THE STRENGTH OF WEAK BOUNDARIES: CATEGORY INFERENCES AND EVALUATION SPILLOVERS IN INTERNATIONAL LEGAL SERVICES MARKET

Lionel PAOLELLA
Judge Business School
University of Cambridge
l.paolella@jbs.cam.ac.uk

ABSTRACT
To explore the consequences of category spanning on audience appeal, most of the studies only take into account an overall evaluation of multiple category members, but not an evaluation for each category spanned. Everything takes place as if a multi-category firm receives a unique and all-encompassing evaluation. Yet, a multi-category firm gets several audience evaluations – one for each category spanned – that affect each other. This paper fills this gap between the empirical tests (1 unique overall evaluation for multi-category members) and the theoretical assumption in the literature (several specific evaluations connected by audiences leading to confusion). In the corporate legal services industry, this paper explores to what extent an organizations’ evaluation in one category is influenced by how appealing audiences perceive the organization to be in other categories. First, I present empirical evidence that the strength of the inferences conveyed by past evaluations in non-focal categories impacts a firm’s evaluation in a focal category. Second when firm’s evaluations across different categories are more dispersed or unclear, a firm is more likely to receive a lower evaluation in the focal category. Third, I suggest that a firm’s evaluation in a focal category is likely to be lower affected by non-focal categories’ evaluation when the latter are more similar to each other.

Key words: categories, inferences, external evaluation, law firm, signal.
THE STRENGTH OF WEAK BOUNDARIES:
CATEGORY INFERENCES AND EVALUATION SPILLOVERS IN INTERNATIONAL LEGAL SERVICES MARKET

1. INTRODUCTION

How firm outcomes are affected by social categories is one of the recent questions of organization theory. As consensual labels attached to similar organizations, categories have been proved to be fertile to explain a wide array of phenomena such as audience appeal (Hsu, 2006; Negro, Hannan, Rao, 2010; Pontikes, 2012), definition of competitors (Porac et al., 1995), institutional change (Rao et al., 2003), evaluative schemas of raters (Durand et al., 2007; Fleisher, 2009; Ruef and Patterson, 2009), strategic positioning opportunities (Granqvist, Grodal, Woolley, 2013; Durand and Vergne, 2014), survival rates of organizations (Bogaert et al. 2010; Carroll et al. 2010), and variation in identity perceptions (Hsu and Elsbach, 2013). Thus categories in economic settings delineate different spaces with specific expectations. As an organization is a member of a given category, it is expected to behave in certain ways in compliance with codes of the category. If categorical boundaries assist audiences in lumping together similar organizations and separating different ones, then does an organization’s evaluation also remain similarly bounded? In other words, when an organization spans categories, are there any spillover effects across categorical boundaries?

Pervasive consensus in literature indicates that category-spanning organizations suffer economic and social disadvantages as their offerings and activities confuse audiences (Hannan, Polos, and Carroll, 2007: 108). Therefore, by showing that multi-membership firms can cloud an audience when they infringe category boundaries, literature assumes that this audience makes a connection between the multiple memberships of firms. However, despite the vast interest and research in this area, less is understood about such a connection in audience’s perceptions and the relationships between the categories spanned by organizations. To date, most of the studies in literature only take into account an overall evaluation of multiple category members, but not an evaluation for each category spanned. Everything takes place as if a multi-category firm receives a unique and all-encompassing evaluation. Yet, a multi-category firm gets several evaluations – one for each category spanned – that affect each other.
In parallel, a stream of research has recently proposed theoretical explanation for cross-firm spillovers based on categorization literature. The wrongdoing behavior of an organization in given category is likely to taint the other organizations in the same category (Barnett and King, 2008; Jonsson et al., 2009; Yu et al., 2008). For instance, further to one firm suffers a toxic chemical accident with an input, investors will punish other users of that input by discounting their stocks (Diestre and Rajagopalan, 2014). What scholars do less effectively is to account for within-firm spillovers beyond category boundaries. Thus inter-category spillover effects represent a case where current knowledge needs amending. We know little about the relationships between categorical boundaries and specifically how evaluation in one category can override boundaries and percolates onto another category. How and to what extent a multi-category organization’s evaluation in non-focal categories can influence its evaluation in a focal category? By ‘focal category’, I name the category a firm \( x \) is in at time \( t \). I label ‘non-focal categories’ the set of categories the firm \( x \) is in at time \( t-1 \).

This paper furthers the literature on the role of categories and evaluation’s spillovers in organizational life by exploring cognitive mechanisms by which an audience extends into a one category the firm’s evaluation gathered from other categories. Rather than starting from scratch when interacting with a multiple category organization, audience members benefit from past information by making inferences from one category membership to another. As inferring “is to transform a piece of information into another piece of information, to go from one category to another” (Robert, 2005: 701), the evaluation of an organization in a given category is likely to be affected by its previous evaluation from another categories. Further, this paper discusses several factors that mitigate these spillover effects as the consistency of a firm’s average evaluation in non-focal categories or the degree of similarity between categories spanned.

These ideas are investigated in a longitudinal analysis of a unique dataset on the corporate legal services industry, including information on the evaluation of law firms’ practice areas between 2000 and 2010. I collected data from three professional guides (The Chambers and Partners, The Legal 500, and PLC Whichlawyer) that reflect clients’ opinions about corporate law firms over eight practice areas of law in three different locations (London, New York City, and Paris). The international corporate legal services market meets all the conditions to test the theory. First, categorization and evaluation of law firms stem from external parties (Hsu and Hannan, 2005: 477): several guides rate law firms across distinct practice areas. Second, these data enable me to unfold the overall evaluation of multi-
practice law firms into a series of evaluations for each of their practice areas (at most eight ratings for the same firm: one rating per practice area). Finally, category-based inferences play only in contexts of imperfect information (Shapiro, 1983) and unobservable quality (Podolny, 1993). In legal industry, asymmetries of information between clients and law firms as well as difficult-to-assess quality of legal services force clients to rely on signals to infer firms’ attributes.

