (the highest rating), we are in a group of 8–10 tasters.” The collectivist nature of these tastings also suggests that the ratings are unlikely to reflect the idiosyncratic tastes of a particular critic.

Interviews with critics reveal that both of the guides employ schemas for the aging styles. This is important, because we want to ensure the evaluations of a wine’s underlying characteristics made by the two critics are based on some criteria (the criteria can still differ, as discussed below). An essential difference with the Gambero Rosso guide is that tasters do not know the identity of the winery whose wine is being rated. However, critics of both guides approach the tastings with the style types in mind. In the introduction to the 2006 edition, the Gambero Rosso editors stated that the tastings are “comparative” in the sense that a wine is judged against other wines from the same zone and vintage. The editors note that their judgment criteria include “factors that cannot be evaluated in merely sensory terms, concepts like a wine’s conformity with or representativeness of its type. . . Wine should be good but should also be representative.” This corroborates previous research suggesting that both Gambero Rosso and Veronelli ratings are generally positively correlated (Negro et al. 2010a).

The coeditor of the 2006 Veronelli guide, Daniel Thomases, also noted that the primary thing he searches for in a wine is the correspondence between the territory and the vintage. Then, two main criteria are applied: “One is the nose, the smell. The second feature is the body, the structure.” This source, however, goes further in detailing the beliefs that accompany the evaluation process. Thomases said, “It happens that some producers were not too sure about which way to go, what technique to use, so they offer two products here and two products there... When that happens, I find it strange. A producer’s line of products should reflect a specific zone as well as that producer’s philosophy. A winery is not a supermarket.”

From these descriptions we expect that, given that attention and resources are limited, a winery spanning multiple styles knows and learns less about any style than a producer focusing on only one. The resulting engagement ought to be reflected in the lower “actual” quality of the wine in blind tastings. Moreover, external earmarks, particularly the identity of a winery, will make evident the spanning of styles in the evaluation of wines by the Veronelli guide because they will be aware of the incongruence between identity and the wines they taste. The evidence of spanning styles will affect evaluations that the wines receive beyond their potentially inferior quality or the specific preference for the style in which they are made.

### 3.3. Data and Measures

Our data cover all Barolo and Barbaresco wine labels listed in the Gambero Rosso and Veronelli guides from their first editions (1988 and 1991, respectively) to 2006. The wineries we study are subject to stringent production standards, and the relevant Italian law sets a minimum aging period before the product can be released in the market. For Barolos and Barbarescos, this period is two and three years, respectively, but it increases to five years for riserva wines. Critics review wines upon release, so this rule imposes a systematic lag structure on the rating process and, consequently, on our data. To allow for comparability across the producers, we end our observation window before the vintage 2001. Our fieldwork revealed that experimentation with cellar techniques was negligible before the 1980s. Despite the fact that the guide goes back to review a few notable wines in earlier vintages, we start our analysis in 1981.

Our sample thus comprises all the labels of Barbarescos and Barolos from the vintages 1981–2000 that were reviewed by the Veronelli and Gambero Rosso guides. This initial sample includes 1,926 label–vintage–producer observations for 164 producers—the same used by Negro et al. (2010a) minus 26 observations lost due to missing values in one covariate not included therein. The estimation generally follows that analysis with minor differences noted below.

#### 3.3.1. Dependent Variable

We used the ratings received by the Veronelli and Gambero Rosso guides to measure how favorably a wine was evaluated. The Veronelli guide uses a four-point scale to communicate assessments of labeled wines and assigns a rating expressed in stars, ranging from one (“good wine”) to three (“excellent wine”). (Since the 1992 edition, the guide assigns a trifolium symbol to products reviewed for the first time, which indicates the overall good quality of the wine, but this does not imply any specific rating. These wines, and consequently the producers...
whose wines are all rated for the first time, are not included in the analyses.) The guide also confers an award of distinction (“super three stars”) to wines considered outstanding, which we coded as four stars. This star rating expresses a general evaluation of a wine’s features, or, as the editor put it, “the wine’s speed cruise.” The Gambero Rosso guide assigns one black glass, two black glasses, two red glasses, and three red glasses (where red is a higher ranking than black). In these instances, we treated two red glasses as a value of 3 and three red glasses as a value of 4.

