Received: 17 November 2018 | Revised: 8 July 2021 | Accepted: 23 July 2021 | Published on: 4 August 2021

DOI: 10.1002/smj.3328

STRATEGIC MANAGEMENT JOURNAL WILEY

RESEARCH ARTICLE

Income or education? Community-level antecedents of firms' category-spanning activities

Heewon Chae

W. P. Carey School of Business, Arizona State University, Tempe, Arizona, USA

Correspondence
Heewon Chae, W. P. Carey School of Business, Arizona State University, PO Box 874006, Tempe, AZ 85287 USA.
Email: heewon.chae@asu.edu

Abstract

Research Summary: This study investigates the relationships between consumer income, consumer education, and firms' propensities to span multiple market categories. Despite their positive correlation, I theorize contrasting effects of income and education on firms' variety-enhancing spanning. Specifically, I propose that the strong purchasing power of high-income communities should reduce the need for firms to operate in multiple categories, but culturally omnivorous preferences among educated elites should encourage firms' spanning. Analyses of 6,072 restaurants in a metropolitan area and a large-scale survey offer support for these predictions. Education, though not income, has a further positive effect on firms' atypicality-enhancing spanning. This study contributes to category and management research by focusing on audience heterogeneity as important antecedents of firms' action and explicating the multifaceted nature of spanning.

Managerial Summary: This study examines how restaurants decide their culinary categories and menus depending on residents' income and education levels of a city they are located in. I find restaurants in higher-income communities tend to be more specialized in a single or fewer categories while those in lower-income communities are more likely to diversify into multiple categories to reach a broader customer base. By comparison, restaurants in more educated communities tend to diversify into multiple categories and provide fusion food because educated cultural omnivores like to explore novelty. These findings imply that retail firms should consider the separate effects of income and education levels of target consumers in determining their business scope and product portfolio.

KEYWORDS

category spanning, community analysis, demand-side strategy, quantitative analysis, restaurant industry

1 | INTRODUCTION

What factors influence firms to engage in category spanning? Categories are social constructions that represent a meaningful consensus *about* organizations' features as shared by actors and audiences. Categories and category spanning in markets are thriving topics across a number of literatures within management and the sociology of organizations (Cattani, Porac, & Thomas, 2017). The main finding of these studies has been that category spanners—organizations "whose features and actions cause them to be assigned to multiple categories" (Kovács & Hannan, 2010, p. 175)—have lower chances of success and survival due to normative and cognitive reasons (Hsu, Hannan, & Koçak, 2009; Negro & Leung, 2013; Zuckerman, 1999). A new line of recent research, by comparison, has begun to investigate more nuanced effects of spanning, highlighting that the consequences of spanning often depend on context, and spanning may be beneficial in some conditions (Merluzzi & Phillips, 2016; Navis & Glynn, 2010; Pontikes, 2012; Smith & Chae, 2016; Vergne, 2012). A few studies also start to investigate a more fundamental question of what factors influence firms to span market categories in the first place to fully understand the consequences they may bring (Durand & Kremp, 2016; Tang & Wezel, 2015). Still, much of the available research assumes that all audiences, especially those within the same group (e.g., investors, consumers, and critics), hold the same categorical expectations, limiting our understanding of the processes and effects of categorization (Durand & Paolella, 2013).

In this study, I emphasize the role of audience heterogeneity within the consumer group and investigate demand-side antecedents of retail organizations' spanning activities. Consumers are critical organizational audiences as they provide opportunities for firms and impose constraints on them (Adner & Zemsky, 2006; Brandenburger & Stuart Jr, 1996; Priem, 2007; Zander & Zander, 2005). Organizations present offerings to audiences and audiences screen, select, consume, and evaluate them using the categories that aid the evaluation. As such, consumers are the ultimate arbiter of value and organizations must provide offerings that appeal to them to remain competitive (Brandenburger & Stuart Jr, 1996; Priem, Wenzel, & Koch, 2018). Importantly, consumers do not have universal criteria for consuming and evaluating organizational offerings. They have a myriad of different preferences and expectations that are embedded in distinct value systems according to factors such as their economic and social status (Allenby & Rossi, 1998). It is, thus, important to acknowledge such differences as they may translate into contrasting attitudes toward the same offering, providing strategic implications for organizational action.

With this consumer heterogeneity in mind and drawing on prior research on organizational niche width and cultural consumption, I analyze how organizations vary in their category-spanning activities in response to the economic and social status of local communities. Distancing from the conventional approach to treat economic and social status interchangeably, I analyze the independent effects of each status, represented by the degree of community income and education, respectively. In doing so, I also distinguish the two different types of category spanning—that is, variety-enhancing and atypicality-enhancing spanning (Goldberg, Hannan, & Kovács, 2015). *Variety-enhancing spanning* concerns an extent to which an organization has membership in multiple market categories. By comparison, *atypicality-enhancing spanning* refers to an extent to which the organization fuses elements from disparate categories to create differentiated offerings within the boundaries set by its organizational membership. I propose differential effects of consumers' degree of income and education on the two types of spanning. More specifically, I suggest that strong purchasing power of high-income communities should reduce the need for firms to claim membership in multiple market categories, as market demand in such communities is sufficient to sustain many specialist organizations. In contrast, I argue that the strong purchasing power of high-income communities should encourage firms to engage in spanning that yields high atypicality as it lowers consumers' relative cost of exploring unfamiliar hybridized offerings. With regard to education, I suggest that educated elites' omnivorous taste that desires novelty will encourage firms to span market categories to increase both variety and atypicality. These arguments are tested using a sample of 6,072 restaurants in 85 cities located in the greater Washington DC metropolitan statistical area, one of the most socially and economically segregated regions in the United States. Results find support for the contrasting effect of the community income and education levels on restaurants' variety-enhancing spanning—both effects are notable by statistical significance and size—whereas only education is found to have a significant and positive effect on their atypicality-enhancing spanning. In addition, I run a survey to further corroborate the proposed mechanisms by which consumer income and education influence restaurants' category spanning. Results from an analysis of data collected from 517 survey participants are consistent with those from the main analyses.

The intended contributions of this article are threefold. First, this study joins a recent scholarly effort to investigate important antecedents of firms' category spanning. While most of the efforts attend to organizational determinants, this study focuses on demand heterogeneity within the same audience group. By so doing, this study offers a novel approach to studying the antecedents of category spanning, emphasizing the endogenous nature of firm decisions, and helps theorize and test "how organizations think of categories as selection filters via which pertinent audiences attend to, perceive, and judge them" (Durand & Paolella, 2013, p. 17). This study also adds to research on firms' location strategies by explicating differential effects of the economic and social status of local communities on organizations' incentives to straddle categories and aiding them in selecting among available resources they can utilize. Third, by unraveling and measuring the two different types of spanning, this study adds to our understanding of the ways in which spanning leads to variations in outcomes and their evaluation (Kennedy & Fiss, 2013) and contributes to recent academic efforts to consider the multifaceted nature of categories and emphasize the generative capabilities of categories (Glynn & Navis, 2013). More contributions on the broader literature on categories and strategic management as well as limitations of this study are presented in the discussion section.

1.1 | Audience heterogeneity as the antecedent of category spanning

Consumers are a crucial audience in formulating firm strategies because "value creation, by offering benefits that induce payments from willing consumers, is a precondition for value capture" (Priem, 2007, p. 219). As Adner and Zemsky (2006) suggest, focusing on the demand environment along with firm resources and competition is essential to address a fundamental concern in strategy of the fit between a firm and its environment. Research in economic sociology and organizational theory has also demonstrated the importance of organizational audiences in influencing and evaluating firms' spanning decision (Bowers, 2015; Pontikes, 2012). According to this line of research, the notion of audience heterogeneity implies that the same entity may be perceived to have a different value by those who evaluate it. Thus, firms would end up capturing a different amount of value by providing the same offerings to different sets of audiences. Most of these studies, while insightful for organizational research, focus on heterogeneity across the different types of audiences [e.g., peers and critics (Cattani, Ferriani, & Allison, 2014); consumers and venture capitalists (Pontikes, 2012); consumers, peers, and experts (Wijnberg & Gemser, 2000)], and have assumed that members of the same audience are homogeneous, namely share the same basic cultural canons and evaluation criteria in evaluating a particular object.[1] We should not, however, "assume *a priori* the existence of consensus about which features audiences" within the same group "attend to, perceive, and use to make sense of organizational reality" (Durand & Paolella, 2013, p. 16). Several recent studies acknowledge heterogeneity within the same audience group and its influence in performance evaluation. For example, Bowers (2015) shows that the same stock may have different evaluations by analysts depending on their previous coverage of stocks. Similarly, Kacperczyk and Younkin (2017) examine differences in audience receptivity in the music industry.

Building on and extending these studies, I attend to heterogeneity within the same audience group—that is, consumers—according to their economic and social status to investigate how varying demand conditions of local consumer communities affect the way retail firms develop their categorical positions in the market. This heterogeneity within the consumer group is important because their evaluation criteria and preferences may be different from one another (and different from producers') and these differences may yield opportunities and pressures for organizations to select strategic decisions with resources available to them. While a number of consumer characteristics may influence firms' category spanning activities, I focus on the economic and the social status of consumer communities, represented by their income and education levels, as important antecedents of firms' category spanning. The two constructs are remarkably robust determinants of variations in virtually all fields of study including social science [e.g., decision making (Bruine de Bruin, Parker, & Fischhoff, 2007); political behavior (Kluegel, 1990); as well as illness and death (Adler et al., 1994)]. While economic and social status are certainly separate constructs, many studies have often used the two interchangeably or even combined them to create a single construct such as socioeconomic status for theoretical and empirical analysis. This may be partly due to the high correlation between the two constructs. Sociologists and labor economists, for example, have found that one's earning generally increases with his or her level of education because education plays a major role in skill sets for acquiring jobs (Psacharopoulos & Patrinos, 2004). Despite this relationship and the convenience of using the combined construct, it is important to disentangle the effect of income from that of education because the mechanisms by which these two constructs influence individual or

[1] I thank an anonymous reviewer for raising this point.

organizational behaviors differ—with one being economic and the other being sociocognitive—and they may, thus, yield varying consequences. Regarding the topic of this study, consumers' income will influence a firm's category spanning by directly affecting their purchasing power and overall market demand while their education level influences preferences for spanning. I explicate these mechanisms in more detail in the hypothesis development section after comparing below the two types of firms' category spanning, the main constructs for the dependent variables of this study.

1.2 | Two facets of category spanning: Variety and Atypicality

Diverging definitions and measures of category spanning have been used as burgeoning literature started to investigate the construct. One way to understand category spanning is to distinguish its two aspects, variety and atypicality (Goldberg et al., 2015). First, variety-enhancing spanning refers to an extent to which an organization, or an entity, has membership in multiple market categories. Firms should decide in what market domains they will operate their business. Some firms pursue their business in a single market category (e.g., a restaurant offering Mexican food) while others straddle multiple categories (e.g., a restaurant offering Mexican and Italian food). Thus, variety-enhancing spanning is related to decisions about an organization's boundaries and concerns niche width the firm appeals to (Hannan, Pólos, & Carroll, 2007). If an organization spans few market categories, its degree of variety-enhancing spanning will be low, appealing to a narrow set of audience. In contrast, if an organization has membership in multiple market categories, it will appeal to a wider set of audience. The categorical boundaries set by the organization's variety-enhancing spanning provide audiences with prototypical expectations, invoking some representations of exemplary members, and audiences rely on the categorical boundaries to identify and make sense of the organization and its features (Glynn & Navis, 2013; Porac, Thomas, Wilson, Paton, & Kanfer, 1995; Zuckerman, 1999). As such, variety-enhancing spanning emphasizes the cognitive and constraining nature of categories by setting the organizational boundaries within which organizations operate business. Studies that focus on the variety aspect often use organizational membership in socially constructed systems of classification to measure the construct. For example, Pontikes (2012) analyzes the number of market labels that software companies claim to belong to and Carnabuci, Operti, and Kovács (2015) use patent class data to investigate the entry of organizations into technological categories.

By comparison, atypicality-enhancing spanning highlights the generative capabilities of categories (Glynn & Navis, 2013). Atypicality-enhancing spanning refers to an extent by which an organization fuses elements from disparate market categories to create hybridized offerings. Organizations may engage in category spanning by bringing elements from different categories and hybridizing them to create new crossbred products (Hsu, Negro, & Perretti, 2012; Montauti & Wezel, 2016). In this hybridizing process, a newly created product develops its own properties that differentiate itself from typical offerings in the focal category. Because the properties of the product violate existing categorical expectations, it may be more difficult to understand and thus deemed less legitimate (Lounsbury & Glynn, 2001). Its atypicality, however, may cause it to be considered novel (Durand & Paolella, 2013). In a study of the film industry, for example, Hsu et al. (2012) argue that while hybridization generally decreases the appeal of products, films that span genres in innovative ways sometimes produce exceptional success.

So atypicality-enhancing spanning provides organizations with a way to differentiate and innovate their products at risk of being ignored, misunderstood, and devalued.