2. THEORY BACKGROUND

2.1. MARKET CATEGORIES AS SOURCE OF INFERENCES

Audience members – a set of homogenous actors who control material and symbolic resources (Hsu & Hannan, 2005) – in markets do not possess perfect information on the quality and value of the commodities potentially exchanged (Podolny, 1993). As such, they try to gather as much information as possible to make decisions when they engage in the selection of an exchange partner. In so doing, actors rely their decision on cues that allow them to substitute “complex tasks of assessing probabilities and predicting values” into “simpler judgmental operations” (Tversky and Kahneman, 1974: 1124). By extracting regularities from some items (products, organizations) and dividing them into clusters based on their similarity (Rosch, 1973), categories offer such cues and information content. Audiences allocate some properties to firms that fall into recognizable categories (Polos et al., 2002). A category is a label with evaluative schemas attached for a set of entities that share one or more properties. For example, if an audience labels a movie as ‘comedy’, it is expecting to laugh. Market categories condense information and predictions about candidates’ properties and make possible connection between producers and buyers over time and through space. They encompass all the necessary and definite features to engage in transactions and exchange. Category membership assures that this firm will respect some codes (Polos et al., 2002) and will pursue certain types of actions in commonly known directions. For example, when clients taste a brand new restaurant of the category nouvelle cuisine, they know what to expect. They anticipate some cooking rules, ingredients, and a certain menu organization (Rao, Durand, Monin, 2003).

Thus categories help audiences to cope with uncertainty on markets. As a cognitive shortcut, category membership allows audiences to identify new entity (e.g., an organization) and then to infer unknown properties and make predictions based on the definitional attributes of the category. That said, a category “delivers specialized packages of inferences to guide an agent’s interactions with particular category members in specific situations” (Barsalou, 2005: 183).
Inferences consist in the process of applying category knowledge to a new object (Murphy, 2004). Category-based inferences describe the mechanism by which audience members are making judgments about firms’ unknown properties based on their known properties. By virtue of a logical association with the observation of firm’s features, audiences derive a conclusion about the other firm’s features unknown. As such, audiences judge what a new member is worth by inferring on already-known properties of the category affiliation and relying on them. They compare candidates’ features to the category prototype’s properties (Durand and Paolella, 2013) and high matching between them increases audience appeal (Hsu, 2006) since inferences occurs more readily for candidates sharing a lot of commonalities with the definitional features of the category. Intra-category inferences depend then on organization-level properties and on how prototypical organizations are of the category as a whole. Based upon the extent to which an organization is associated with a category, audiences formulate predictions applicable in the given category.

Further, while categories are useful to make predictions about novel members, they are also meaningful to retrieve some features about its members perceived in another category. This type of inferences across categories carries within-firm evaluations of audiences from one known domain of firm’s activities to another unknown domain. Inter-category inferences depend then more on boundary-level properties and on how associated categories are to one another. Based upon the accuracy and the coverage of category-based evaluations, audiences formulate predictions applicable in another category. Specifically, this paper develops a series of hypotheses on how evaluation in one category can override boundaries and percolates onto another category.

### 2.2. CATEGORY-BASED INFERENCES AND SPILLOVER EFFECTS

The need to predict the likelihood of an unknown property (e.g., quality in one domain) given some known properties (e.g., quality in another domain) confronts actors on markets. Category-based evaluation can be used to guide such predictions. Good or bad category evaluation is respectively a positive or negative signal (Spence, 1973; Shapiro, 1983) that gives information on firm attributes and reduces uncertainty for potential partners. For audiences inspecting a multi-category firm, an assessment made in one category enjoys some relevancy in another category. Therefore, the signal of category evaluation is likely to cross boundaries as long as firms span categories. The perception of a category-spanner in a focal category triggers retrieving general information about it from its other memberships. Due to the inference-based mechanism, the evaluation of an organization in a given category
is thus likely to be affected by its evaluation from other categories. In that case, an organizations’ evaluation in a focal category will be influenced by how appealing audiences perceive the organization to be in other categories.

Inter-category inferences made by audience members generate an \textit{a priori} – positive or negative – about the organization that drives some expectations (Corter and Gluck, 1992). Audiences derive expectations via inferences from a category membership to another one and thus extend some features of an organization (Hampton, 1987). For a highly valued organization in some activities, features and information retrieved by clients are likely to bolster the appeal of the organization in another domain of activity. For instance, the firm reputation in category “tax” will provide a signal for audience members of the category “real estate” regarding firm skills and quality in its practice area. Drawing positive (or negative) conclusions about one category membership of an organization based on knowledge of its memberships in another categories will be as great as the signal be strongly positive (or negative). For a multi-category organization, a positive average evaluation among several practice areas fuels audiences’ inferences and increases the evaluation in a focal practice area. Therefore:

\textbf{Hypothesis 1. The higher the evaluation about an organization in non-focal categories, the higher will be its evaluation in a focal category.}