3.3.2. Independent Variables. The first independent variable is producers’ spanning across wine styles. Each producer can demonstrate categorical adherence by producing wines that are more or less focused in a particular style—traditional or international. Producers can release multiple wines each year, and each wine can be vini
died using one of the two different techniques, botti or barriques (or a mix), as described previously.

We follow the notion that the two categories corre
dpond to styles associated with each method of produc
tion (aging). With the understanding that wines are released each year, a winery’s portfolio, and therefore spanning, can change depending on the methods used to make their wines. A winery can focus and not span by producing wines only in one style. For example, Bartolo Mascarello and Giacomo Conterno are representative of the traditional category and use only botti; La Spinetta is representative of the international style and uses only barriques. Wineries such as Poderi Luigi Einaudi make wines in all styles and show a significant degree of span
ning. Style spanning is expressed by an index of con
centration (1 minus the Simpson’s index of diversity), calculated as the sum of the squared proportions of the wines in a winery’s portfolio made in each style in a vintage. More formally,

\[ \text{Style spanning} (\mu(s, x, t)) = 1 - \sum_{s \in S} \mu^2(s, x, t), \]

where \( s \in S \) is the set of winemaking styles in our analysis for winery \( x \) at time \( t \). This measure ranges from 0 to 0.5, where 0.5 means they produced wines equally in each of the two styles (and are maximally spanning), and 0 means they only released wines in one style (and are not spanning at all).

The hypotheses are tested by including style spanning as a covariate when estimating the ratings assigned by the wine guides. As a control for the underlying quality characteristics of the wine (Benjamin and Podolny 1999), we add on the right-hand side of the regression equation the rating assigned blindly by the Gambero Rosso guide to each wine in the sample.

3.3.3. Control Variables. Our models include controls for additional observable product characteristics, producer characteristics, and contextual factors. First, we include two dichotomous variables equal to 1 if the wine is a riserva or a cru (a “first-growth” place). Reserve wines age for a longer period, and cru wines are vinified from particular vineyards. Second, we recognize that spanning can occur at the product level as well as the producer level and that critics may have preferences for specific styles of wine. We created a pair of dichotomous variables, barrique-only wine and barrique–botti wine, and set them equal to 1 if the wine was aged using only barrique or was aged using barriques and botti together, respectively, and 0 otherwise. This leaves the traditional wines (botti) as the reference category.

We add a control for the status of the winery to account for halo effects that can influence the ratings and include a variable measuring the lagged maximum rating awarded to a winery by the Veronelli guide. Missing values in this covariate reduce the sample by 26 observations compared to that used by Negro et al. (2010a), which did not include this variable in the esti
mates. A variable counting the number of Barolo and Barbaresco labels released by the winery in each vintage, number of labels, is added to address that style spanning increases with the number of different wines made by a producer. Finally, we add (but do not report) fixed effects for each vintage to account for weather-related fluc
tuations and other unobserved time-dependent effects. A summary of the variables in the regressions and their correlations is reported in Tables 1 and 2.

The star rating outcome is an ordered categorical variable. However, cut-point estimates obtained from an ordered logistic regression of the number of stars received by a wine did not suggest significant devi
ations from linearity. Thus, to ease interpretation, the estimates we report are obtained using ordinary least squares (OLS) regression with robust standard errors clustered by producer to adjust for within-winery corre
clation. Results of these OLS estimates are presented in Tables 3 and 4. In Table 5 we present additional results obtained using winery fixed-effects regression and gen
eralized random effects at the winery level to address whether the effects of category spanning stem from between-producer variation or within-producer variation.