Thus, the two types of category spanning are distinct constructs from each other. Whereas variety-enhancing spanning determines a firm's overall business boundaries by demonstrating the diversity of market categories in which it operates its business, atypicality-enhancing spanning provides a way to position its products within those boundaries. This distinction implies that even if firms have membership in the same market categories, their product portfolios may look different depending on the extent to which they engage in atypicality-enhancing spanning (e.g., Some Mexican restaurants provide only typical Mexican dishes while other Mexican restaurants offer unique dishes by engaging in atypicality-enhancing spanning). This distinction is also consistent with the idea of Glynn and Navis (2013) that organizations that meet minimal standards for valid category membership "retain considerable discretionary autonomy to differentiate" (p. 1128) their offerings within the category or categories to which they belong in order to achieve optimal distinctiveness. While most previous empirical research on category spanning has focused on only one aspect of spanning over the other or has not considered the conceptual difference between the two types, I consider both types of spanning in formulating hypotheses on the influence of consumers' economic and social status.

2 | HYPOTHESIS DEVELOPMENT

2.1 | Consumers' income, purchasing power, and firms' variety-enhancing spanning

Due to fixed costs, local retail establishments are made available only if there is a sufficient demand that wants and can afford them (Waldfogel, 2008). Consumers' income level influences their purchasing power—their financial ability to buy products and services. Higher household income generally implies greater purchasing power among local residents, and lower household income implies limited purchasing power, all other things being equal (Meltzer & Schuetz, 2012; Porter, 1995). As the overall purchasing power of lower-income communities is smaller than that of higher-income communities, retailers in the lower-income communities often face the small market problem, which occurs when certain firms' revenues are not sufficient to cover their costs due to low demand (Waldfogel, 2008). If an organization located in a low-income community operates its business only in a single market segment, the potential consumer base will be even smaller, serving a limited group of consumers in the community. Thus, to survive, the organization should extend its business to multiple market segments to attract a larger base of consumers.

Organizational niche-width theory supports this prediction. An organization's niche tells "what regions of a social space a producer can exploit" (Hannan, 2010, p. 166). Organizations targeting a wide diversity of consumers by straddling multiple market categories are considered to have broad niches and are labeled "generalists" and those that limit their focus to a narrow region in the social space are labeled "specialists" (Hannan et al., 2007). The relative fitness of generalists versus specialists differs depending on various environmental conditions. Specialists concentrate their capacities on a single market segment to efficiently and reliably perform in the target market niche, while generalists assign their capacities across many different kinds of activities. Under a stable environment, specialists are expected to outperform generalists in areas they both target (Hannan & Freeman, 1989). But generalists are likely to outlast specialists given a highly variable distribution of resources because they spread risk across multiple segments and garner attention from a greater number of consumers (Dobrev, Kim, & Hannan, 2001; Hsu, 2006).

To apply this implication to the focus of this study, demand for each market category in low-income communities is small and unstable. Targeting a specific niche, accordingly, may not enable an organization to garner enough resources to survive. Thus, organizations would be more likely to target a broader niche by claiming multiple category memberships to accumulate greater attention from potential customers. By comparison, the concern about limited purchasing power is relatively insignificant for retail firms in wealthier communities. Generally, higher-income communities have a greater and more stable aggregated demand for each category, and an organization's need to span multiple market categories would thus be smaller. Rather, concentrating on a narrow market niche and devoting its capacities to the focused domain will be a more effective way to appeal to consumers in the given market segment.

> **Hypothesis (H1).** *Firms located in higher-income communities are less likely to engage in variety-enhancing spanning than firms located in lower-income communities.*

2.2 | Education and cultural omnivorousness

Education is an important indicator of one's social status and influences one's perception of categorical boundaries (Bryson, 1996). In the sociology of culture and consumption, a stream of research initiated by Peterson and his colleagues argues that since the 1990s, educated individuals have shown a tendency toward cultural omnivorousness—"an openness to appreciating everything" (Peterson & Kern, 1996, p. 904). Burgeoning empirical evidence has documented a shift in the orientation of highly educated individuals toward an inclusive range of cultural preferences that traverse categorical boundaries (e.g., Goldberg, 2011; Hahl, Zuckerman, & Kim, 2017). With social structural and value changes, greater tolerance of difference and understanding of the coexistence of multiple values have become virtues that social elites are expected to possess. And the elites' inclusive taste and openness to novelty have become "a new cultural logic of distinction" (Goldberg, 2011, p. 1411)—a way of differentiating themselves from others. Holt (1998) also finds, from a series of in-depth interviews in an ethnographic study, that those with high cultural capital (HCC) in the United States who have at least a bachelor's degree "understand their social world much more expansive" than those with low cultural capital (LCC) who have at most a high school education (p. 12). He further finds that while "HCCs and LCCs enjoy variety in their consumption to a greater or lesser extent," "they differ in their subjective understandings of what constitutes variety. What is exotic for LCC is mundane for HCCs, and what is exotic for HCCs in unfathomable or repugnant for LCCs." (p. 13), implying that the breadth of the perceived social world and the openness to diversity and novelty of highly educated groups are more expansive than those of groups with less education.

In line with the view, Merluzzi and Phillips (2016) also suggest that generalists may be better evaluated than specialists when audiences can easily determine the quality of offerings. They argue that strong screening capabilities diminish the advantage of a focused player by minimizing evaluators' uncertainty about the quality of offerings. Compared with less educated audiences, highly educated individuals who have a better ability to evaluate offerings and have openness to diversity may welcome generalists. With the inclusive taste and screening capabilities, thus, highly educated groups are expected to be more open and favorable to organizations that pursue business in multiple market categories. Therefore, I hypothesize:

> **Hypothesis (H2).** *Firms located in more educated communities are more likely to engage in variety-enhancing spanning than firms located in less educated communities.*

2.3 | Atypicality-enhancing spanning: Receptivity to hybridized offerings

Atypicality-enhancing spanning often brings novel and unfamiliar offerings to the market by recombining and hybridizing elements from disparate categories and bridging categorical boundaries (Durand & Paolella, 2013). Many attempts at recombination end in failure and offerings that are poorly appreciated because atypicality comes with unfamiliarity. Highly atypical offerings may be untested and incompletely understood and product definitions may be ambiguous and thus regarded not legitimate (Navis & Glynn, 2010). Which audience, then, would be more likely to take risks and explore untested hybridized offerings?

Sociologists have long argued that wealthy consumers are more likely to explore novel offerings and create trends (Bourdieu, 1984; Simmel, 1957). They are among the first group to try novel offerings in the market and create fashion, and once a trend begins to be followed by the lower classes, they pursue new trends (Simmel, 1957). Importantly, what enables them to do so is their economic affluence. Exploring atypical hybridized offerings is often costly and entails risk-taking because they deviate from prototypical features and may not meet audiences' expectations. Thanks to their wealth, the relative cost of trying hybridized offerings is small for consumers with higher income compared with those who are worse off. Consumers with lower income, in contrast, do not have the luxury to actively pursue hybridized offerings. Considering their tight budget, they will be more likely to make a safer choice to maximize utility by selecting category-typical products that are familiar and conform to their established expectations. Retail firms in lower-income neighborhoods would, thus, be encouraged to provide category-typical offerings to absorb an existing demand for the offerings that conform to established expectations rather than providing atypical offerings that are unfamiliar to target audiences.

The devastating impact of the COVID-19 crisis on the restaurant industry provides descriptive evidence that is consistent with this argument. The restaurant industry got hit hard during the pandemic and ended 2020 with total sales that were $240 billion below the pre-pandemic forecast for the year by the National Restaurant Association (https://restaurant.org/articles/news/new-report-measures-pandemics-effect-on-business). Increase in to-go and delivery sales was not enough to save restaurants and more than 110,000 establishments were closed for business temporarily, or for good, as of December 1, 2020 (https://www.cnbc.com/2020/12/27/covid-pandemic-restaurant-revenue-has-fallen-despite-delivery-boom.html). To overcome sudden decrease in consumer demand and purchasing power, restaurants have streamlined menus to be cost effective. They also provide more typical food as consumers look for comfort food such as burgers and pizza (https://restaurant.org/articles/news/new-report-measures-pandemics-effect-on-business). In line with this change, experts in the restaurant industry in the greater Washington DC metropolitan area—chefs, owners, real estate brokers, accountants—predicted in their recent interview with the Washingtonian magazine that restaurants would provide more fail-safe dishes to survive during and after the COVID-19 crisis.

For example, John Asadoorian, a real estate broker in the DC area, said, "Menus that are meat and potatoes and have broad appeal will be safer than very high-end complicated restaurants."[2] As such, category-typical offerings will have a broader reception than atypical offerings for audiences who are on a tight budget due to lower income. Thus, I hypothesize:

Hypothesis (H3). *Firms located in higher-income communities are more likely to engage in atypicality-enhancing spanning than firms located in lower-income communities.*

I also suggest that organizations targeting highly educated communities will be more likely to engage in atypicality-enhancing spanning for a different reason. As mentioned earlier, the omnivorous taste of social elites influences their perception of categorical boundaries and preferences for hybridized offerings. Research in cultural consumption has demonstrated that the openness of elite omnivores is articulated in several ways, ranging from a passive tolerance of different forms to an active desire to discover new and challenging items (Ollivier, 2008; Roose, van Eijck, & Lievens, 2012). In other words, elite individuals' openness entails both "the capacity and willingness to learn and choose as opposed to the inability or unwillingness to do so" (Ollivier, 2008, p. 125). First, it entails competencies and knowledge that enable educated individuals to understand and tolerate difference. This ability leads elites to impose less strong categorical schemas when evaluating products. For them, categories and categorical schemas are less of a checklist to strictly follow and more of a convenient tool to aid their own comparison and evaluation. As Glynn and Navis (2013) argue, categories are a cultural resource or toolkit (Swidler, 1986) that organizational audiences use to make sense of and evaluate the organization and its offerings. Educated elites will be more competent to *use* categories for making decisions and will not be simply dominated by the categorical imperative. Exposed to a wide range of values and knowledge, the educated exert more agency when selecting and evaluating organizational offerings and make a conscious self-driven decision (Hannerz, 1990; Kacperczyk & Younkin, 2017).

In addition to their ability to better understand hybridized offerings, omnivorousness leads educated elites to develop a preference for such offerings. Research in cultural consumption and marketing science has long found a positive relationship between higher levels of education and inclination to seek intellectual stimulation and novel ideas (Chan & Goldthorpe, 2007; Rogers, 1995). Combined with the tendency to pursue novelty, omnivorousness paves the way for the elites to favor crossbred goods in practical or technical domains (Ollivier, 2008; Roose et al., 2012). They find hedonic value in consuming innovative hybrids (Bianchi, 2002). Holt (1998) also shows that educated elites in the United States seek out and desire "consumption objects far removed conceptually from what is considered to be normative within a category" (p. 13) while those with less education "find comfort in objects that are familiar" (p. 13). As such, while atypicality-enhancing spanning may result in creating confusing products that are difficult to understand among some audiences, certain audiences—such as educated elites—enjoy experiencing and understanding offerings that present a challenge to easy classification.

Hypothesis (H4). *Firms located in more educated communities are more likely to engage in atypicality-enhancing spanning than firms located in less educated communities.*

[2]From "How Washington's Dining Scene Will Change—Explained by Those Who Know It Best" (April 23, 2020) https://www.washingtonian.com/2020/04/23/how-washingtons-dining-scene-will-change-explained-by-those-who-know-it-best/

3 | METHODS

3.1 | Sample and data

I perform quantitative analyses on a large and detailed data set of the restaurant industry in the greater Washington DC metropolitan area. The restaurant industry is large and ubiquitous and touches nearly every household's daily life in the United States (Carroll & Torfason, 2011). The National Restaurant Association estimates restaurant industry sales to be $659 billion in 2020, representing about 4% of the U.S. GDP and employing 10% of the overall U.S. workforce. In addition to its economic and social significance, the restaurant industry is an excellent empirical setting for several reasons. First, the industry contains submarkets with many culinary categories that appear to be broadly understood and schematized (Rao, Monin, & Durand, 2005). Also, the industry provides an opportunity to examine the two different types of category spanning simultaneously because many establishments claim to provide multiple cuisine styles and often create atypical and differentiated offerings within the styles (Kovács & Johnson, 2014). Second, restaurants usually serve a small geographic area (Waldfogel, 2008) and the industry is a field of fierce competition in which many organizations rise and fall. According to Parsa, Self, Njite, and King (2005), approximately 26% of restaurants fail during the first year of operation, so it is imperative for restaurants to reflect the economic conditions and preferences of local demand.

Most information on the sample restaurants is collected from the website Yelp, a major online source of information on local businesses. Yelp has been a frequent source of data for category research because it provides an entire coverage of organizations in sample regions and uses a classification system that is based on social constructions shared by relevant audiences (Zerubavel, 2009). All establishments listed in the "restaurants" group on Yelp were included in the sample. Information on menus is extracted from the website Allmenus, the largest local menu guide in the United States owned by GrubHub, the website gathers restaurant menus from a variety of sources including restaurant websites, restaurants' submissions, and its own team of menu collectors, and the menus are shown in a standard format. The richness of the data on the menus enabled to perform a detailed analysis of the atypicality-enhancing spanning of the restaurants in the focal region. All the information on restaurants was downloaded on February 27, 2014.