Category-based inferences are characterized by the generalization of a finite series of specific observations into a constant rule. If the evaluations about a category-spanner are consistent with each other, generalization will be stronger. Inter-category inferences are thus favored by higher but also homogeneous prior evaluation in given activities. The strength of inferences achieved by the common category system that economic actors share hinges on the ability for the audience to develop in mind a single and univocal evaluation about organizations. Conversely, category-based inferences may be complicated by plural and dissonant category evaluations of organizations. For multiple category members, divergences in evaluation across different categories could fade the inferences make by audiences. Divergent evaluations across the source categories do not provide clients with accurate and relevant information but with contradictory cues about firm’s potential quality services in the target category. If the assessment of an organization is unclear, that is, divergent, audiences will not be able to generate consistent inferences. Thus inconsistent evaluations across categories prevent audience members from making “generalizations sufficiently strong that they seem like laws” (Hacking, 1995: 352). Let assume a law firm that provides legal services
in three practice areas such as litigation, bankruptcy, and corporate. If this law firm receives in the meantime a positive evaluation in litigation, but a poor evaluation in corporate, what could clients predict for its practice in bankruptcy? That is, a discrepant evaluation of firm in some activities prevents audience members from making strong inferences about this firm in other activities. In presence of divergent signals, clients cannot make robust inferences. Furthermore, divergent evaluations will be interpreted as a negative cue from adverse-risk clients. Therefore, inter-category inferences will be as poor as previous evaluations in non-focal categories are disparate.

**Hypothesis 2.** The greater the inconsistency of evaluation about an organization in non-focal categories, the poorer will be its evaluation in a focal category.

Categories on markets delineate distinctive spaces that are interrelated. Similarity varies along with the different interrelated categories in which a firm operates. Indeed, it is difficult to assume that all structural relationships between categories spanned by a firm are equal. Thus audiences’ inter-category inferences will not be initiated in identical manners. The richness of information on which inferences are based matters (Osherson et al., 1990). Research in psychology shows that people rely on multiple categories to make inferences when they face with uncertainty (Heit, 1988, Anderson, 1991) in order to gather as much information as possible. Taking multiple categories into account is likely to increase the range of information and thus the inferences’ accuracy. Inter-category inferences will be more reliable based on a broader range of information than on focused and redundant (i.e. similar) information. Therefore, the greater the range of information retrieved from the source categories, the greater the inferences are likely to affect the target category. Further, beyond the mere number of non-focal categories the inferences are based upon, they will be lower if grounded on two similar non-focal categories than two dissimilar ones. The latter provide richer information than the former. For instance, in corporate legal market, the practice area ‘Competition/Antitrust’ is more similar with the practice ‘Corporate/M&A’ than the practice ‘Intellectual Property’ as the two first practice areas share many firms in common and less with the third one. I hypothesize that inter-category inferences between ‘Competition/Antitrust’ and ‘Corporate/M&A’ will be lower and so have a weaker impact on a third practice area’s evaluation. In the same line of thought, if a general counsel knows that a firm $x$ offers good legal services in “Corporate/M&A” and “Litigation” – quite distant areas of law – he would be confident that this firm $x$ is doing well in “Real Estate”. In contrast, suppose the same client knows that the firm $y$ offers legal services is “Corporate/M&A” and
“Tax” – quite similar areas of law – he would be probably less confident that the firm is doing well in “Real Estate” as he gets much less information about the firm (or more precisely two information quite redundant). Therefore:

**Hypothesis 3.** The greater the similarity between an organization’s non-focal categories, the poorer will be its evaluation in a focal category.

3. DATA AND METHODS

3.1. EMPIRICAL SETTING: THE LEADING CORPORATE LAW FIRMS IN LONDON, NEW YORK CITY, AND PARIS

To test my hypotheses, I needed an industry where quality of service is difficult to assess, uncertainty is high, and transactions occur repeatedly. These three conditions strengthen the key role of inter-category inferences in economic settings. The international corporate legal service market meets all these requirements. First, as law firms’ ability to deliver quality services is difficult and costly to observe, clients rely on social evaluation in lieu of actual quality observations to economize on search costs (Podolny, 1993). In this line, the results of a 2009 survey of 200 general counsels from 60 countries, conducted by a professional journal (International In-house Counsel Journal, 2010), confirmed that “market reputation” is the first choice selection criteria of legal providers. Law firms thus have to establish their reputation and market credibility being known as expert in their practice areas. Second, clients are much less expert than lawyers to identify and classify their legal issues. They face difficult-to-categorize legal problems and they cannot foresee all the issues their legal cases cover. In that case, they are likely to use their past evaluations of legal providers about a specific case to select them for a new issue. Third, inter-category inferences imply memory retrieval as the basis for judgments in repeated transactions with same economic entities. As the occasion to predict the likelihood of an unknown property given some new properties occurs very often on legal services market, my setting is well suited for studying this mechanism in this regard. Clients have a pool of privileged providers but they face new legal problems and groundbreaking cases. For example, clients in Corporate-M&A, Litigation or Intellectual Property will be potential clients in Tax, Real Estate, or Employment and vice versa as its outside legal needs will evolve along its activities. Overall, under these three conditions, clients are likely to spread their evaluation of one practice area to another to assess service quality, reduce uncertainty, and engage in new transactions.
During the last decades of the 20th century, a trend toward the judicialization of the economic world led to the emergence of an international corporate legal services market based on the American model (Dezalay & Garth, 2004; Galanter & Henderson, 2008). In today’s legal arena, corporate clients have manifold projects that often encompass different law practice areas (Kor & Leblebici, 2005; Wilkins, 2009). Unlike in the past, clients today have less need for a lawyer who acts exclusively for their litigation cases, preferring instead a partner who can navigate them through their whole legal life (e.g., addressing both social and fiscal dimensions) and advise on major corporate events (e.g., mergers or acquisitions, or the creation of foreign subsidiaries). Thus, law firms across countries are mushrooming into large legal “department stores”, offering a panoply of services delivered by many lawyers (Galanter & Palay, 1994; Harper, 2013). Given that law firms are normally structured by practice disciplines (often regarded as “silos”), corporate law firms have been offering combinations of their various competencies rather than offering them each independently in an effort to fit their clients’ expectations (Chayes & Chayes, 1985; Harper, 2013):