Table 1 Summary Statistics of Variables Included in the Analysis of the Determinants of the Ratings of Barolo and Barbaresco Wines—Veronelli Guide

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>SD</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Veronelli rating</td>
<td>2.366</td>
<td>0.790</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>Style spanning</td>
<td>0.097</td>
<td>0.185</td>
<td>0</td>
<td>0.5</td>
</tr>
<tr>
<td>Gambero Rosso rating</td>
<td>2.152</td>
<td>0.744</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>Riserva wine</td>
<td>0.050</td>
<td>0.218</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Cru wine</td>
<td>0.881</td>
<td>0.324</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Barrique-only wine</td>
<td>0.272</td>
<td>0.445</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Barrique–botti wine</td>
<td>0.248</td>
<td>0.432</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Maximum prior rating</td>
<td>2.567</td>
<td>0.969</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>Number of labels</td>
<td>6.577</td>
<td>3.242</td>
<td>1</td>
<td>19</td>
</tr>
</tbody>
</table>
the spanning by a winery across styles, the lower the ratings of its wines ($X^2 = 4.25$, $p = 0.039$, df = 1), as judged by critics blind to the identity of the winery. The effect of prior maximum stars, our control for the status of a winemaker, is significant and positive, suggesting that past demonstrations of value lead to success in future years as well (Roberts and Reagans 2007). We contend that this effect is attributable to superior product quality because the raters taste the wines blindly.

In Table 4 we estimate the effects of category spanning on the ratings assigned to wines by the Veronelli guide. Model 1 estimates the effect of the control variables. Of the ones that show a significant effect, they generally behave as expected. Note again the positive and statistically significant effect of prior maximum stars ($X^2 = 264.02$, $p = 0.000$, df = 1). Wines that are designated as reserve or come from a first-growth vineyard are more favorably rated. Wines aged in barriques are preferred by the critics as well, perhaps because these wines, aged only in small barrels, tend to reach a more mature taste earlier, as they lose hard tannins faster.

Model 2 includes the covariate measuring category spanning, and we find that greater spanning across winemaking styles reduces ratings significantly ($X^2 = 7.80$, $p = 0.005$, df = 1). This result suggests preliminary
support for our second hypothesis, but at this stage, the estimates do not yet include our control for the actual quality of the wine; therefore, we are unable to draw conclusions as to which mechanism is driving the effect of spanning. We can, however, interpret the coefficient. For a one-standard-deviation increase in style spanning by a winemaker, the wine produced is predicted to have a 0.04 star discount (−0.2165 × 0.185), holding all other variables constant.

Model 3 adds the covariate measuring the Gambero Rosso ratings, which are obtained from blind tastings and, as demonstrated above, act as a control for the underlying quality of the wine. Not surprisingly, the covariate has a positive and significant effect on the star rating received from the Veronelli guide, suggesting that nonblind evaluations also incorporate “actual” quality ($X^2 = 115.5, p < 0.000, df = 1$). Interestingly, and more to our point, the negative effect of style spanning by the individual winery continues to reduce ratings, despite controls for the Gambero Rosso glasses ($X^2 = 5.47, p = 0.019, df = 1$). Including controls for blind ratings, a one-standard-deviation increase in style spanning is predicted to decrease the star ratings a wine receives by 0.03 (−0.174 × 0.185), holding all other variables constant. Note that our models also include controls for the methods with which the wine was aged; thus, wineries spanning styles receive lower ratings even after accounting for a wine’s specific style. This result now lends direct support to our second hypothesis.

Figure 1 plots the predicted star ratings the Veronelli guide would award a wine as a function of the observed range of style spanning from the estimates contained in Models 2 and 3. The figure shows that as style spanning increases, the predicted star rating by the Veronelli guide drops, when all other variables are at their means. The lower, solid line represents the reduction in Veronelli star ratings that a wine will receive as a function of both revealed identity and blindly assigned Gambero Rosso rating, which was derived from Model 2 above as the Veronelli rating without a control for blind rating. The effect on Veronelli star ratings of identity only, upon the inclusion of a control for Gambero Rosso rating, is represented by the upper, dashed line. Here, we see that spanning by a winery continues to negatively affect how their wines are rated, even controlling for actual quality.