The demand-side characteristics are examined at the city level. To avoid reverse causality and endogeneity issues, information on city-level characteristics was collected from the 2008–2012 American Community Survey (ACS) data by the Census Bureau. In the case of retail business, customers are drawn primarily from the immediate vicinity, and establishments thus mostly likely reflect the characteristics of neighborhood residents (DiPasquale & Wheaton, 1996). The greater Washington DC metropolitan area, composed of the District of Columbia and municipalities in 22 counties in Virginia, Maryland, and West Virginia, is one of the most economically and socially segregated metro areas in the United States (5th most segregated metropolitan area from the 2008–2012 ACS); consequently, the setting provides enough variance across cities for analysis. The final sample includes 6,072 restaurants in 85 cities.

3.2 | Dependent variable

3.2.1 | Variety-enhancing spanning

Variety-enhancing spanning is defined as the extent to which an organization has membership in multiple market categories. All restaurants listed on Yelp are classified into one or more of

108 categories.[3] Some categories concern culinary origins such as American and French. Others refer to key dishes such as pizza and chicken wings. And still others are about food codes such as vegan and kosher. Following previous research, I use all the categories as they are shown on the website to measure restaurants' degree of variety-enhancing spanning because these categories are the actual heuristics that restaurants and (potential) customers rely on to identify and evaluate the restaurants (Goldberg et al., 2015; Kovács & Johnson, 2014). In other words, these categories are typically the first information consumers check about the restaurants; as a result, they work as salient market categories used as a classificatory schema in the market. About half of the sample belongs to one cuisine style (56%), 28% of the restaurants are in two styles, and the rest belong to three or more cuisine styles (16%). Because some categories share more attributes and are considered closer than others, I take into account the distance between the categories a restaurant spans when measuring the degree of variety-enhancing spanning, following Kovács and Hannan (2010) and Chang, Kim, and Chae (2020);

$$\text{Variety}-\text{enhancing spanning} = \begin{cases} 0 & \text{if } numcate = 1; \\ numcate * \overline{d}x & \text{if } numcate > 1 \end{cases}$$

where, *numcate* denotes the number of categories the organization has membership and $\overline{d}x$ denotes the average distance between the categories spanned. The average distance between the categories was calculated using the cosine distance between each pair of categories on a co-occurrence matrix, which shows what categories are claimed together by restaurants in the sample. Including cosine distance between categories allows the similarity structure of the spanned categories to be captured (Leahey, Beckman, & Stanko, 2017) and thus enables to distinguish the degree of spanning between restaurants that span the same number of different categories.

3.2.2 | Atypicality-enhancing spanning

While I follow previous literature to measure the degree of variety-enhancing spanning, I create a measure of the degree of atypicality-enhancing spanning by calculating average product typicality of each restaurant to its claimed category (or categories). Previous studies using typicality to measure the degree of category spanning have shown that typicality to a category decreases as category spanning increases (Kovács & Johnson, 2014; Smith, 2011). To establish the product typicality of restaurants, I analyze every word used in the menus of all the restaurants in the sample and used word-category collocation mapping, a commonly utilized computational linguistics approach.

[3]The categories of restaurants shown on Yelp are determined jointly by restaurants and consumer users. Most restaurants manage their Yelp page including the category or categories they claim and belong to. And if Yelp users do not agree with the categorization that is shown, they can report it to Yelp. Once a majority of Yelp users do not agree with the categorization, the website changes the categories accordingly. In addition to the direct report to Yelp, Yelp automatically analyzes reviewers' comments and other restaurant information to correct the categories assigned to better represent category membership of restaurants that is shared by the restaurants and users.

$$\text{Atypicality} - \text{enhancing spanning of a restaurant}$$
$$= 1 - \text{average of the highest dish typicalities in the claimed categories}$$

First, I calculate the typicality of each of the words used in dish names and dish descriptions in each of the 108 categories by computing the Jaccard similarity of the word to the category (Kovács & Johnson, 2014). I exclude all prepositions, conjunctions, and interjections. This left 12,356 unique words in 108 categories. These numbers are comparable to 12,323 unique words in 91 categories, the numbers for the sample of restaurants in San Francisco reported by Kovács and Johnson (2014). In the next step, I calculate dish typicality in the claimed category by taking the average of the word typicality for all words used to describe each dish. Then I average these scores to calculate the average dish typicality of a restaurant. In cases of restaurants that claim multiple categories, I choose the highest score among dish typicalities in the claimed categories for each dish to get an average dish typicality of a restaurant. Finally, because typicality and category spanning are inversely related, I subtract the average dish typicality of a restaurant from 1 to get a degree of atypicality-enhancing spanning for a restaurant. Because the values of from the process are low in absolute number, I rescale the values so that the maximum observed value of atypicality-enhancing spanning is 1 and the minimum is 0 for better interpretability. See Appendix for a detailed illustration of how atypicality-enhancing spanning is calculated and the online Appendix S1 for a correspondence analysis on a selected set of words to show how typical those words are in some of the categories. Figure 1 is a scatterplot of the variety-enhancing and atypicality-enhancing spanning of the sample. As expected, there is a variance in atypicality-enhancing spanning for a given degree of variety-enhancing spanning, indicating that organizations in the same category differentiate their offerings.

3.2.3 | Independent variables

The first independent variable is the community income level. It is measured as five-year estimates of median household income, in tens of thousands of dollars, of each city between 2008 and 2012. Second, the degree of education is measured as the proportion of the population aged

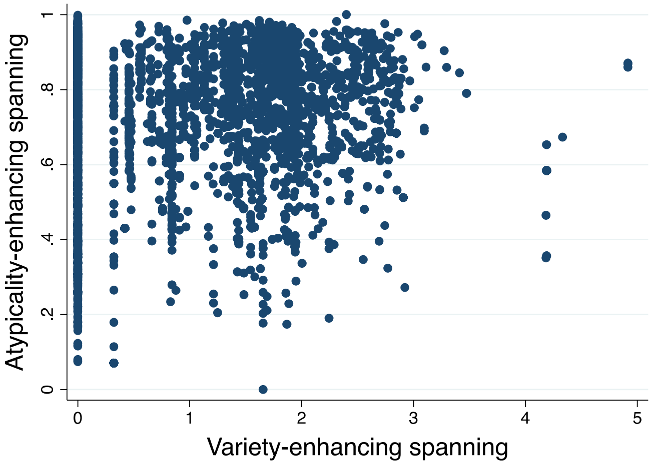

FIGURE 1 Scatterplot of variety-enhancing and atypicality-enhancing spanning by restaurants (n = 6,072)

over 25 who have at least a bachelor's degree, averaged across 2008 through 2012. In an additional set of analysis to investigate the effect of education as well as to address a multicollinearity issue between income and education, I split the proportion of the people with at least a bachelor's degree into those with a bachelor's, master's, and Ph.D. degree, respectively, as their highest degree received. Importantly, results of regressing income on the proportions of the three different levels of degree holders indicate that the proportions of bachelor's and master's degree holders have a positive and significant relationship with income ($p = .006$ and $p = .037$, respectively) whereas the proportion of Ph.D. degree holders does not have a significant relationship with income ($p = .743$). In other words, the highest level of education is not conducive to the highest level of income. This non-linear relationship between income and education and the additional analyses using the separate bachelor's, master's, and Ph.D. degree variables address a concern for multicollinearity. In addition, it gives support to the idea that the effect of the two variables be theorized and analyzed separately.

3.2.4 | Control variables

A number of control variables are included for analysis. First, to investigate the effects of the demand-side characteristics, it is important to isolate the effects from supply side explanations. Thus, I employ several organizational level variables as controls. First, I control for the price level of restaurants. Yelp classify the price level of a restaurant into four groups, indicating "the approximate cost per person for a meal, including one drink, tax, and tips": $ denotes "under US$10," $$ denotes "US$11–US$30," $$$ denotes "US$30–US$61," and $$$$ denotes "above US $61." In my sample, 43% of the restaurants are in the lowest price group, 51% are in the $$ group, 5% are in the $$$ group, and the remaining 1% are in the $$$$ group. To allow for the nonlinear effect of price, I include dummy variables for all levels of the price range except $, which is used as a baseline. I also include a dummy variable for fast food and chain restaurants to account for their unique business style (Carroll & Torfason, 2011). The age of the restaurant is also included, which I measure with the number of years since the first review of the restaurant, following Kovács and Johnson (2014). Next, one important assumption of this study is that consumers are residents of the town in which the restaurant is located. Previous literature has argued that customers of restaurants tend to be local residents (Waldfogel, 2008). However, there may also be other clientele, such as commuters. To control for this issue, I include dummy variables for restaurants that were closed before 7 p.m. or on Sundays, based on the presumption that the main customer base for these restaurants is not local residents but commuters from other towns.[4] Finally, to control for category-specific effects, I include 107 dummy variables for the categories the restaurant is in and use Austrian as a baseline.

At the city level, the ethnic diversity of the town is included because it may influence demand for each cuisine style and for spanning. It is calculated as the Blau index of races in each town (Blau, 1977). Second, as mentioned earlier, an important assumption of this study is that restaurant clientele are residents of the town in which the restaurant is located. In addition to including dummy variables for restaurants that were closed before 7 pm and on Sundays, I add two more control variables. Due to the central location of Washington, DC in the metropolitan area, its restaurants serve a large number of commuters and tourists as well as residents.

[4]Dropping these restaurants from the sample does not change the significance of the results. Results are available from the author.

Thus, to control for its specificity, I include a dummy variable for Washington, DC. By the same token, I also include a dummy variable for central counties, as designated by the Census Bureau.[5] Next, I include the population density (thousand people/mile^2) and the total number of restaurants in the city to control for the effect of overall demand and industry competition in the region. The number of restaurants in the city in the category (or categories) to which the focal restaurant belongs is also included as well as its squared term to take into account the possible curvilinear effects of the category population (Hannan & Carroll, 1992).

In addition, the proportion of the population aged 20–29 in a city is controlled for based on the findings that young adults tend to engage in high risk-taking behaviors and demonstrate an openness to diversity and novelty, which may influence restaurants' likelihood of engaging in spanning (Gardner & Steinberg, 2005). Finally, I include median housing rent of the city as a proxy for overhead costs of restaurants in the focal town. Descriptive statistics and correlations for major study variables are provided in Table 1.

3.2.5 | Analysis

I test my hypotheses using the generalized linear model (GLM) framework (McCullagh & Nelder, 1989). The first dependent variable, variety-enhancing spanning, is nonnegative outcome and has many zero values. To account for the skewedness of the variable, I use the GLM model with a log link (Hardin & Hilbe, 2012). The second dependent variable, atypicality-enhancing spanning, has a minimum value of 0 and a maximum value of 1. I apply the fractional logit model (Papke & Wooldridge, 1996) using a GLM with a binomial distribution and a logit link function. I estimate both models using robust standard errors and cluster observations by city to account for possible heterogeneity among cities that is not explained by the city-level control variables in the models.

4 | RESULTS

4.1 | Variety-enhancing spanning

Models 1–8 in Table 2 shows the results of the regression models predicting restaurants' variety-enhancing spanning. In Model 1, only control variables are included. Models 2 and 3 are partial models, introducing the income level and the education level, respectively. While the coefficient for education is positive and marginally significant in Model 3, the coefficient for income in Model 2 is not significant though it is negative as predicted in H1. I speculate that the lack of significance is due to the omission of the education variable, which is fairly highly correlated with the income variable (0.68) and hypothesized to have the opposite effect on spanning. Without controlling for the effect of the other, it is difficult to estimate the effect of income reliably in Model 2. Model 4 shows a result of a full model where both independent variables are included. The coefficient for income does become negative and significant (b = −0.0247, p = .017, two-tailed) while the coefficient for education stays positive and significant

[5]According to the Census Bureau, central counties in metropolitan statistical areas are defined as counties that (a) have at least 50% of their population in urban areas of at least 10,000 population; or (b) have within their boundaries a population of at least 5,000 located in a single urban area of at least 10,000 population.