Our firm has historically understood our clients’ needs and how we can best help them address those needs. Our clients want sophisticated and responsive legal service [and] are looking for a full package of services. We need to offer a set of diversified practices in different regional areas. (Interview with a UK law firm partner, London office)

If we want to develop our more profitable practice [areas], which are corporate/M&A or litigation, we need skills in tax, intellectual property, real estate, employment, occasionally in environmental law and so on. To close the deal, you generally need expertise in diverse areas. So, having partners and teams in these areas helps us to enhance our core practice to attract clients and make deals. That is generally true because otherwise, our competitors would say to clients ‘don’t give them this case because they won’t be able to handle this aspect of the deal, or this issue in, for example, intellectual property, tax, or employment.’ So for us, being a multi-practice firm is necessary to close the deal. (Interview with a US law firm partner, Paris office)

Therefore, suitable to test my conjectures, I use original data on corporate legal services collected in three professional legal directories (The Chambers and Partners, The Legal 500, and PLC Whichlawyer) that rank law firms both by categories – i.e. by practice areas – and by location. These directories are the most widespread over the business legal market. Based on extensive independent research, they track the most important trends in the legal profession and provide rankings of corporate law firms operating in various practice areas. These guides reflect the market’s opinions by collecting informed feedback from lawyers and clients (International In-house Counsel Journal, 2010: 21; Coates et al., 2011).
They do not directly assess legal services; they conduct interviews and their publications reflect the opinions of various audiences. Thus these rankings provide with a snapshot of activities and position of law firms in the market over a year. The research coverage reflects market conditions in each location taking into account on several factors and considerations. Law firms are ranked in their practice-area on the basis of their “technical legal ability, professional conduct, client service, commercial astuteness, diligence, commitment, and other qualities most valued by the client” (Chambers and Partners Editorial). A ranking in a given practice area relates to the firm’s department in this specific practice area, not to the firm as a whole. When a firm has several departments specializing in different areas of law, some of its departments may be ranked and some others not.

The legal directories adopt the same research methodology to identify the leading law firms and lawyers across the three jurisdictions. They judge effectiveness and capability of each department first by assessing the actual work done – deals, cases, reported contentious issues – via law firms’ submissions and second by interviews with those active in the market – mainly clients, in-house lawyers and peers. First, submissions provided by the law firms contain factual information which, when used in conjunction with other sources, help the directories to determine whether a firm’s department should be covered or not. Submissions provide the type and volume of deals achieved during the year. Second, the most important part of the methodology is feedback from clients. These legal directories have large team of researchers that contact lawyers and clients directly. They conduct thousands of interviews with clients, market commentators, lawyers, judges and others. Interviewees are selected on the basis of submissions put forward by law firms and extensive database of law firms’ clients. Throughout the interviews, they explore the main qualities valued by clients (value for money, professional conduct, commercial astuteness, service delivery, diligence, industry knowledge, commitment, technical ability) and assess recent work done. Confidential client interviews are given priority, as they tell whether lawyers truly provide the services clients want.

3.2. DATA

I collected rankings from 2000 to 2010 for 3 locations (London, New York City, and Paris) in eight different practice areas: Competition-Antitrust, Litigation, Intellectual Property, Real Estate, Tax, Corporate-M&A, Bankruptcy, and Employment. Contrary to an overall evaluation of multiple category membership, this dataset provides several grades for a given firm (one grade per practice area). Thus I can unpack multi-category firms’ evaluation
to study the spillover effects of evaluation from one practice area to another one. The risk set includes all the law firms that have been ranked at least in two practice areas in each guide. The level of analysis is the triad ‘firm-practice-guide’ in each location. For instance, ‘Linklaters-Chambers-real estate’ and ‘Linklaters-Legal500-real estate’ are two distinct entities in the dataset. In addition, ‘Linklaters-Legal500-real estate’ in Paris is distinct from ‘Linklaters-Legal500-real estate’ in London as well as in New York City.

**Dependent Variable. Evaluation in the focal practice.** Every practice of law firms covered by the three guides is placed in ranking bands. These bands range from 1 to n, with 1 being the best. However, in the two guides *Chambers* and *Legal 500*, law firms are ranked in band but the grading scale can change across practice areas. For example, 4 bands can be used to rank the practice “Tax”, and 6 bands to rank “Real Estate” (band 1 being always the best). Instead, the guide *PLC* uses a four-point grading scale: leading; highly recommended; recommended; recognized. Therefore, I took into account the position of the firm-practice-guide triad compared to its competitors by calculating for each guide *g* the evaluation of a firm *x* in a given practice *p* as follows:

\[
\text{Evaluation } x_{g,p} = 1 - \frac{\text{number of firms ahead } x_{g,p}}{\text{total firms ranked}_{g,p}}
\]

As such, best firm-practice-guide triads obtain always 1 as an evaluation that is in a decreasing order. Those firms at the bottom of their classification gain ratings closer to zero.