More specifically, a winery that moves from aging wines exclusively in botti to aging half of their wines with barriques (representing a move from 0 spanning to 0.5) sees a reduction in star ratings for its wines from actual and perceptual effects of 0.11 stars ($−0.2165 × 0.185$), holding all other variables constant.

<table>
<thead>
<tr>
<th>Table 5 Determinants of the Ratings of Barolo and Barbaresco Wines—Veronelli Guide (Additional Analyses)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 4</td>
</tr>
<tr>
<td>GEE</td>
</tr>
<tr>
<td>Fixed effects</td>
</tr>
<tr>
<td>Constant</td>
</tr>
<tr>
<td>Style spanning</td>
</tr>
<tr>
<td>Gambero Rosso rating</td>
</tr>
<tr>
<td>Gambero Rosso first</td>
</tr>
<tr>
<td>Riserva wine</td>
</tr>
<tr>
<td>Cru wine</td>
</tr>
<tr>
<td>Barrique-only wine</td>
</tr>
<tr>
<td>Barrique–botti wine</td>
</tr>
<tr>
<td>Maximum prior rating</td>
</tr>
<tr>
<td>Number of labels</td>
</tr>
<tr>
<td>Time</td>
</tr>
<tr>
<td>Gambero Rosso rating × Time</td>
</tr>
<tr>
<td>Wald $X^2$</td>
</tr>
<tr>
<td>Model $R^2$</td>
</tr>
<tr>
<td>Number of observations</td>
</tr>
</tbody>
</table>

Notes. Robust standard errors adjusted for clustering on winery are shown. The specifications include fixed effects for vintage.

$^* p < 0.05$ (two-tailed).
The same winery risks a reduction in star ratings of 0.09 (−0.174 × 0.5) merely from their demonstrations of identity-confusing features. Although this seems like a modest effect, we can evaluate a market implication. Our analysis in the appendix indicates that the average effect of Veronelli’s star ratings on retail prices is worth about 7.3 dollars. The effect of style spanning then translates into approximately 75 cents per bottle. Considering the very small size of the wineries in the Langhe, which typically sell 30,000–50,000 bottles of Barolo and Barbaresco wines yearly, the overall impact of spanning can reduce the value of retail sales on the international market by about 30 thousand dollars, a not insignificant amount.

Another way to examine this effect is to look at the relative importance of spanning on actual versus perceptual discounts. In this particular case, approximately 80% (0.174/0.217) of the rating decrease due to style spanning for the Veronelli guide is associated with the audience connecting the identity of the winery to the schemas of winemaking styles. We return to the substantive implication of this finding in the Discussion section.

Of note, in Model 3, a winery’s status, as measured by its prior maximum stars, continues to increase ratings. The finding that status is correlated with and indeed a reward for value is not novel; often, however, status processes are difficult to separate from quality or the role of power. Because we control for underlying value of the product, this effect can reflect deference and/or power. Recall that in our setting producers are of very small size, thereby making it unlikely they will occupy dominant positions that make them capable of threatening critics, although our informants noted that some winemakers may weaken their good relations if rated unfavorably. Then, the maximum stars likely capture a significant deference effect.

Next we deal with two empirical concerns. One is that our clustered standard errors can still be subject to misspecification of the correlation structure between observations affecting the consistency of our estimates. Second, our analyses can imply both within- and between-producer effects of style spanning. Model 4 estimates panel regressions with generalized random effects; we used the method of generalized estimation equations (GEEs) implemented with an exchangeable error structure at the producer level. This allows a random producer effect subject to the constraint that the random effects are uncorrelated between producers and that the correlation over time of the random effects is constant. Then, in Model 5 we present estimates of fixed-effects regressions. Whereas the previous model specifications use all the observed variation, fixed-effects regression considers the relationship between changes in style spanning and critical ratings within each producer. The advantage of this method is that it removes the effect of any unobservable producer characteristics that can bias the impact of spanning. Results in both analyses continue to support the argument that style spanning exerts a negative effect on nonblind ratings by the Veronelli guide in the presence of controls for actual quality.