TABLE 1 Descriptive statistics and correlations ($n = 6{,}072$)

	Mean	SD	Min	Max	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
1. Atypicality	0.66	0.20	0	1	1																		
2. Variety	0.66	0.89	0	4.92	0.24	1																	
3. Income (in \$10,000 s)	8.75	2.87	3.73	25.00	−0.04	−0.02	1																
4. Education	0.52	0.15	0.13	0.93	0.12	0.06	0.68	1															
5. Ethnic diversity	0.59	0.11	0.10	0.76	−0.02	0.02	−0.45	−0.28	1														
6. Population density (in 1,000 s/miles^2)	5.92	2.49	0.10	8.50	0.19	0.07	−0.39	0.23	0.22	1													
7. Total # of orgs. (in 100 s)	5.46	6.33	0.06	16.08	0.24	0.10	−0.46	0.06	0.16	0.67	1												
8. # of orgs. in category	46.91	72.68	1	493	0.10	0.29	−0.34	0.03	0.12	0.48	0.72	1											
9. % of population aged 20–29	0.17	0.05	0.04	0.37	0.13	0.07	−0.64	−0.12	0.29	0.57	0.54	0.39	1										
10. DC	0.25	0.43	0	1	0.22	0.09	−0.47	−0.05	0.15	0.58	0.98	0.70	0.46	1									
11. Central county	0.93	0.25	0	1	0.08	0.03	0.20	0.39	0.14	0.32	0.19	0.13	0.09	0.16	1								
12. Median rent (in \$1,000 s)	1.44	0.27	0.63	2.00	−0.07	−0.01	0.81	0.55	−0.18	−0.32	−0.48	−0.35	−0.49	−0.52	0.28	1							
13. Org. age (in years)	4.53	2.35	0	9.54	0.08	0.06	0.09	0.21	0.01	0.13	0.10	0.06	0.03	0.07	0.12	0.07	1						
14. Chain	0.61	0.49	0	1	−0.35	−0.07	0.03	−0.18	−0.03	−0.26	−0.27	−0.07	−0.15	−0.24	−0.13	0.05	−0.16	1					
15. Fast food	0.16	0.37	0	1	−0.13	−0.20	−0.03	−0.17	−0.02	−0.16	−0.16	−0.12	−0.07	−0.13	−0.09	−0.02	−0.15	0.28	1				
16. Price==“\$\$”	0.51	0.50	0	1	−0.07	0.12	0.08	0.09	−0.01	0.00	−0.02	−0.04	−0.03	−0.03	0.01	0.07	0.13	−0.10	−0.44	1			
17. Price==“\$\$\$”	0.05	0.22	0	1	0.19	0.04	0.00	0.05	−0.03	0.05	0.11	0.02	0.03	0.11	0.01	−0.04	0.05	−0.19	−0.10	−0.23	1		
18. Price==“\$\$\$\$”	0.01	0.09	0	1	0.09	0.03	0.01	0.03	−0.02	0.02	0.04	0.00	0.01	0.04	0.00	−0.01	0.07	−0.02	−0.04	−0.09	−0.02	1	
19. Closed on Sundays	0.06	0.24	0	1	0.12	0.06	−0.06	0.05	0.01	0.12	0.19	0.17	0.11	0.18	0.04	−0.07	0.01	−0.10	−0.05	−0.09	0.05	0.06	1
20. Closed before 7 pm	0.05	0.21	0	1	0.11	0.08	−0.06	0.04	0.00	0.13	0.20	0.25	0.09	0.19	0.05	−0.07	0.01	−0.06	−0.05	−0.10	−0.03	−0.01	0.41

TABLE 2 GLM regression predicting variety-enhancing and atypicality-enhancing spanning

	Model1	M2	M3	M4	M5	M6	M7	M8	M9	M10	M11	M12	M13	M14	M15	M16
	Variety-enhancing spanning								Atypicality-enhancing spanning							
Income		−0.0037		−0.0247	−0.0189	−0.0322		−0.0146		0.016		0.0005	−0.003	0.0001		0.0102
		(.587)		(.017)	(.058)	(.008)		(.171)		(.005)		(.956)	(.767)	(.993)		(.145)
Education			0.1826	0.4269	0.388	0.4556		–			0.3129	0.3078	0.3064	0.2755		
			(.074)	(.004)	(.011)	(.005)		–			(.004)	(.062)	(.083)	(.123)		
Neighbor income (10 miles radius)						−0.0033		–						0.0133		
						(.867)								(.390)		
Neighbor education (10 miles radius)						−0.0767		–						−0.1232		
						(.810)								(.638)		
# of orgs. in 10 miles radius						0.0038		–						0.003		
						(.060)		–						(.169)		
Bachelor's							−0.1318	−0.1357							0.0205	0.0139
							(.749)	(.735)							(.955)	(.969)
Master's							0.0977	0.4539							0.2866	0.0517
							(.811)	(.325)							(.399)	(.881)
Ph.D.							1.3033	1.2079							1.2509	1.2946
							(.026)	(.020)							(.062)	(.041)
Variety-enhancing spanning									0.1913	0.1913	0.1895	0.1896	0.1561	0.1962	0.1883	0.1887
									(.000)	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)
Ethnic diversity	−0.1898	−0.2194	−0.113	−0.21	−0.2033	−0.0874	−0.1055	−0.1682	−0.2288	−0.1203	−0.1325	−0.1306	−0.1293	−0.0446	−0.1417	−0.1007
	(.031)	(.031)	(.263)	(.045)	(.057)	(.506)	(.263)	(.103)	(.028)	(.273)	(.110)	(.177)	(.213)	(.714)	(.095)	(.279)
Population density	−0.0005	−0.0009	−0.0027	−0.0085	−0.0052	−0.0168	−0.009	−0.0124	−0.0017	0.0002	−0.0044	−0.0043	−0.0083	−0.0092	−0.0064	−0.0042
	(.938)	(.894)	(.678)	(.260)	(.460)	(.078)	(.137)	(.039)	(.774)	(.969)	(.422)	(.481)	(.183)	(.154)	(.223)	(.457)
Total # of orgs.	−0.0491	−0.0483	−0.0563	−0.0601	−0.0705	−0.0679	−0.0531	−0.0545	0.0657	0.0612	0.0503	0.0504	0.057	0.0478	0.0605	0.0614
	(.000)	(.000)	(.000)	(.000)	(.003)	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)

TABLE 2 (Continued)

	Model1	M2	M3	M4	M5	M6	M7	M8	M9	M10	M11	M12	M13	M14	M15	M16
	Variety-enhancing spanning								Atypicality-enhancing spanning							
# of orgs. in same category	0.004	0.004	0.004	0.004	0.0074	0.004	0.0057	0.0057	−0.0037	−0.0037	−0.0037	−0.0037	−0.0089	−0.0039	−0.0038	−0.0037
	(.000)	(.000)	(.000)	(.000)	(.001)	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)	(.001)	(.000)	(.000)	(.000)
Squared # of orgs. in same category(-ies)	−9.8E-06	−9.8E-06	−9.8E-06	−9.7E-06	−9.8E-06	−9.7E-06	−1.2E-05	−1.3E-05	7.35E-06	7.36E-06	7.41E-06	7.41E-06	0.0001	7.61E-06	7.44E-06	7.43E-06
	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)
% of age in 20s	0.4364	0.3725	0.5294	0.2238	0.4401	−0.1507	0.5321	0.3125	−0.1552	0.1448	0.052	0.0583	0.034	−0.0304	−0.0154	0.1426
	(.043)	(.120)	(.005)	(.317)	(.040)	(.550)	(.013)	(.217)	(.521)	(.499)	(.747)	(.759)	(.881)	(.880)	(.921)	(.438)
DC	0.4936	0.4816	0.5886	0.633		0.6896	0.4409	0.4567	−0.4388	−0.3763	−0.2382	−0.2394		−0.2188	−0.3814	−0.3915
	(.000)	(.001)	(.000)	(.000)		(.001)	(.018)	(.016)	0.000	(.001)	(.023)	(.031)		(.093)	(.002)	(.002)
Central county	−0.0075	−0.0027	−0.0219	−0.0073	−0.0298	−0.0566	0.007	0.0168	0.0407	0.0188	0.0138	0.0136	0.0116	−0.0141	0.0302	0.0234
	(.850)	(.948)	(.604)	(.872)	(.481)	(.190)	(.890)	(.745)	(.227)	(.570)	(.688)	(.694)	(.727)	(.747)	(.380)	(.499)
Median rent	0.0578	0.0785	0.0115	0.087	0.0816	0.0857	0.0169	0.0674	0.0748	−0.0118	0.0011	−0.0005	0.0111	−0.0311	0.0023	−0.0306
	(.220)	(.206)	(.808)	(.123)	(.104)	(.164)	(.787)	(.439)	(.156)	(.822)	(.981)	(.992)	(.844)	(.563)	(.964)	(.548)
Org. age	0.0006	0.0007	0.0002	0.0001	0.0029	0.0002	0.0035	0.0036	−0.001	−0.0018	−0.0025	−0.0025	0.0092	−0.0037	−0.0023	−0.0024
	(.890)	(.877)	(.963)	(.973)	(.633)	(.959)	(.362)	(.349)	(.919)	(.861)	(.806)	(.806)	(.229)	(.715)	(.826)	(.816)
Chain	−0.2237	−0.2244	−0.2208	−0.2218	−0.188	−0.2232	−0.2992	−0.3007	−0.4322	−0.43	−0.4288	−0.4288	−0.3943	−0.4215	−0.4289	−0.4282
	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)
Fast food	−1.0209	−1.0212	−1.0173	−1.0147	−0.994	−1.0069	−1.0805	−1.0814	−0.2299	−0.2274	−0.2261	−0.2261	−0.2857	−0.2145	−0.226	−0.2251
	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)	(.001)	(.000)	(.000)
Price = "$$"	0.1803	0.1807	0.1788	0.1797	0.3017	0.1878	0.1203	0.1206	−0.1801	−0.1814	−0.1814	−0.1814	−0.1458	−0.1777	−0.1798	−0.1804
	(.015)	(.015)	(.015)	(.015)	(.000)	(.015)	(.001)	(.001)	(.000)	(.000)	(.000)	(.000)	(.002)	(.000)	(.000)	(.000)
Price = "$$$"	0.1408	0.1418	0.1368	0.1383	0.2003	0.1567	0.0243	0.0264	0.086	0.0832	0.0808	0.0808	0.0804	0.0852	0.0824	0.0809
	(.231)	(.229)	(.244)	(.238)	(.243)	(.192)	(.818)	(.803)	(.013)	(.017)	(.019)	(.019)	(.144)	(.017)	(.016)	(.018)
Price = "$$$$"	0.3801	0.3816	0.3796	0.3881	0.3053	0.3962	0.1778	0.1839	0.4232	0.414	0.4146	0.4144	0.511	0.4057	0.4156	0.4112
	(.000)	(.000)	(.000)	(.000)	(.033)	(.000)	(.067)	(.055)	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)

TABLE 2 (Continued)

	Model1	M2	M3	M4	M5	M6	M7	M8	M9	M10	M11	M12	M13	M14	M15	M16
	Variety-enhancing spanning								**Atypicality-enhancing spanning**							
Closed on Sundays	0.0469	0.0473	0.0443	0.0438	0.0611	0.0442	0.1304	0.1307	−0.0703	−0.0735	−0.0751	−0.0751	−0.1067	−0.0777	−0.0739	−0.0748
	(.181)	(.175)	(.194)	(.196)	(.307)	(.207)	(.022)	(.022)	(.020)	(.020)	(.018)	(.018)	(.007)	(.025)	(.018)	(.019)
Closed before 7 pm	0.0474	0.0473	0.0483	0.0491	−0.0234	0.0476	0.1021	0.1028	0.2443	0.245	0.2458	0.2458	0.18	0.2556	0.2443	0.2443
	(.169)	(.170)	(.157)	(.151)	(.719)	(.195)	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)
Constant	−2.7414	−2.7147	−2.7898	−2.6725	−2.9747	−2.512	−2.3384	−2.2511	1.133	1.0225	1.0605	1.0581	0.9963	1.0322	1.1138	1.0521
	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)
Fixed effects, 108 categories	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of cities	85	85	85	85	84	69	79	85	85	85	85	85	84	69	79	85
N	6,072	6,072	6,072	6,072	4,534	5,592	5,856	6,072	6,072	6,072	6,072	6,072	4,534	5,856	6,072	6,072
Log-pseudo-likelihood	−4,194.43	−4,194.40	−4,194.05	−4,193.33	−2,944.51	−4,065.38	−4,397.08	−4,396.80	−2,504.84	−2,504.60	−2,504.37	−2,504.37	−1918.08	−2,411.60	−2,504.26	−2,504.20

Note: Robust standard errors are clustered at the city level and *p*-values are reported in parentheses. Two-tailed tests.

(b = 0.4269, p = .004). While the results may support H1 and H2, it is necessary to check whether or not multicollinearity is driving the results, given the high correlation between the two variables. Variance inflation factor (VIF) scores for the income and the education variables after an OLS regression of the full model are 10.98 and 6.08, respectively, relieving some concerns for high multicollinearity. While the multicollinearity may still exist, the presence of high multicollinearity makes it more difficult to reject the null hypothesis as confidence intervals for coefficients becomes wider (Belsley, Kuh, & Welsch, 2005). To further address the possible multicollinearity problem, I split the education variable into three separate variables—the proportion of people with bachelor's, master's, and Ph.D. degree as their highest degree received, respectively. Results of the analysis are consistent with the main model and are described in detail in the next section.

In terms of magnitude change, the results of postestimation margins analysis using the full model demonstrate a 6.8% decrease in the predicted mean of variety-enhancing spanning if median household income rises by one standard deviation from the mean. By comparison, the model indicates that the predicted mean of spanning will increase by 6.6% if the education variable rises by one standard deviation from the mean. Figure 2 visualizes this contrast effect of income and education on variety-enhancing spanning.