**Independent Variables. Average evaluation in the non-focal practice areas.** To test the first hypothesis, I calculated the average evaluation in the non-focal categories as follows:

\[
\text{Average evaluation in non focal practices } x_{g,p} = \sum_{n=1}^{7} \frac{x_n}{n}
\]

where *x* is a firm covered in guide *g* both in a focal practice *p* and non-focal practices *n*. Take as an example the firm *Linklaters* in Paris covered by the *Chambers and Partners* in three practice areas ‘Litigation’, ‘Tax’, and ‘Employment’ respectively with the evaluations 1, 0.7, 0.3. The values of the independent variable for the three distinct triads ‘Linklaters-Litigation-Chambers’, ‘Linklaters-Tax-Chambers’ and ‘Linklaters-Employment-Chambers’ will be respectively 0,5; 0,65; and 0,85 (e.g., 0,5 = (0.7+0.3)/2). I computed the variable in a similar way for the two other guides.
**Standard deviation of evaluation in the non-focal practice areas.** To test the second hypothesis and based on the first independent variable, I computed the standard deviation of evaluations obtained in the non-focal practice areas. Turning back to the previous example of the firm *Linklaters* covered by the directory *Chambers and Partners* in ‘Litigation’ (score: 1), ‘Tax’ (score: 0.7) and ‘Employment’ (score: 0.3), the values of the standard deviation for ‘*Linklaters-Litigation-Chambers*’, ‘*Linklaters-Tax-Chambers*’ and ‘*Linklaters-Employment-Chambers*’ will be respectively 0.2; 0.35; and 0.15.

**Similarity of the non-focal practice areas.** To test the third hypothesis, I calculated the average similarity between each pair of non-focal practice areas a firm is engaged. Following previous studies (Hsu, 2006; Hannan and Kovacs, 2011), I used the Jaccard similarity index to capture the similarity in terms of members between each pair of non-focal practice areas. The Jaccard coefficient takes the following form: 

\[ J(A, B) = \frac{A \cap B}{A \cup B} \]

where 

\( A \cap B \) indicates the cardinality of the set of firms covered in both practice areas A and B, and 

\( A \cup B \) the cardinality of the set of firms covered in A and/or B. The range of the independent variable is between 0 and 1. Some practice areas show no similarity, whereas others reach partial or full similarity in their members. For example, in the year 2001 in Paris the two practice areas ‘Tax’ and ‘Employment’ in *Chambers and Partners* count 9 firms in common for a total number of covered firms equal to 30. The degree of similarity in terms of members between these two categories is thus equal to \( 9/(30-9) = 0.428 \). As the firm *Linklaters* is only engaged in three categories ‘Litigation’, ‘Tax’ and ‘Employment’, the value of the independent variable for the firm-practice-guide triad ‘*Linklaters-Litigation-Chambers*’ in Paris will be 0.428. I reproduced the same calculation for the two other locations.

**Control Variables.** Previous research has shown that clarity regarding the meaning of categories increases the appeal of all members (Kovacs & Hannan, 2010; Kuilman & Li, 2009) and the signaling effect (Negro, Hannan and Fassiotto, 2014). Multi-category firms blur the saliency of the categories they are affiliated with, leading to audiences reacting negatively to such dissolution in clarity. I therefore controlled for the contrast of the focal category. Based on previous literature (e.g., Kovacs and Hannan, 2010: 184-185), I measured the contrast of a category as the sum of the grades of membership of the category members divided by the total number of members belonging to the category. The grade of membership is equal to the ratio: 1/number of categories the firm is in. For example, the grade of
membership of ‘Linklaters-Litigation-Chambers’ engaged in the three categories “Litigation”, “Tax” and “Employment” is then equal to 1/3. As an example of categorical contrast, if there are 3 members in the category ‘Litigation’ with respectively a grade of membership equals to 0.2, 1, and 0.6: the categorical contrast of the ‘Litigation’ category is equal to (0.2+1+0.6)/3=0.6. I also included the tenure of coverage of the firm in the focal practice (Pontikes, 2012) by calculating the cumulative number of years for successive presence in the guide that covers the firm-practice. I expect a positive relationship because the longer firms are engaged in the focal category, the closer they are to its core definitional attributes (Navis and Glynn, 2010).

At the firm level, I controlled for the scope of the firm, measured by the number of categories in which the firm-location dyad is covered by a guide, divided by the total number of categories covered by the guide (Phillips and Zuckerman, 2001). The visibility of law firms on the market may increase their audience appeal and evaluation (Karpik, 2010: 163). I therefore controlled for the local size of the firm, using the log of the total number of partners in each location. Due to globalization, clients expect their lawyers to provide services across many countries in which they have deals or cases and so are more likely to value highly internationalized law firms. I captured this effect of internationalization with the percentage of lawyers outside home country of the firm.

These directories contain a section where firms can buy professional cards within each jurisdiction. The full-page profiles are based on information provided by the participating firms. This profile has been approved by the firms prior to publication, and is completely separate and different from the editorial section (ranking and comment). However, to control for any ‘pay per play’ bias, I therefore controlled for the potential effect of advertisement by counting the average number of page profile purchased in each guide covering the firms. I also included dummies to control for the nationality of firms (i.e. headquarter’s location).

As some practice areas of law are more prestigious than others, their presence in category portfolio of firms may affect clients’ evaluation. I therefore created 8 dummy variables for the coverage of the firm in the given practice area in at least one guide. To capture guide-specific effects and some potential measurement errors in reflecting clients’ opinions that may influence evaluation of firms, I included dummy variables to flag the firm-practice’s presence in each guide. I also captured time effects and location effects by including a set of dummy variables in our models. The descriptive statistics and correlations for the variables used in the analysis of evaluation’s spillovers can be found in Tables 1 and 2.
Table 1. Descriptive Statistics (n=1671).