The ratings could be influenced by aims of differentiation or conformity of one critic vis-à-vis the other. If this were the case, then the timing of a review versus another could identify whether the critics are “following” each other or not. Both guides are published every year in October within days from each other and mostly review newly released wines, and we believe that the opportunity for influence is limited. However, in our data, 137 wines were reviewed in an earlier edition by the Gambero Rosso guide, and these reviews could be “used” by the Veronelli guide to distinguish from or imitate the other guide. Model 6 reestimates our main model by adding a dummy for wines that were reviewed first by the Gambero Rosso guide. This variable does not affect the Veronelli ratings, suggesting that these wines are not evaluated differently, nor interfere with other effects. In unreported analyses, we did not find any significant interaction with the main covariates.

We proffer that the value of underlying quality can change over time and that our estimation approach to include the Gambero Rosso ratings as a control does not measure the producer-side effect of category spanning equally well over the course of the analysis period. If that were the case, one could detect time variation of the estimates of the Gambero Rosso ratings coefficient. In Model 7 we add to the main specification a calendar year clock and its interaction term with the Gambero Rosso ratings. The interaction, while leaving the other findings unaltered, does not show statistical significance and does not improve the explanatory power of the model over the main specification of Model 3.

5. Discussion

The findings reported above support our general hypothesis that producers spanning multiple categories experience lower market success in the form of lower ratings assigned by specialized critics. Our empirical setting allowed us to distinguish between two conditions under which wines were rated—one where the identity of the producer was known and the other where it was unknown. This allowed us to separately identify a skill-based mechanism from a perception-based one, both of which have been invoked to explain the category-spanning discount (Zuckerman 1999, Hsu 2006, Hannan et al. 2007, Hannan 2010).

In particular, our analysis indicated that once we controlled for the underlying quality of wines, as expressed by ratings assigned by critics conducting blind tastings, the wines made by producers spanning wine styles continued to receive lower ratings. A noteworthy result is
that the category-spanning discount we observe seems to be largely intact after actual effects are controlled for; that is, our comparison of the decrease in ratings attributed to perceptions of identity, after the addition of the blind ratings, suggests an approximate 20% reduction of the effect of identity on nonblind ratings. This underscores the idea that the identity of a producer plays a dominant role in the evaluation process, at least in these estimates.

A possible limitation to our test of the first hypothesis is the risk that the Gambero Rosso critics inferred the identity of the winemaker at the tasting. That is, when tasting a wine blindly, the critic could guess the categorical identity of the winemaker and thereby judge the wine accordingly. This is a reasonable concern to the extent that there is a different taste to the wines produced by different styles and a correlation between a producer’s categorical identity and the types of wines they produce. Given the nonexperimen-
tal nature of our design, we cannot rule out this possibility for the first hypothesis. However, if the ratings from blind tasting act as a “control” for underlying features, our test for the second hypothesis is actually more conservative; that is, if the blind tastings account partially for “guessed” identity, then the fact that we find an effect after this control suggests a very strong effect of “visible” identity on Veronelli ratings (or it suggests that the Gambero Rosso critics are not guessing or not guessing well).

In other words, we are unable to definitely identify whether our measure for quality is completely subsumed by the Veronelli guide or whether the Gambero Rosso critics take into account additional characteristics of the wines that may not be considered by the Veronelli guide. The image this conjures is the potential for two Venn diagrams: one whereby quality characteristics as evaluated by the Gambero Rosso guide are a complete subset of criteria by which the Veronelli guide rates their wines and another whereby the Gambero Rosso and the Veronelli criteria are two intersecting but not completely overlapping lists. The intersection, in either scenario, is the overlap in quality characteristics that the Veronelli and Gambero Rosso guides both rate on. (There is an overlap because of the correlated nature of the ratings by the two guides.) However, we investigated this further by flipping our analyses by predicting Gambero Rosso ratings with the Veronelli ratings. If the Veronelli guide’s criteria completely subsume all of the Gambero Rosso guide’s, we can still expect a significant effect. And indeed, we find that the nonblind ratings affect the blind ratings. We also find that once we include the Veronelli ratings, producer spanning does not significantly affect the blind ratings, which further points to the subset interpretation.