4.1.1 | Robustness checks

I perform several additional analyses for robustness checks of the effects. In the main models, I included a dummy variable for restaurants located in Washington, DC to control for the locational specificity due to a large number of commuters and tourists. Acknowledging that the dummy variable may be a rough control, I drop, in Model 5, all of the 1,538 restaurants in Washington, DC from the sample to obtain a more conservative analysis. The coefficients of both the income and the education variable stay significant (p = .058, p = .011, respectively), supporting H1 and H2.

Second, in Model 6, I include the average levels of income and education of neighboring cities that are located within 10 miles from a sample restaurant's longitude and latitude to better isolate the effect of the local community in which the restaurant is embedded. Also, in the model, I replace the total number of restaurants in the focal city with the total number of restaurants that are located within 10 miles to reflect the influence of geographic proximity rather than administrative demarcation. Restaurants that do not have neighboring cities within

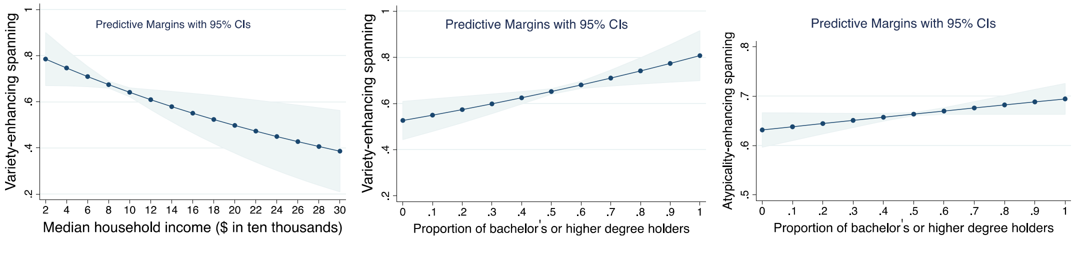

FIGURE 2 Postestimation margins analyses of the effects on variety-enhancing and atypicality-enhancing spanning. Dashed lines represent the average income and education level of the 85 cities, respectively. Stars indicate the average income and education level of the cities in each of the four quadrants, respectively

10 miles are dropped in the analyses. The results of the effect of residents' income and education stay significant ($p = .008$, $p = .005$, respectively), consistent with the previous models.

Third, to conduct a more deliberate analysis of the effect of education as well as to address the potential multicollinearity issue, in Models 7 and 8, I remove the education variable (i.e., the aggregated proportion of bachelor's or higher degree holders) and separately include the proportion of people with bachelor's, master's, and Ph.D. degree as their highest degree received, respectively. As mentioned earlier, the proportions of bachelor's and master's degree holders as their highest degree received have a positive and significant relationship with income whereas the proportion of Ph.D. degree holders does not. Thus, the analysis of the Ph. D. *variable* may help isolate the effect of education from that of income that may have not been fully separated in the main education variable due to their high correlation. In both Models 7 and 8, only the coefficient for the Ph.*D. variable* is positive and significant ($p = .026$, $p = .020$, respectively), indicating that restaurants located in highly educated communities are more likely to engage in variety-enhancing spanning than those located in less educated communities. The coefficient income variable in Model 8 stays negative although it is not significant ($b = -0.0146$, $p = .171$).

4.2 | Atypicality-enhancing spanning

Models 9–16 reports the findings of the regression models predicting restaurants' atypicality-enhancing spanning. The coefficient of income was significant and positive in the partial model (Model 10) but became insignificant in the full model (Model 12). Whereas we cannot reject the null hypothesis that income is not a significant factor that increases atypicality-enhancing spanning, the effect of education stays significant in the partial and the full models, supporting H4. This result supports my argument that highly educated consumer groups are more likely to develop preferences for hybridized offerings, which encourages restaurants to engage in atypicality-enhancing spanning. Though, the magnitude of the effect of education on atypicality-enhancing spanning is not as big as that on variety-enhancing spanning. The full model indicates that the predicted mean of atypicality-enhancing spanning will increase by 1.4% if the education variable rises by one standard deviation from the mean.

4.2.1 | Robustness checks

I took the same approaches applied earlier to check the robustness of the models predicting atypicality-enhancing spanning. First, Model 13 reports the regression result that excludes restaurants located in Washington DC. Like in Model 12, the coefficient of income is not significant while that of education stays positive and marginally significant ($p = .083$). Second, the result does not change much in Model 14 that adds the income and education levels of neighboring cities within 10 miles and the number of restaurants in the neighboring cities. Third, Models 15 and 16 demonstrate the results of the regressions that include the three separate education variables. In both models, the coefficient for only the Ph.D. *variable* is positive and significant ($p = .062$, $p = .041$, respectively). Overall, these results support H4 on the effect of education but not H3 on the effect of income.

Several control variables are worth noting. First, as expected, the coefficient estimates of chain and fast food restaurants are negative and significant in all of the models that predict both types of spanning. These negative effects suggest that chain and fast food restaurants offer more category-typical products in a given category with standardized processes and menus (Carroll & Torfason, 2011). Next, price is found to have a nonlinear effect on category spanning. Restaurants with the highest price level show the highest degree of spanning in both types. This result is consistent with the findings of previous literature that high-status producers, who usually charge high price, create and introduce novel offerings to the field by fusing disparate elements together (Rao et al., 2005). The result, however, also suggests that atypicality-enhancing spanning may occur at relatively low cost as well; restaurants that provide the cheapest offerings (Price = "$") are more likely to engage in atypicality-enhancing spanning than those who provide moderately priced offerings (Price = "$$"). Third, whereas the total number of organizations has a significant and negative effect on restaurants' variety-enhancing spanning (Model 4), it has positive and significant effect on restaurants' atypicality-enhancing spanning (Model 12). This result suggests that increased competition creates pressure for restaurants to differentiate themselves within a narrow market domain to stand out from numerous competitors. Furthermore, the number of organizations in the same categories was found to have a curvilinear relationship with spanning, consistent with the findings from population ecology (Hannan & Freeman, 1989).

As a side note, in another supplementary analysis, the total number of cuisine categories offered by all of the restaurants in a city shows a positive relationship with the level of education of the city ($b = 81.6$, $p = .000$), after controlling for the income level. This finding is consistent with the suggested mechanism of the cultural omnivorousness of educated social elites—that is, social elites express openness to diversity. The income level has a significant and negative relationship with the total number of categories but the coefficient converges to zero ($b = -0.0003$, $p = .001$).

4.3 | Combined effects of income and education

Figure 3 illustrates the income and education levels of 85 cities in the sample. Each node represents a city, and the gray dashed vertical and horizontal lines indicate the average level of income and education, respectively. As income and education have a high correlation, more cities are located in the first and the third quadrant. Still, 18.8% of the cities are located in the off-diagonal quadrants (i.e., 11 cities in the second and 5 cities in the fourth quadrant), demonstrating that cities with above-average income and below-average education levels, and vice versa, exist.

The four stars represent the average income and education level of cities located in each quadrant, respectively, demarcated by the dashed lines. This illustration is made to give an idea about the general application of the estimated combined effects of income and education on restaurants' variety-enhancing and atypicality-enhancing spanning. H1 and H2 suggest that, when controlling for other variables, the degree of variety-enhancing spanning will be the lowest for restaurants located in the fourth quadrant where income levels are high and education levels are low, and the highest for those located in the second quadrant where income levels are low and education levels are high. The results of postestimation margins analysis on the four star points using the full model (Model 4) are consistent with the prediction; a variety-enhancing spanning score is the lowest for the star in the fourth quadrant (0.577) and the highest in the second quadrant (0.671), which is 16.3% greater than the lowest value.

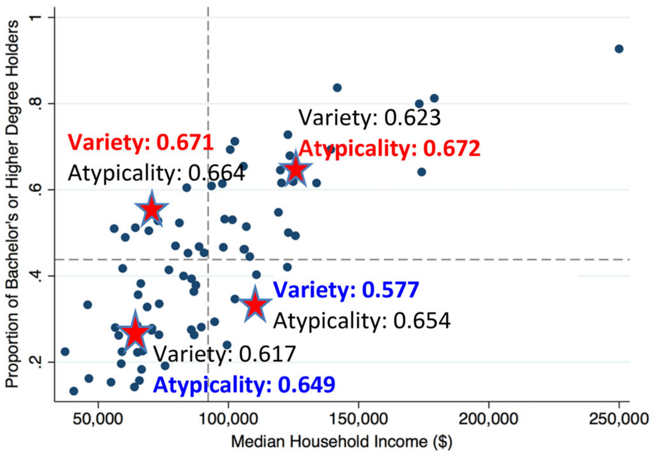

FIGURE 3 Income and education levels of cities ($N = 85$) and estimated spanning scores

On atypicality-enhancing spanning, H3 and H4 suggest that restaurants in the third quadrant will be least likely and those in the first quadrant are expected to most actively engage in atypicality-enhancing spanning. The results of postestimation analysis again are consistent with the prediction although the change of magnitude is smaller; an atypicality-enhancing spanning score is the lowest for the star in the third quadrant (0.649) and the highest in the first quadrant (0.672), 3.5% greater than the lowest value. More strategic implications of the combined effects with some example restaurants can be found in the online Appendix S1.

4.4 | Getting into mechanisms: A survey

In addition to the analyses using the large-scale data, I conducted a survey about consumers' preferences for restaurants to provide descriptive evidence and to better understand the mechanisms by which income and education influence restaurant's spanning. I recruited a total of 627 participants using Amazon's Mechanical Turk, an online crowdsourcing marketplace that allows requesters to post tasks for workers to complete for a fee. To check internal and external validity, I recruited 321 participants from the Washington DC metropolitan area and 306 participants in other areas in the United States. Twenty-four participants who did not complete all required tasks were eliminated from the final sample. I further culled the data according to a handful of attention checks, which will be described in detail below.[6] This reduction leads to a final sample consisting of 517 participants. Results from analysis using the entire responses are consistent with those presented here. Participants' demographic information is summarized in Table 3. All demographic information was collected after participants had completed the survey.

[6]In addition to attention checks described henceforth, participants were asked to read the following during the survey: "In order to facilitate our research on decision making we are interested in knowing certain factors about you, the decision maker. We are interested in whether you actually take the time to read the directions; if not, some of our questions will be unclear. So, in order to demonstrate that you have read the instructions, please ignore the preferences form below, and simply write 'I read the instructions" in the 'other' space provided below." Responses from those who did not follow the instruction are considered not paying attention to the survey and, thus, dropped from analysis.

TABLE 3 Survey participant demographics

	Sample used in analysis ($N = 517$)	Total sample collected ($N = 603$)
Household income in 2019: Mean (*SD*)	$83,207 ($69,017)	$79,240 ($67,589)
Age: Mean (*SD*)	36.3 (11.69)	36.0 (11.60)
Household size: Mean (*SD*)	2.9 (1.27)	3.0 (1.30)
Education level		
Less than high school	2	2
High school	33	43
Some college or professional certificate	114	131
Bachelor's degree	265	308
Masters' degree or professional degree	94	110
Ph.D.	9	9
Race		
Asian	47	51
Black	75	93
Hispanic/Latino	38	42
Pacific Islander	3	3
White/Caucasian	340	398
Other	14	16

Note: Responses from participants who fail to correctly respond to attention check questions are not included in the main analysis. For robustness checks, I ran models including the entire responses from 603 participants and there is no statistical difference of coefficients for main explanatory variables.

Participants were asked about their preferences for restaurants depending on the degrees of restaurants' category spanning. Participants from the Washington DC metropolitan area were randomly assigned into one of two scenarios; the first involved their preferences toward variety, and the second toward atypicality. Participants from the other areas were assigned into both scenarios. There was no significant difference in responses between the two groups. First, participants in the variety scenario read explanations about multiple category membership and several examples of restaurants that offer one, two, three, or more categories of food. To assess whether participants understood the concept of variety, I asked them to answer the following question: "We want to make sure if you understand the concept of multiple category memberships. Which of the following restaurants offers three categories of food?" Participants were asked to select an answer from the following three choices, "A restaurant that offers Italian food", "A restaurant that offers Mexican and Korean food", and "A restaurant that offers American, Spanish, and Japanese food". Only the participants who chose "A restaurant that offers American, Spanish, and Japanese food" were included in the final sample for analysis. Participants were then asked to answer, on a seven-point Likert scale (1: "Very unlikely"; 7: "Very likely"), how likely they would try a restaurant that offers one, two, three, and four or more categories of food, respectively. Cronbach's alpha for the latter two questions ("How likely is it that you would try a restaurant that offers *three* categories of food?" and "How likely is it that you would try a restaurant that offers *four or more* categories of food?") is 0.90 and I use the combined two-item measure, *VARIETY*, for analyses.[7]

Next, participants in the atypicality scenario read explanations about the atypicality of dishes and several examples of restaurants that offer typical or atypical dishes. To assess whether participants understood the concept of atypicality, I asked participants to answer the following question: "We want to make sure if you understand the concept of dish typicality. Which of the following is more *atypical* Italian food? Participants were asked to select an answer from the following three choices, "Bolognese spaghetti (spaghetti with meat sauce)," "Kung Pao spaghetti (spaghetti with scallions, peanuts, and hot red chilies)," and "Pizza margherita (with tomatoes and mozzarella cheese)." Only the participants who chose "Kung Pao spaghetti (spaghetti with scallions, peanuts, and hot red chilies)" were included in the final sample for analysis. Participants were then asked to answer, on a seven-point Likert scale (1: "Very unlikely"; 7: "Very likely"), how likely they would try a restaurant that offers atypical dishes from the category it belongs to. I used it to measure, *ATYPICALITY*, for analyses.