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>evaluation in focal practice</td>
<td>0.664</td>
<td>0.234</td>
<td>0.054</td>
<td>1</td>
</tr>
<tr>
<td>avg. evaluation other practice</td>
<td>0.810</td>
<td>0.203</td>
<td>0.034</td>
<td>1</td>
</tr>
<tr>
<td>std. dev. other practice</td>
<td>0.124</td>
<td>0.081</td>
<td>0</td>
<td>0.417</td>
</tr>
<tr>
<td>similarity between other practice</td>
<td>0.367</td>
<td>0.160</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>contrast focal practice</td>
<td>0.342</td>
<td>0.104</td>
<td>0.128</td>
<td>0.875</td>
</tr>
<tr>
<td>tenure in focal practice</td>
<td>5.064</td>
<td>2.794</td>
<td>1</td>
<td>11</td>
</tr>
<tr>
<td>firm scope</td>
<td>0.639</td>
<td>0.293</td>
<td>0.143</td>
<td>1</td>
</tr>
<tr>
<td>size</td>
<td>5.020</td>
<td>1.081</td>
<td>0</td>
<td>6.802</td>
</tr>
<tr>
<td>internationalization degree</td>
<td>27.433</td>
<td>22.901</td>
<td>0</td>
<td>85.995</td>
</tr>
<tr>
<td>advertisement in guide</td>
<td>0.858</td>
<td>0.862</td>
<td>0</td>
<td>4.5</td>
</tr>
</tbody>
</table>

Table 2. Pairwise Correlations (n=1671).

<table>
<thead>
<tr>
<th>Variables</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 evaluation in focal practice</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 avg. evaluation in other practice</td>
<td>0.42</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 std. dev. other practice</td>
<td>0.03</td>
<td>0.01</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 similarity between other practice</td>
<td>0.06</td>
<td>0.08</td>
<td>0.44</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 firm scope</td>
<td>0.10</td>
<td>0.03</td>
<td>-0.12</td>
<td>-0.25</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6 contrast focal practice</td>
<td>0.33</td>
<td>0.21</td>
<td>0.05</td>
<td>0.03</td>
<td>-0.13</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7 tenure in focal practice</td>
<td>0.33</td>
<td>0.43</td>
<td>0.44</td>
<td>0.47</td>
<td>-0.17</td>
<td>0.25</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8 size</td>
<td>0.26</td>
<td>0.33</td>
<td>0.13</td>
<td>0.07</td>
<td>-0.01</td>
<td>0.11</td>
<td>0.35</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9 internationalization degree</td>
<td>0.24</td>
<td>0.30</td>
<td>0.13</td>
<td>0.13</td>
<td>-0.04</td>
<td>0.13</td>
<td>0.40</td>
<td>0.54</td>
<td></td>
</tr>
<tr>
<td>10 advertisement in guide</td>
<td>0.11</td>
<td>0.13</td>
<td>0.09</td>
<td>0.09</td>
<td>-0.06</td>
<td>-0.02</td>
<td>0.13</td>
<td>0.00</td>
<td>0.06</td>
</tr>
</tbody>
</table>

3.3. MODELS

To explore the relationship between my predictors and the dependent variable within each entity (i.e. firm-practice-guide), I used fixed-effects linear models. Fixed-effects regression model estimates the within variation of the regressors for each entity over time. It allows a limited form of endogeneity as the regressors may be correlated with the unobserved entity-level effect. I used a robust version of the Hausman test proposed by Wooldridge (Wooldridge, 2002: 290-291) to see if the random effects would be consistent. The test rejected the null hypothesis ($p < .001$) and so confirmed that fixed effects regression model is appropriate. In addition, I ran a test that shows the time effects were jointly significant. Accordingly, I ran a two way fixed effect model (i.e. within entity and time effect). A likelihood-ratio test comparing the homoskedastic and heteroskedastic error term models concluded the presence of groupwise heteroskedasticity. I also tested the potential for serial correlation in the idiosyncratic error of the full model (Drukker, 2003; Wooldridge, 2002), showing the presence of autocorrelation. A last test (Pasaran Cross sectional dependence) was used to explore whether the residuals are correlated across entities sine cross sectional
dependence can lead to bias in tests results. This test failed to reject the null hypothesis that residuals are not correlated. For these three traditional issues in time-series panel data, I performed a fixed effects within regression with Driscoll and Kraay standard errors (1998) which is robust to general forms of heteroskedasticity, cross sectional as well as temporal dependence, independently of the number of observations and time periods (Hoechle, 2007: 286). All independent and control variables are lagged to enhance causal inference.

Table 3. Two-way Fixed Effects estimations: Effect on Focal practice’s evaluation.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Model 1 Control</th>
<th>Model 2 H1</th>
<th>Model 3 H2</th>
<th>Model 4 H3</th>
</tr>
</thead>
<tbody>
<tr>
<td>avg. evaluation non-focal practice</td>
<td>0.124***</td>
<td>0.153***</td>
<td>0.151***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.021)</td>
<td>(0.020)</td>
<td>(0.028)</td>
<td></td>
</tr>
<tr>
<td>std. dev. non-focal practice</td>
<td>-0.140***</td>
<td>-0.127***</td>
<td></td>
<td>-0.031*</td>
</tr>
<tr>
<td></td>
<td>(0.027)</td>
<td>(0.028)</td>
<td></td>
<td>(0.016)</td>
</tr>
<tr>
<td>similarity between other practice</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>contrast focal practice</td>
<td>0.384***</td>
<td>0.366***</td>
<td>0.373***</td>
<td>0.370***</td>
</tr>
<tr>
<td></td>
<td>(0.047)</td>
<td>(0.045)</td>
<td>(0.045)</td>
<td>(0.045)</td>
</tr>
<tr>
<td>tenure in focal practice</td>
<td>0.015***</td>
<td>0.015***</td>
<td>0.014***</td>
<td>0.014***</td>
</tr>
<tr>
<td></td>
<td>(0.003)</td>
<td>(0.003)</td>
<td>(0.003)</td>
<td>(0.003)</td>
</tr>
<tr>
<td>firm scope</td>
<td>0.041***</td>
<td>0.034*</td>
<td>0.033**</td>
<td>0.039**</td>
</tr>
<tr>
<td></td>
<td>(0.011)</td>
<td>(0.011)</td>
<td>(0.012)</td>
<td>(0.012)</td>
</tr>
<tr>
<td>size</td>
<td>0.009</td>
<td>0.007</td>
<td>0.008</td>
<td>0.008</td>
</tr>
<tr>
<td></td>
<td>(0.007)</td>
<td>(0.007)</td>
<td>(0.006)</td>
<td>(0.006)</td>
</tr>
<tr>
<td>internationalization degree</td>
<td>0.002</td>
<td>0.003</td>
<td>0.002</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td>(0.007)</td>
<td>(0.007)</td>
<td>(0.005)</td>
<td>(0.005)</td>
</tr>
<tr>
<td>advertisement in guide</td>
<td>0.001</td>
<td>0.001</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>(0.005)</td>
<td>(0.005)</td>
<td>(0.004)</td>
<td>(0.004)</td>
</tr>
<tr>
<td>guide dummies</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>practice dummies</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>year dummies</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>location dummies</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>nationality dummies</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>Constant</td>
<td>0.251***</td>
<td>0.173***</td>
<td>0.164***</td>
<td>0.167***</td>
</tr>
<tr>
<td></td>
<td>(0.043)</td>
<td>(0.040)</td>
<td>(0.036)</td>
<td>(0.036)</td>
</tr>
<tr>
<td>Observations</td>
<td>16741</td>
<td>16741</td>
<td>16741</td>
<td>16741</td>
</tr>
<tr>
<td>Number of firm-practice-guide triad</td>
<td>2940</td>
<td>2940</td>
<td>2940</td>
<td>2940</td>
</tr>
</tbody>
</table>