This study’s main contribution to the growing literature on market categories is to produce direct evidence that identity underlies the link between category specialization and greater market success. A previous study of category defection in the wine industry (Negro et al. 2011) demonstrated that higher spanning by some producers leads to reduced likelihood of traditional producers leaving their category, ostensibly because those producers who decided not to span across categories find value in staying. In this paper, we provide complementary evidence of this process by showing how those same spanning decisions can lead to both quality and identity discounts from other members of the audience, particularly critics.

The study also offers a general methodological contribution. Evaluations in blind and nonblind conditions are appropriate to test the influence of actual versus perceptual mechanisms on audience assessment. In other domains, researchers can replicate our approach and estimate how much each mechanism matters. These conditions are not unique to wine tastings and are found elsewhere—for example, screened auditions in arts (Goldin and Rouse 2000) or commercial releases under pseudonyms (Phillips and Kim 2009).

Our regressions account for a winery’s status, and we find that lagged demonstrations of value are associated with higher ratings (Roberts and Reagans 2007). In additional analyses, we explored whether the effect of style focus varies with producer status. We did not find any significant effects of interactions with lagged maximum stars, and these models showed no significant fit improvement over the models presented in Table 4. Additionally, we found nonsignificant interactions of style focus with the botti and barrique indicators, which can also be used to test differences in status of the product categories (details of these tests are available from the authors upon request). So the category-spanning result is significant not only because it isolates the role of identity from that of underlying characteristics but also because it holds net of the influence of the status of the producer or product (Fleischer 2009, Waguespack and Sorenson 2011, Roberts et al. 2011).

We were also curious as to how spanning differentially affects wines dependent on their ratings. In unreported analyses, we found that the negative effects of spanning were greatest for wines that were rated two or three stars, suggesting that wines that are clearly high or low quality are not as affected by their producer’s identity. We speculate this can be due to the lower uncertainty surrounding these wines. On the other hand, there was likely more uncertainty to rating wines in the middle range. Consequently, the Veronelli critics seemed to be more influenced by a producer’s identity in these decisions. Much work has suggested that under conditions of uncertainty, market participants look to alternative information of a producer or product’s quality (Spence 1973, Zuckerman et al. 2003). In our setting, the clear identity of a producer seems to act as a positive cue to wine critics to increase ratings for wines in this middle range.
In unreported analyses, we started exploring one extension of the test presented here. Recent research argues that the perception of category boundaries can shape audience assessments. Hannan et al. (2007) proposed the notion of contrast to describe the degree of distinctiveness categories have against their social background. A category with high contrast means that it stands out more, in the sense that it has sharper and more easily recognizable boundaries. One can expect that with greater contrast the appeal of any offering that fits well into a category will increase. Members of the audience react positively to such clarity (Zuckerman 2004, Carroll et al. 2010, Kovács and Hannan 2010, Negro et al. 2010a).

We followed this approach and examined the effect of style contrast, which measures the overall agreement on the schemas applied by the audience to Barolo and Barbaresco wines, calculated as the average style spanning of all wineries in each vintage (Bogaert et al. 2010, Kovács and Hannan 2010, Carroll et al. 2010, Negro et al. 2010a). We found that style contrast increased the Veronelli ratings when the Gambero Rosso ratings were not controlled for, and separately increased the Gambero Rosso ratings. However, it did not significantly affect the Veronelli ratings after accounting for the number of glasses awarded by the Gambero Rosso guide. If the role of contrast depended on the priming of a producer’s identity, we would not expect to observe a significant effect on the ratings conducted with a blind method. Instead, we did. That higher contrast increases appeal whether the identity of the producer was observ- able or not is consistent with the idea that the activation of schemas does not need deliberation (DiMaggio 1997, Murphy 2004).