4.4.1 | Mediators

After answering the main questions described above, all participants were asked to answer questions about their attributes on a seven-point scale. I used these responses to construct four potential mediators for analyses. Based on the argument of cultural omnivorousness discussed earlier in the theory section (Ollivier, 2008; Roose et al., 2012), I classified the questions into the four groups that represent the various attributes of cultural omnivores: *DIVERSITY*, *NOVELTY*, *OWN*, and *RISK* (see Table 4 for specific questions asked and descriptive statistics of the mediating variables). Mediation tests using these variables may help identify what specific attributes of cultural omnivores contribute to their choice of restaurants and check if the suggested mechanisms are at play in influencing consumers' preference for spanning.

4.4.2 | Analyses and results

Table 5 presents the ordinary least squares regression results using robust standard errors predicting the four potential mediators (Models 17–20) and *VARIETY* (Models 21–22) as dependent variables. Participants' annual household income in 2019 (in $1,000) was used to measure the income variable and a dummy variable for those who received a bachelor's or higher degree is used to measure the education variable. The correlation between the two variables was 0.20, relieving the multicollinearity concern. In all analyses, I controlled for participants' age, race, and household size as they were significant predictors of the dependent variables in some models.

First, participants' household income was not significantly associated with any of the four mediators (Models 17–20) but it was negatively and significantly associated with their preference for *VARIETY* (Models 21–22). For the effect of income (H1), I suggested the mechanism of

[7]Alpha for the three questions ("How likely is it that you would try a restaurant that offers *two* categories of food?", "How likely is it that you would try a restaurant that offers *three* categories of food?" and "How likely is it that you would try a restaurant that offers *four or more* categories of food?") was 0.76. Results using the combined three-item measure are consistent with those using the combined two-item measure. Results are available from the author.

TABLE 4 Questions from the survey to create mediating variables

Variables	Questions asked	Cronbach alpha	Mean	*SD*
DIVERSITY	• I am open to diversity. • Diversity is an important value to me. • I can understand and tolerate difference.	0.7956	5.6	0.97
NOVELTY	• I am the kind of person who would try any new product or service once. • When I eat out, I like to try the most novel and unusual items the restaurant serves even if I am not sure I would like them.	0.7568	4.4	1.52
OWN	• My own evaluation criteria are more important in making decisions than what others think. • I use my own perspective to evaluate anything instead of relying on others' opinions.	0.6565	5.3	1.00
RISK	• I like to take risks. • The greater the risk the more fun the activity.	0.8343	4.1	1.51

consumers' purchasing power and overall demand in a focal community based on their income. While this mechanism needs the community-level observation of market demand and thus is unable to be directly tested in the current mediation analysis, I found a result that is consistent with the mechanism; the result of a logistic regression predicting the likelihood of participants' dining out using robust standard errors demonstrates that income is significantly associated with dining out ($b = 0.0202$, $p = .014$, education and all the control variables included). This result indicates that those who have lower income are less likely to dine out, implying that overall demand for restaurants in a lower-income neighborhood will be smaller than that in a higher-income neighborhood, consistent with H1.

By comparison, education was significantly associated with participants' preference for novelty (*NOVELTY*) and risk taking (*RISK*) (Models 18 and 20, respectively). I used the PROCESS macro for mediation analysis (Hayes, 2017) with 10,000 bootstrap samples to generate 95% confidence intervals (CIs) for the indirect effects of income and education. Table 6A shows the point estimates and CIs of the indirect effects. Only the indirect effect of education on *VARIETY* via *NOVELTY* was significant and positive ($b = 0.1742$, 95% CI: 0.0491, 0.3351). This result is in line with the idea of H2 that educated individuals' preference for unusual and novel items leads them to like restaurants that have multiple category memberships.

Similarly, on predicting participants' preference for atypicality, only *NOVELTY* was a significant mediator between education and *ATYPICALITY*, corroborating H4 that highly educated people, who has preference for novel and unusual items are more likely to prefer restaurants that engage in atypicality-enhancing spanning. Regarding the effect of income, the mediation analysis (Hayes, 2017) suggests that *DIVERSITY* mediates the effect of income on *ATYPICALITY* ($b = 0.0007$, 95% CI: 0.0001, 0.0021 in Table 6B). While this result is consistent with the idea suggested in H3 that small relative cost of exploration would lead people with higher income to be more open to diversity, the magnitude of the indirect effect as well as the effect of income on *DIVERSITY* is close to zero, giving weak support for H3.

TABLE 5 Results of OLS regression from the survey data

	Model 17 *DIVERSITY*	M18 *NOVELTY*	M19 *OWN*	M20 *RISK*	M21 *VARIETY*	M22 *VARIETY*	M23 *DIVERSITY*	M24 *NOVELTY*	M25 *OWN*	M26 *RISK*	M27 *ATYPICALITY*	M28 *ATYPICALITY*
Income	0.0013	−0.0012	0.0001	−0.0012	−0.0038	−0.0035	0.0017	0.0004	−0.0002	0.0007	0.0016	0.0006
	(.107)	(.337)	(.888)	(.362)	(.007)	(.009)	(.134)	(.657)	(.814)	(.523)	(.259)	(.523)
Education	−0.1588	0.7314	0.0092	0.6254	0.0227	−0.2126	−0.1673	0.5476	0.0094	0.5885	0.4129	0.1416
	(.198)	(.000)	(.937)	(.000)	(.914)	(.311)	(.215)	(.005)	(.940)	(.001)	(.074)	(.465)
Age	−0.0038	−0.0127	0.0052	−0.0206	0.0002	0.0071	−0.0114	−0.0256	0.0105	−0.0241	−0.0144	0.0074
	(.464)	(.090)	(.277)	(.004)	(.979)	(.342)	(.024)	(.000)	(.025)	(.000)	(.082)	(.304)
Household size	0.0085	0.0936	0.0228	0.1677	0.1929	0.1483	0.0096	0.1256	0.0047	0.1693	0.0269	−0.0576
	(.832)	(.120)	(.585)	(.005)	(.003)	(.017)	(.840)	(.059)	(.912)	(.009)	(.732)	(.386)
Black	0.3112	0.5743	0.4293	0.8747	0.9253	0.6592	0.1102	0.6898	0.4045	0.5726	0.8133	0.4241
	(.008)	(.007)	(.001)	(.000)	(.000)	(.001)	(.509)	(.014)	(.016)	(.043)	(.011)	(.125)
Hispanic	−0.0709	0.1337	0.0488	0.4615	0.093	0.0127	0.0219	0.3019	0.5482	0.586	0.1866	0.0544
	(.715)	(.563)	(.808)	(.098)	(.773)	(.968)	(.920)	(.331)	(.009)	(.076)	(.628)	(.835)
Asian	−0.2202	−0.3233	−0.1702	−0.3427	−0.1255	0.0170	0.0306	−0.2344	−0.0453	−0.494	−0.3514	−0.2015
	(.203)	(.223)	(.369)	(.198)	(.745)	(.965)	(.854)	(.334)	(.761)	(.039)	(.239)	(.453)
Other race	−0.0237	−0.0908	−0.4797	−0.2353	−0.6188	−0.6044	−0.0543	0.0915	0.0843	0.3405	0.1279	0.0827
	(.932)	(.813)	(.213)	(.558)	(.340)	(.387)	(.832)	(.814)	(.819)	(.266)	(.741)	(.842)
DIVERSITY						0.1526						0.4009
						(.135)						(.000)
NOVELTY						0.2381						0.5141
						(.003)						(.000)
OWN						−0.0902						−0.1641
						(.363)						(.101)

TABLE 5 (Continued)

	Model 17 *DIVERSITY*	M18 *NOVELTY*	M19 *OWN*	M20 *RISK*	M21 *VARIETY*	M22 *VARIETY*	M23 *DIVERSITY*	M24 *NOVELTY*	M25 *OWN*	M26 *RISK*	M27 *ATYPICALITY*	M28 *ATYPICALITY*
RISK						0.1379						0.0993
						(.055)						(.118)
Constant	5.6868	4.1397	5.0526	3.8917	4.6017	2.6671	5.8758	4.3496	4.8815	3.8842	4.5612	0.3851
	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)	(.527)
R-sq	0.0307	0.0928	0.039	0.1407	0.0859	0.1702	0.0392	0.1069	0.0525	0.1336	0.0544	0.3873
N	366	366	366	366	366	366	301	301	301	301	301	301

Note: Robust standard errors are estimated and p-values are reported in parentheses. Two-tailed tests.

TABLE 6 Mediation analysis results using 10,000 bootstrap samples

IV	Mediator	DV	Point estimate of indirect effect	95% CI for indirect path
A. ($n = 366$)				
Income	*DIVERISTY*	*VARIETY*	0.0002	[−0.0001, 0.0008]
	NOVELTY		−0.0003	[−0.0013, 0.0002]
	OWN		0.0000	[−0.0002, 0.0002]
	RISK		−0.0002	[−0.0008, 0.0002]
Education	*DIVERISTY*	*VARIETY*	−0.0242	[−0.0880, 0.0195]
	NOVELTY		0.1742	[0.0491, 0.3351]
	OWN		−0.0008	[−0.03398, 0.0318]
	RISK		0.0862	[−0.0002, 0.1945]
B. ($n = 301$)				
IV	**Mediator**	**DV**	**Point estimate of indirect effect**	**95% CI for indirect path**
Income	*DIVERSITY*	*ATYPICALITY*	0.0007	[0.0001, 0.0021]
	NOVELTY		0.0002	[−0.0008, 0.0016]
	OWN		0.0000	[−0.0003, 0.0004]
	RISK		0.0001	[−0.0003, 0.0004]
Education	*DIVERSITY*	*ATYPICALITY*	−0.0671	[−0.1351, 0.0659]
	NOVELTY		0.2815	[0.0815, 0.4918]
	OWN		−0.0015	[−0.0465, 0.0406]
	RISK		0.0584	[−0.0108, 0.1581]

5 | DISCUSSION AND CONCLUSION

Drawing on prior research on organizational niche width and the sociology of cultural consumption, I have argued that audience heterogeneity affects organizations' likelihood of engaging in two different types of category spanning. The results from the quantitative analyses indicate the contrasting effect of communities' income and education levels on firms' variety-enhancing spanning activities, the extent to which they have multiple categorical memberships. The results are consistent with my argument that while economic affluence enhances market potential and stability, enabling firms to devote their resources to a narrow market niche, higher education leads social elites to be cultural omnivores who have inclusive tastes and a preference for firms that operate in multiple categorical domains.

With regard to atypicality-enhancing spanning, only education has a positive effect and remains statistically significant in the main models. While both wealthy and educated individuals are generally regarded as trendsetters who drive fashion and contribute to the emergence of new trends for atypical hybridized offerings, I have suggested different mechanisms by which the two constructs influence firms' atypicality-enhancing spanning; economic affluence reduces the relative cost of exploring atypical hybridized offerings, whereas higher education influences individuals' preference for those offerings. The significant effect of only the education variable—especially the Ph.D. *variable*—implies that the consumers' preference has a stronger influence than economic affluence on their exploration of hybridized offerings. The results from the mediation analysis using the survey data are also in line with the idea that educated elites who have preference for novelty and unusual items are more likely to desire organizations that engage in category spanning.

One possible alternative explanation for the insignificant effect of income on atypicality-enhancing spanning is related to a scope condition of this study on functional versus non-functional consumption. When new offerings are created as a result of atypicality-enhancing spanning, some of the offerings may retain the original properties and functions of their imported elements (i.e., element aggregation) while others transform into a totally new product with different functions and utilities. The hybridized offering that aggregates functions from disparate categories may appeal to lower-income consumer groups because they may reduce overall cost by choosing the hybridized offering instead of buying multiple single-function offerings (killing two birds with one stone!). I developed my argument focusing on the novelty of the hybridized products because even the aggregator type of hybridized products can provide hedonic utilities to rich trendsetters due to the very newness they bring to the market by collecting disparate functions into one offering, while this newness may be translated as unfamiliarity and uncertainty by budget-constrained consumers. But because the multifunctionality of these products may attract budget-constrained consumers as well as rich trendsetters for a different reason, H3 would work stronger in a context of nonfunctional consumption that particularly emphasizes the novelty of offerings rather than multifunctionality.