Clustered-Robust standard errors in parentheses

*** p<0.001, ** p<0.01, * p<0.05, + p<0.1
Table 4. Robustness Checks. Effect on Focal practice’s evaluation.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Model 5 Highest Score</th>
<th>Model 6 GEE Estimates</th>
<th>Model 7 New focal practice at time t+1</th>
</tr>
</thead>
<tbody>
<tr>
<td>avg. evaluation non-focal practice</td>
<td>0.201***</td>
<td>0.272**</td>
<td></td>
</tr>
<tr>
<td>highest evaluation in non-focal practice</td>
<td>0.165***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>std. dev. non-focal practice</td>
<td>-0.064*</td>
<td>-0.197**</td>
<td>-0.467*</td>
</tr>
<tr>
<td>similarity between other practice</td>
<td>-0.036*</td>
<td>-0.022*</td>
<td>-0.853**</td>
</tr>
<tr>
<td>contrast focal practice</td>
<td>0.355***</td>
<td>0.320***</td>
<td>0.146</td>
</tr>
<tr>
<td>tenure in focal practice</td>
<td>0.014***</td>
<td>0.026***</td>
<td></td>
</tr>
<tr>
<td>firm scope</td>
<td>0.056***</td>
<td>0.069***</td>
<td>0.017*</td>
</tr>
<tr>
<td>size</td>
<td>0.008</td>
<td>0.010*</td>
<td>0.221</td>
</tr>
<tr>
<td>internationalization degree</td>
<td>0.003</td>
<td>0.014***</td>
<td>0.345*</td>
</tr>
<tr>
<td>advertisement in guide</td>
<td>0.001</td>
<td>0.007</td>
<td>0.077</td>
</tr>
<tr>
<td>guide dummies</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>practice dummies</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>year dummies</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>location dummies</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>nationality dummies</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>Constant</td>
<td>0.273***</td>
<td>0.073+</td>
<td>0.56</td>
</tr>
<tr>
<td>Observations</td>
<td>16741</td>
<td>16741</td>
<td>1646</td>
</tr>
<tr>
<td>Number of firm-practice-guide triad</td>
<td>2940</td>
<td>2940</td>
<td>253</td>
</tr>
</tbody>
</table>

Clustered-Robust standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05, + p<0.1

4. RESULTS

Table 3 shows fixed-effects regression results on the predicted evaluation in the focal practice. Model 1 includes controls only. Consistent with previous literature on categorization (Negro, Hannan, and Rao, 2010), the contrast of the focal practice has a significant positive impact on the evaluation, meaning that the clarity of practice area in which firms operate
matter in the eye of clients. The tenure of coverage in the focal practice area affects positively the dependent variable ($\beta = 0.015$, $p < .001$), meaning that the longer the organizations have been covered by the directories, the higher their evaluation. I also find that firms with a broader scope (i.e. covered in multiple practice areas) received a better evaluation than focused firms. This result confirms my preliminary qualitative findings that clients favor “full service” law firms, i.e. with a large range of practice areas. Size and internationalization of firms as well as advertisement pages bought in the directories do not significantly impact the focal practice’s evaluation.

In Model 2, I introduce the first independent variable. From Hypothesis 1, based upon the category inferences mechanism, I expect that the higher the average evaluation about a firm’s non-focal practice, the higher will be the evaluation of its focal practice. The results from model 2 support the Hypothesis. The coefficient is positively significant ($\beta = 0.124$, $p < .001$) over all the models and provides evidence that clients rely on past evaluation of firms to assess its current value in another domain.

Model 3 estimates the impact of the standard deviation of non-focal practice areas’ evaluation. The negative and significant coefficient ($\beta = -0.140$, $p < .001$) confirms Hypothesis 2, which posits that the inconsistency of evaluation about a firm’s non-focal practice areas will fade the focal practice’s score. A greater inconsistency of evaluation prevents audience members from inferring firms’ properties from categories to another and acts as a bad signal related to its consistency and future capacities in the focal category.