To us, rather than indicating that style contrast does not systematically affect ratings, this pattern suggests that the perception of category boundaries affects similarly the evaluation processes of the two critics even if they use different tasting methods. Higher contrast indicates greater awareness of category membership, leading to a greater likelihood that particular category schemas will be activated. Instances of a category will be easier to recognize even without priming a producer’s identity, and with increased recognition comes increased liking (Zajonc 1968).

This interpretation also mirrors a central idea in sociological theories of organizations that posits that cognitive representations assume taken-for-granted character and operate tacitly to influence perceptions and behaviors of market agents (Meyer and Rowan 1977, Hannan and Freeman 1989, DiMaggio and Powell 1983). We must, of course, caution that these implications were examined indirectly, yet we hope that future studies will examine more in depth under what specific conditions schemas for market categories can be activated.

Acknowledgments

This paper draws on work done in collaboration with Michael Hannan and Hayagreeva Rao. The authors thank them, Jerker Demrell, Dave McKendrick, Chris Rider, Peter Roberts, Amanda Sharkey, Olav Sorensen, Jim Wade, Ezra Zuckerman, two anonymous reviewers, and seminar participants at the 2007 Academy of Management Meetings in Philadelphia, American Sociological Association New York, and Stanford University for helpful comments on earlier versions of this manuscript. The authors also thank their informants for their willingness to spend time with them and for allowing them to quote them by name. Contributions were equal, and all errors remain the authors’ responsibility.

Appendix. Determinants of Internet Retail Prices of Barolo and Barbaresco Wines (Estimates of OLS Regressions)

<table>
<thead>
<tr>
<th></th>
<th>Coeff.</th>
<th>SE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>30.821</td>
<td>6.652</td>
</tr>
<tr>
<td>Veronelli rating</td>
<td>7.286</td>
<td>1.836</td>
</tr>
<tr>
<td>Gambero Rosso rating</td>
<td>3.880</td>
<td>1.791</td>
</tr>
<tr>
<td>Riserva wine</td>
<td>−6.183</td>
<td>5.519</td>
</tr>
<tr>
<td>Cru wine</td>
<td>−3.690</td>
<td>2.796</td>
</tr>
<tr>
<td>Barrique-only wine</td>
<td>3.539</td>
<td>4.659</td>
</tr>
<tr>
<td>Barrique–botti wine</td>
<td>−2.459</td>
<td>3.211</td>
</tr>
<tr>
<td>Number of labels</td>
<td>1.303</td>
<td>0.481</td>
</tr>
<tr>
<td>Model $R^2$</td>
<td>0.183</td>
<td></td>
</tr>
<tr>
<td>Number of observations</td>
<td>422</td>
<td></td>
</tr>
</tbody>
</table>

Notes. Robust standard errors adjusted for clustering on winery are shown. The specification includes fixed effects for vintage. $^* p < 0.05$ (two-tailed).

Endnotes

1 I Vini d’Italia (1988–) is published by Gambero Rosso Editore and Slow Food Editore, Roma, Italy; and I Vini di Veronelli (1991–) is published by Veronelli Editore, Bergamo, Italy.

References


Negro and Leung: “Actual” and Perceptual Effects of Category Spanning

Giaccomo Negro is an associate professor of organization and management at the Goizueta Business School, Emory University. He studies the ecology of organizations, categories, and identities in markets.

Ming D. Leung is an assistant professor at the Haas School of Business, University of California, Berkeley. His research area is in organizational and economic sociology with a focus on the effects of classification schema on market participant decisions and outcomes. Currently, he is studying online markets for freelance labor and personal loans. His current work examines how signals of credibility and reputation are interpreted in these virtual settings and how they may lead to biased outcomes.