6 | LIMITATIONS

Along with the possible alternative explanation for H3, my approach and analysis are not without several notable limitations. First, it remains to be seen if the findings are generalizable to other industries. This study focuses on the restaurant industry, where target demand groups are mainly determined by location. While the idea of tuning firms' business portfolio to align with target demand groups can certainly go beyond geographical boundaries and may be applied to any retail industry, where consumers with different economic and social status directly interact with firms and they have clear categorical expectations, additional empirical analyses of other industries would help confirm or find boundary conditions of the results of the present study.

Also, the restaurant industry provides an appropriate context to assume that overall demand for market categories increases with consumers' income levels because eating out is not an essential activity but an optional one (i.e., you may choose not to eat out at all if your budget is tight). In different contexts, however, aggregated demand will not necessarily change regardless of consumers' income levels and some products or product categories may be wanted or needed by most of the members of a community, making it sufficient for specialists to survive. This boundary condition could also be explored in future research by investigating different industries or market categories.

Second, while the results from the quantitative analyses and the descriptive evidence from the survey provide support to the contrasting effect of consumers' income and education levels on firms' variety-enhancing spanning, a degree to which this contrasting effect is manifested in the real world may be difficult to capture. Because of the high positive correlation between income and education, the negative effect of income may be offset by the positive effect of education in many communities. The contrasting effect, however, may be more salient under some circumstances. For example, 16 out of 85 cities in the study sample were located in the off-diagonal quadrants of the income-education coordinates. The contrasting effect will be more salient in such cities with above-average income and below-average education levels, or vice versa, as illustrated in the online Appendix S1. Also, the effect may be found more prominent in communities that experience sudden change in income levels while education levels remain consistent, like in the case of the COVID-19 crisis or other economic crises under which many people lose their jobs.

Third, this study has investigated a single industry and thus category spanning is examined as straddling culinary boundaries within the industry rather than across industries. While the mechanisms by which the demand influences firms' spanning should be applicable to cross-industry spanning, further empirical analyses should be accompanied to extend and confirm the validity of the mechanisms. Relatedly, category spanning in this study is examined as a business-level strategy, that is, how individual restaurants position themselves in the industry and construct their product offerings. Category spanning as a corporate-level strategy, such as diversification of a restaurant group by having various restaurants, is beyond the scope of this article but warrants attention for further investigation.

A fourth limitation relates to the study sample. I used large cross-sectional data of the restaurant industry. To prevent reverse causality and potential endogeneity, I used the ACS averaged census data from 2008 to 2012 for the independent variables and the spanning information from 2014 for the dependent variables. The sample covers entire organizations in the industry with detailed information about their features, enabling the analysis of the effect of difference in each community on restaurants' spanning. However, the reverse imagery may be possible, while unlikely, whereby a group of people with certain economic and social status moves into neighborhoods that have restaurants that appeal to them. To address this endogeneity issue, I ran a Friedman test (Friedman, 1937) to check if the distribution of the ranks of income and education levels of the cities included in the sample changed between the Census 2000 and the ACS 2008–2012 data. Friedman's chi-square for income ranks has a value of 0.05 and a p-value of .82; it is not statistically significant. For education ranks, Friedman's chi-square has a value of 0.33 and a p-value of .56, also not significant. Hence, there is no evidence that the distributions of income and education ranks among the sample cities have changed between the 13-year period, which relieves concerns about endogeneity. Also, I ran additional analyses using a panel dataset by collecting extra data of restaurants in 2017 located in some of the cities included in the main analysis. Results are consistent with what is shown in this study.[8]

Relatedly, this study analyzes the entire set of organizations in the sample region. Thus, the data do not capture the decision of a restaurant entering the market. Rather, the study examines the outcome of how community features influence the kinds of restaurants that inhabit it. While not capturing the entrants' decision, the result of the study holds important implications for them as to how they should reflect the community features when engaging in spanning activities upon their entry. Indeed, while entry barriers to the restaurant industry may be relatively low compared with industries such as pharmaceutical or semiconductor, for example, a non-negligible number of new restaurants fail during the first year of operation (Parsa et al., 2005). A part of the high failure rate may be due to the misalignment between restaurants' features including their category spanning and community demand.[9] The result of the study

[8]The 2016 American Community Survey I used to collect for the city-level variables covers only cities with populations of 65,000 or more so only the restaurants that were located in these cities in 2017 are included in the panel dataset. The results are available by the author.

[9]In an additional analysis for robustness checks, I ran models with a subsample that excludes 1,159 restaurants of which the age is less than a year. The results of the analysis without the new restaurants are consistent with (and stronger for some models) the results in the main models.

thus highlights the importance of serving local demand when making spanning decisions regardless of how old or new organizations are and the importance of the fit between a firm and its environment, a fundamental concern in strategy. Indeed, the differences of local communities can aid the firm in selecting among available resources they have and help them generate novel items by recombining the resources.

7 | CONTRIBUTIONS

Notwithstanding its limitations, this study makes several contributions to research on category spanning and strategic management. First, this study joins a recent scholarly effort to investigate important antecedents of firms' category spanning and highlights the endogenous nature of firm decisions. Strategic decisions are not randomly made so it is crucial to understand the rationale behind the decisions to fully apprehend the consequences they may bring (Tang & Wezel, 2015). By comprehensively examining the full spectrum of the demand side—from the poor to the rich and from the non-educated to the highly educated—, this study supports the idea that the logic of the categorical imperative may not be universal. Audiences even within the same group may differ in how they interpret category spanning, and this difference translates into contrasting preferences for the same behavior. Some are market-takers and use categories passively as a guiding rule to get what meets their categorical expectations (Pontikes, 2012). Others are more open to perceiving category spanners as flexible, novel, and appealing. For them—such as educated elites—categories may be nothing but a useful tool that aids understanding and exploration of novel organizational forms (Navis & Glynn, 2010; Swidler, 1986). Rather than assuming that consumer groups are monolithic, this article suggests that audience heterogeneity provides different organizational opportunities and constraints for firms' action and contributes to a growing interest in the ways in which category spanning leads to variations in outcomes and their evaluation (Durand & Paolella, 2013; Kennedy & Fiss, 2013).

Furthermore, I believe disentangling the effects of consumers' economic and social status has broader implications for future research on socioeconomic disparities and for strategy formulation. Whereas consumers' economic status and social status are frequently used interchangeably in a simplified manner in a variety of academic fields, this study sheds some light on elucidating the relationship between different facets of socioeconomic position and various other social phenomena. The findings of the study demonstrate the unique roles that income and education play in organizational decisions. This refined approach provides evidence of how the interrelated but still distinct aspects of socioeconomic position can drive opposing results on firm behavior.

Having said that, the results of the study also add a sociological voice to research on firms' location strategies. The implication of the study is consistent with increasing recognition by strategy scholars that "research needs to investigate how contextual factors affect competition, performance, and the development of sustainable competitive advantages" (Marquis & Raynard, 2015, p. 295). This study provides an example of investigating "socio-cultural bridging strategies" (p. 294) which deal with socio-cultural and demographic issues that influence economic performance of organizations. Organizations would benefit from considering how their business scope and product portfolio should be aligned with the communities in which they are, or plan to be, located.

Finally, this study joins recent efforts by category scholars to consider the complex nature of categories and categorization (Durand & Paolella, 2013; Goldberg et al., 2015). In their review article, Glynn and Navis (2013) argue for an understanding of multifaceted nature of categories that is "enabling as well as constraining" (p. 1132). They seek to move beyond the constraining power of categories and to emphasize the generative capabilities of categories. The simultaneous analysis of both market membership and product properties in this study helps theorize and capture both aspects of spanning. Especially, by utilizing detailed product information into computational text analyses, this study contributes to identifying a generative process by which organizations use categories to bring unfamiliar elements into focal domains to enable creative recombination.

ACKNOWLEDGEMENTS

This article is based on my dissertation research. Associate Editor Rodolphe Durand and three anonymous referees greatly improved the article. I am grateful to Ned Smith, Jim Westphal, Gautam Ahuja, Mark Mizruchi, Jerry Davis, David Zhu, Don Lange, Joe Porac, Balázs Kovács, Chris Rider, David Tan, Lori Yue, Filippo Wezel, and numerous seminar participants for their comments on previous drafts. I would also like to thank Kookjin Lee for excellent research assistance. All remaining errors are my own.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID

Heewon Chae https://orcid.org/0000-0003-3159-8060

REFERENCES

Adler, N. E., Boyce, T., Chesney, M. A., Cohen, S., Folkman, S., Kahn, R. L., & Syme, S. L. (1994). Socioeconomic status and health: The challenge of the gradient. *American Psychologist*, *49*(1), 15.

Adner, R., & Zemsky, P. (2006). A demand-based perspective on sustainable competitive advantage. *Strategic Management Journal*, *27*(3), 215–239.

Allenby, G. M., & Rossi, P. E. (1998). Marketing models of consumer heterogeneity. *Journal of Econometrics*, *89*, 57–78.

Belsley, D. A., Kuh, E., & Welsch, R. E. (2005). *Regression diagnostics: Identifying influential data and sources of collinearity*. Philadelphia, PA: John Wiley & Sons.

Bianchi, M. (2002). Novelty, preferences, and fashion: When goods are unsettling. *Journal of Economic Behavior & Organization*, *47*(1), 1–18.

Blau, P. M. (1977). *Inequality and heterogeneity: A primitive theory of social structure*. New York, NY: Free Press.

Bourdieu, P. (1984). *Distinction: A social critique of the judgment of taste*. Cambridge: Harvard University Press.

Bowers, A. (2015). Relative comparison and category membership: The case of equity analysts. *Organization Science*, *26*(2), 571–583.

Brandenburger, A. M., & Stuart, H. W., Jr. (1996). Value-based business strategy. *Journal of Economics & Management Strategy*, *5*(1), 5–24.

Bruine de Bruin, W., Parker, A. M., & Fischhoff, B. (2007). Individual differences in adult decision-making competence. *Journal of Personality and Social Psychology*, *92*(5), 938.

Bryson, B. (1996). "Anything but heavy metal": Symbolic exclusion and musical dislikes. *American Sociological Review*, *61*(5), 884–899.

Carnabuci, G., Operti, E., & Kovács, B. (2015). The categorical imperative and structural reproduction: Dynamics of technological entry in the semiconductor industry. *Organization Science*, *26*(6), 1734–1751.

Carroll, G. R., & Torfason, M. T. (2011). Restaurant organizational forms and community in the US in 2005. *City & Community*, *10*(1), 1–24.
Cattani, G., Ferriani, S., & Allison, P. D. (2014). Insiders, outsiders, and the struggle for consecration in cultural fields: A core-periphery perspective. *American Sociological Review*, *79*(2), 258–281.
Cattani, G., Porac, J. F., & Thomas, H. (2017). Categories and competition. *Strategic Management Journal*, *38*(1), 64–92.
Chan, T. W., & Goldthorpe, J. H. (2007). Class and status: The conceptual distinction and its empirical relevance. *American Sociological Review*, *72*(4), 512–532.
Chang, W., Kim, S., & Chae, H. (2020). A regularized spatial market segmentation method with Dirichlet process—Gaussian mixture prior. *Spatial Statistics*, *35*, 100402.
DiPasquale, D., & Wheaton, W. C. (1996). *Urban economics and real estate markets*. Englewood Cliffs, NJ: Prentice Hall.
Dobrev, S. D., Kim, T. Y., & Hannan, M. T. (2001). Dynamics of niche width and resource partitioning. *American Journal of Sociology*, *106*(5), 1299–1337.
Durand, R., & Kremp, P. A. (2016). Classical deviation: Organizational and individual status as antecedents of conformity. *Academy of Management Journal*, *59*(1), 65–89.
Durand, R., & Paolella, L. (2013). Category stretching: Reorienting research on categories in strategy, entrepreneurship, and organization theory. *Journal of Management Studies*, *50*(6), 1100–1123.
Friedman, M. (1937). The use of ranks to avoid the assumption of normality implicit in the analysis of variance. *Journal of the American Statistical Association*, *32*(200), 675–701.
Gardner, M., & Steinberg, L. (2005). Peer influence on risk taking, risk preference, and risky decision making in adolescence and adulthood: An experimental study. *Developmental Psychology*, *41*(4), 625.
Glynn, M. A., & Navis, C. (2013). Categories, identities, and cultural classification: Moving beyond a model of categorical constraint. *Journal of Management Studies*, *50*(6), 1124–1137.
Goldberg, A. (2011). Mapping shared understandings using relational class analysis: The case of the cultural omnivore reexamined. *American Journal of Sociology*, *116*(5), 1397–1436.
Goldberg, A., Hannan, M. T., & Kovács, B. (2015). What does it mean to span cultural boundaries? Variety and atypicality in cultural consumption. *American Sociological Review*, *81*(2), 215–241.
Hahl, O., Zuckerman, E. W., & Kim, M. (2017). Why elites love authentic lowbrow culture: Overcoming high-status denigration with outsider art. *American Sociological Review*, *82*(4), 828–856.
Hannan, M. T. (2010). Partiality of memberships in categories and audiences. *Annual Review of Sociology*, *36*, 159–181.
Hannan, M. T., & Carroll, G. R. (1992). *Dynamics of organizational populations: Density, legitimation, and competition*. New York, NY: Oxford University Press.
Hannan, M. T., & Freeman, J. (1989). *Organization ecology*. Cambridge, MA: Harvard University Press.
Hannan, M. T., Pólos, L., & Carroll, G. R. (2007). *Logics of organization theory: Audiences, codes, and ecologies*. Princeton, NJ: Princeton University Press.
Hannerz, U. (1990). Cosmopolitans and locals in world culture. *Theory, Culture, & Society*, *7*, 237–251.
Hardin, J. W., & Hilbe, J. M. (2012). *Generalized linear models and extensions* (3rd ed.). College Station, TX: Stata Press.
Hayes, A. F. (2017). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. New York, NY: Guilford Publications.
Holt, D. B. (1998). Does cultural capital structure American consumption? *Journal of Consumer Research*, *25*(1), 1–25.
Hsu, G. (2006). Jacks of all trades and masters of none: Audiences' reactions to spanning genres in feature film production. *Administrative Science Quarterly*, *51*(3), 420–450.
Hsu, G., Hannan, M. T., & Koçak, Ö. (2009). Multiple category memberships in markets: An integrative theory and two empirical tests. *American Sociological Review*, *74*(1), 150–169.
Hsu, G., Negro, G., & Perretti, F. (2012). Hybrids in Hollywood: A study of the production and performance of genre-spanning films. *Industrial and Corporate Change*, *21*(6), 1427–1450.
Kacperczyk, A., & Younkin, P. (2017). The paradox of breadth: The tension between experience and legitimacy in the transition to entrepreneurship. *Administrative Science Quarterly*, *62*(4), 731–764.