Hypothesis 3 states that the similarity in terms of members between non-focal practice areas offers poorer information to audience members in order to infer firms’ properties from a source domain to a target domain. The full model provides evidence that redundant information of two similar categories impact negatively the focal category’s evaluation ($\beta = -0.031$, $p < .001$). The coefficients of the first two independent variables remain consistent with the previous models’ results.

As robustness checks, I ran supplementary models with alternative measures of our independent variables. Regarding the first independent variable, instead of taking the average evaluation in non-focal practice areas, I used the highest score obtained by a firm in one of its practice portfolio. I obtained the same results (Table 4, Model 5). Second, to ensure validity of statistical methods, I ran generalized estimating equation (GEE) models that provide robust estimation of standard errors by using the observed variability in the data rather than the variability predicted by an underlying probability model (Negro and Leung, 2013). Results
are similar (Table 4, Model 6) for the variables of interest. Third, to mitigate the reverse causality issue, I reduced the sample to the focal practice areas newly covered at time t for a given firm (N=1646). As such, I looked at the impact of the average and the standard deviation of evaluation, and the similarity of existing non-focal practices at time t-1 on the focal practice area newly included at time t. Results in Model 7 (Table 4) support my hypotheses.

5. DISCUSSION

Insights into categorization mechanisms have improved the understanding of economic and organizational life. Market categories shape what audiences perceive, how they store and retrieve information, and how they organize experience. The operation of categorization consists for audience members to assign a particular organization into a category “so that they can understand and draw inferences about it” (Loken, Barsalou and Joiner, 2008: 133). This paper advances the literature on categories and classification processes by highlighting the role of inference mechanisms. Most research to date has ignored the relationships among the categories being studied. Here, I explored the line of demarcation of category systems exploring evaluation spillovers via inferences. I show that inference-based judgments are grounded on the observation of an organization’s past evaluation in source categories weighted by its homogeneity and its distance with target category. Audiences make a connection between multi-category memberships and are confident with organizations whose they are familiar with and they already know properties. By transferring known features from a source category to the domain of a target category, audiences are able to extend conventional knowledge and instantiating new insights about organizations (Lakoff & Johnson, 1980). As such, audience members collect reliable information by relying on categories which are the basis for imputing knowledge, rights, obligations, and typical activities conventionally associated with these categories.

This study makes a broader contribution to research on category spanning. I support the idea that lower appeal due to cognitive confusion arises for category-spanners that external evaluators rate divergently among categories. The variance of multiple evaluations across various market categories dilutes and weakens the whole social evaluation of firms. In that case, several conflicting perceptions of firms exist and divergently drive inferences and expectations of audiences. It is less the number of category memberships in and of itself that matters than the propensity of audiences to apply category knowledge to make predictions. In this line of research, Loken, Barsalou and Joiner (2008) have shown that audience members
do not focus on the entire information contained in a category, but on a portion of the available knowledge therein. This selection is influenced by both the accessibility of information in memory or in the environment and the relevance of information in achieving specific goals. This selection bias reinforces the importance of being viewed concordantly across categories organizations span.

The study also offers a perspective on category dynamics. While scholarship in this line of thought has amply demonstrated the importance of fitness with cognitive representations, in regularly shaping a variety of economic behaviors and outcomes, there remains many circumstances in which the impact of cognition on organizations and markets is much more than just linear and static. Here I described what factors lead to a better evaluation and how such process, in turn, trigger self-reinforcing dynamics in a given domain.

This paper also provides insight into the relationship between categories and social evaluation. It connects the cognitive foundations of categories with their market outcomes through the prism of social evaluation hierarchies. Social evaluation or status of firms are mainly studied in only one domain or market category in which firms are engaged. To one firm corresponds a single and one-dimensional evaluation. Here, I examine organizations that evolve among various categories at the same time and then receive different levels of social evaluation. To one firm corresponds a plural evaluation. This paper highlights that audiences attribute to a new activity the privilege associated with high past evaluation received in another domain. That said, both within and across various domains, the inertia of the actor’s social position does not result from stable intrinsic differences (in terms of skills for example) but from a self-reinforcing process and self-validating expectations (Gould, 2002).

Finally, this investigation also adds to the literature on diversification by providing a complementary cognitive explanation to the benefits of related diversification (Greenwood et al., 2005; Montgomery and Wernerflet, 1988; Rumelt, 1974). Law firms contain many specializations and exhibit great variety since each internal department displays its own codes. Each practice area (e.g., banking, intellectual property, employment) has its specific routines, rules of working, professional reviews and role models. Managing multiple practice areas in order to level them and enhance their cognitive complementarity is a strategic issue for organizations (Elsbach and Kramer, 1996; Peteraf and Shanley, 1997). In order to signal sharp and consistent category affiliations, law firms have to reconcile and homogenize the ratings of their subunits. They have to bring their multiple category memberships to the same level and forge linkages between them to leverage practice diversification and become a global brand.
The primary goal of this paper is to advance the line of research on the role of categories and classification processes in markets by establishing inferences mechanisms. In this paper, I rely on Ross and Murphy (1996) who showed that there is no cognitive limit or inability to use multiple categories in inferences. Therefore, I make the assumption that audience knows all the category membership of firms and they do not operate any distinction among each category. However, some practice areas of law may be more highly valued than others. Some categories might be more informative than others, based on knowledge of clients, etc. How many categories are used in induction? Does it exist primary/auxiliary categories regarding inference-based mechanisms? More broadly speaking, it would be fruitful to examine whether there are differences in logical competences and capacities of actor to schematize information and make inferences. Audience members can make fallacies and errors due to abusive or too extensional generalization (Robert, 2005: 709). Future research could explore the role of knowledge, expertise, and the influence of social position in inference-based mechanisms.
REFERENCES