Kennedy, M. T., & Fiss, P. C. (2013). An ontological turn in categories research: From standards of legitimacy to evidence of actuality. *Journal of Management Studies*, *50*(6), 1138–1154.

Kluegel, J. R. (1990). Trends in whites' explanations of the black-white gap in socioeconomic status, 1977-1989. *American Sociological Review*, *55*(4), 512–525.

Kovács, B., & Hannan, M. T. (2010). The consequences of category spanning depend on contrast. *Research in the Sociology of Organizations*, *31*, 175–201.

Kovács, B., & Johnson, R. (2014). Contrasting alternative explanations for the consequences of category spanning: A study of restaurant reviews and menus in San Francisco. *Strategic Organization*, *12*(1), 7–37.

Leahey, E., Beckman, C. M., & Stanko, T. L. (2017). Prominent but less productive: The impact of interdisciplinarity on scientists' research. *Administrative Science Quarterly*, *62*(1), 105–139.

Lounsbury, M., & Glynn, M. A. (2001). Cultural entrepreneurship: Stories, legitimacy, and the acquisition of resources. *Strategic Management Journal*, *22*(6–7), 545–564.

Marquis, C., & Raynard, M. (2015). Institutional strategies in emerging markets. *The Academy of Management Annals*, *9*(1), 291–335.

McCullagh, P., & Nelder, J. A. (1989). *Generalized linear models*. Boca Raton, FL: CRC Press.

Meltzer, R., & Schuetz, J. (2012). Bodegas or bagel shops? Neighborhood differences in retail and household services. *Economic Development Quarterly*, *26*, 73–96.

Merluzzi, J., & Phillips, D. (2016). The specialist discount negative returns for MBAs with focused profiles in investment banking. *Administrative Science Quarterly*, *61*(1), 87–124.

Montauti, M., & Wezel, F. C. (2016). Charting the territory: Recombination as a source of uncertainty for potential entrants. *Organization Science*, *27*(4), 954–971.

Navis, C., & Glynn, M. A. (2010). How new market categories emerge: Temporal dynamics of legitimacy, identity, and entrepreneurship in satellite radio, 1990–2005. *Administrative Science Quarterly*, *55*(3), 439–471.

Negro, G., & Leung, M. D. (2013). "Actual" and perceptual effects of category spanning. *Organization Science*, *24*(3), 684–696.

Ollivier, M. (2008). Modes of openness to cultural diversity: Humanist, populist, practical, and indifferent. *Poetics*, *36*(2), 120–147.

Papke, L. E., & Wooldridge, J. M. (1996). Econometric methods for fractional response variables with an application to 401(k) plan participation rates. *Journal of Applied Econometrics*, *11*, 619–632.

Parsa, H. G., Self, J. T., Njite, D., & King, T. (2005). Why restaurants fail. *Cornell Hotel and Restaurant Administration Quarterly*, *46*(3), 304–322.

Peterson, R. A., & Kern, R. M. (1996). Changing highbrow taste: From snob to omnivore. *American Sociological Review*, *61*(5), 900–907.

Pontikes, E. G. (2012). Two sides of the same coin how ambiguous classification affects multiple audiences' evaluations. *Administrative Science Quarterly*, *57*(1), 81–118.

Porac, J. F., Thomas, H., Wilson, F., Paton, D., & Kanfer, A. (1995). Rivalry and the industry model of Scottish knitwear producers. *Administrative Science Quarterly*, *40*(2), 203–227.

Porter, M. E. (1995). The competitive advantage of the inner city. *Harvard Business Review*, *73*(3), 55–71.

Priem, R. L. (2007). A consumer perspective on value creation. *Academy of Management Review*, *32*(1), 219–235.

Priem, R. L., Wenzel, M., & Koch, J. (2018). Demand-side strategy and business models: Putting value creation for consumers center stage. *Long Range Planning*, *51*(1), 22–31.

Psacharopoulos, G., & Patrinos, H. A. (2004). Returns to investment in education: A further update. *Education Economics*, *12*(2), 111–134.

Rao, H., Monin, P., & Durand, R. (2005). Border crossing: Bricolage and the erosion of categorical boundaries in French gastronomy. *American Sociological Review*, *70*(6), 968–991.

Rogers, E. (1995). *Diffusion of innovations*. New York, NY: Free Press.

Roose, H., van Eijck, K., & Lievens, J. (2012). Culture of distinction or culture of openness? Using a social space approach to analyze the social structuring of lifestyles. *Poetics*, *40*(6), 491–513.

Simmel, G. (1957). Fashion. *American Journal of Sociology*, *62*(6), 541–558.

Smith, E. B. (2011). Identities as lenses: How organizational identity affects audiences' evaluation of organizational performance. *Administrative Science Quarterly*, *56*(1), 61–94.

Smith, E. B., & Chae, H. (2016). "We do what we must, and call it by the best names": Can deliberate names offset the consequences of organizational atypicality? *Strategic Management Journal*, *37*(6), 1021–1033.

Swidler, A. (1986). Culture in action: Symbols and strategies. *American Sociological Review*, *51*(2), 273–286.
Tang, Y., & Wezel, F. C. (2015). Up to standard? Market positioning and performance of Hong Kong films, 1975–1997. *Journal of Business Venturing*, *30*(3), 452–466.
Vergne, J. P. (2012). Stigmatized categories and public disapproval of organizations: A mixed-methods study of the global arms industry, 1996–2007. *Academy of Management Journal*, *55*(5), 1027–1052.
Waldfogel, J. (2008). The median voter and the median consumer: Local private goods and population composition. *Journal of Urban Economics*, *63*(2), 567–582.
Wijnberg, N. M., & Gemser, G. (2000). Adding value to innovation: Impressionism and the transformation of the selection system in visual arts. *Organization Science*, *11*(3), 323–329.
Zander, I., & Zander, U. (2005). The inside track: On the important (but neglected) role of customers in the resource-based view of strategy and firm growth. *Journal of Management Studies*, *42*(8), 1519–1548.
Zerubavel, E. (2009). *Social mindscapes: An invitation to cognitive sociology*. Cambridge, MA: Harvard University Press.
Zuckerman, E. W. (1999). The categorical imperative: Securities analysts and the illegitimacy discount. *American Journal of Sociology*, *104*(5), 1398–1438.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Chae, H. (2021). Income or education? Community-level antecedents of firms' category-spanning activities. *Strategic Management Journal*, *43*(1), 93–129. https://doi.org/10.1002/smj.3328

APPENDIX

Calculation of atypicality-enhancing spanning

This section provides an illustration of how atypicality-enhancing spanning scores are calculated using a hypothetical example. The front part of the computational process (Tables A1–A3) follows that of Kovács and Johnson (2014). Then, I develop and introduce a novel measure for atypicality-enhancing spanning.

In this hypothetical example, there are six restaurants, each with only two items (Table A1).

TABLE A1 Hypothetical example of a list of restaurants

Restaurant	Categories	Menu	Restaurant	Categories	Menu
1	Korean	Kimchi bibimbop, bulgogi	4	American	Macaroni and cheese, hamburger
2	Korean	Kimchi bibimbop, bulgogi	5	Korean, American	Kimchi macaroni, bulgogi hamburger
3	American	Macaroni and cheese, hamburger	6	Korean, American	Kimchi bibimbop, macaroni and cheese

First, using their cuisine categories and words on the menu, I make a category-word occurrence table (Table A2). In the case of multiple-category restaurants, I divide the occurrence of the menu items by the number of categories the restaurant belongs to, following Kovács and Johnson (2014).

TABLE A2 Occurrence table

	Kimchi	Bibimbop	Bulgogi	Cheese	Macaroni	Hamburger	Total
Korean	3	3.5	2.5	0.5	2.5	0.5	12.5
American	1	0.5	0.5	2.5	3	2.5	10
Total	4	4	3	3	5.5	3	22.5

Next, I compute the Jaccard similarity index for all word-category pairs (Table A3) to get typicality scores, using the following formula:

$$Typicality\left(word_i, category_j\right) = \frac{\#\left(word_i \& category_j\right)}{\#(word_i), + \#\left(category_j\right) - \#\left(word_i \& category_j\right)}..$$

where, $\#(word_i\&category_j)$ denotes the number of times the word *i* appears on menus in category *j*, $\#(word_i)$ denotes the total number of times the word *i* appears on the menus of all restaurants, and $\#(category_j)$ denotes the total number of words in category *j*. For example, the word *kimchi* appears three times in the Korean category and four times total, and there are 12.5 words in total in the Korean category. Then,

$$Typicality(\text{“}kimchi\text{”}, \text{“}Korean\text{”}) = 3/(4 + 12.5 - 3) = 0.22.$$

TABLE A3 Word typicality: The Jaccard similarity index for all word-category pairs

	Kimchi	Bibimbop	Bulgogi	Cheese	Macaroni	Hamburger
Korean	0.22	0.27	0.19	0.03	0.16	0.03
American	0.08	0.04	0.04	0.24	0.24	0.24

As shown in Table A3, the word “kimchi” has a high typicality score in Korean and a low typicality score in American whereas the word “cheese” has a low typicality score in Korean and a high typicality score in American. From Table A3, I calculate the typicality of each dish in each claimed category by taking the weighted average of the Jaccard similarities of the menu words used in the item description. For example, the typicality of the dish “kimchi macaroni” of Restaurant 5 is 0.19 (=[0.22 + 0.16]/2) in the Korean category and 0.16 (=[0.08 + 0.24]/2) in the American category.

After obtaining dish typicalities in each category, the highest score among dish typicalities in the categories is chosen for each dish, and these scores are averaged to calculate the average dish typicality of a restaurant (Table A4). For example, "kimchi macaroni" has a higher typicality score in the Korean category than in the American category, so the score in the Korean category is selected. "Bulgogi hamburger" has a higher value in the American category, so the value in the American category is selected. And by averaging the two values (i.e., 0.19 and 0.14), the average dish typicality of Restaurant 5 is calculated (=0.165). Because the dishes of Restaurant 5 mix disparate elements together, even the highest dish typicality is relatively low compared with the case of Restaurant 6, which provides a very typical Korean dish and a very typical American dish. Thus, the average dish typicality of Restaurant 5 is lower than that of Restaurant 6. Finally, because typicality and category spanning are inversely related, I subtract the average dish typicality of a restaurant from 1 to get the degree of atypicality-enhancing spanning for a restaurant. Because the values of atypicality-enhancing spanning are low in absolute number "due to the division by the count of words in the Jaccard formula" (Kovács & Johnson, 2014, p. 12), for better interpretability I rescale the values to make the maximum observed value of atypicality-enhancing spanning 1 and the minimum 0.

TABLE A4 Dish typicality and atypicality-enhancing spanning of restaurants 5 and 6

Restaurant	5		6	
Categories	Korean, American		Korean, American	
Menu	Kimchi macaroni	Bulgogi hamburger	Kimchi bibimbop	Macaroni and cheese
Typicality in Korean	0.19 = (0.22 + 0.16)/2	0.11	0.25	0.1
Typicality in American	0.16 = (0.08 + 0.24)/2	0.14	0.06	0.24
Dish typicality	0.19	0.14	0.25	0.24
Average dish typicality	0.165		0.245	
Atypicality-enhancing spanning	0.835 (= 1–0.165)		0.755	