Received: 27 November 2024 | Revised: 26 August 2025 | Accepted: 1 September 2025

DOI: 10.1002/smj.70018

RESEARCH ARTICLE

SMS | Strategic Management Journal

WILEY

Right on cue? Category-switching in online marketplaces

Karl Taeuscher[1] | Eric Yanfei Zhao[2] | Michael Lounsbury[3]

[1]Alliance Manchester Business School, University of Manchester, Manchester, UK

[2]Saïd Business School, St Hugh's College, New College, University of Oxford, Oxford, UK

[3]Alberta School of Business, University of Alberta, Edmonton, Canada

Correspondence
Karl Taeuscher, Alliance Manchester Business School, University of Manchester, Booth Street West, Manchester M15 6PB, UK.
Email: karl.taeuscher@manchester.ac.uk

Abstract

Research Summary: When and why do producers change the categorization of their offerings? Prior categorization research assumes that producers engage in ongoing efforts to proactively optimize their categorical positioning, but this assumption may not hold for many producers due to their limited attentional capacity. Our theoretical account instead highlights the role of expectation violation cues—salient pieces of information indicating a violation of audience expectations—as triggers that can lead producers to revise their category choice. Our longitudinal study of 84,667 Airbnb hosts' categorization choices finds that negative customer reviews—an important form of expectation violation cue—significantly increase the likelihood of category-switching, particularly in categories with heterogeneous expectations. Our study suggests that many producers might be less proactive about their category choices than previous research assumed.

Managerial Summary: How businesses categorize their products and services influences their commercial success; yet, there exists very limited understanding of when, why, and how businesses revise their category choices. Our study, which tracked the category choices of over 80,000 Airbnb hosts over time, reveals that Airbnb hosts most commonly switch categories after receiving negative customer reviews, particularly when a review indicates that the accommodation did not meet customer expectations and if the previously chosen category lacks a clearly defined set of expected features. When switching categories, hosts tend to choose categories that are relatively similar to their prior choices and that seem to accommodate a wide variety of offerings. These patterns suggest that many businesses might be less proactive about their category choices than previous research assumed.

KEYWORDS

categories, expectation violations, online marketplaces, problemistic search, strategic categorization

1 | INTRODUCTION

Categories play a central role in how audiences evaluate producers and their offerings. Over the past decades, scholars have developed rich theoretical accounts of how categories and categorization processes shape market cognition and guide evaluation processes (for reviews, see Cattani et al., 2017; Durand & Thornton, 2018; Vergne & Wry, 2014). Categories function by anchoring audience expectations because they embody "a meaningful consensus about some entities' features as shared by actors grouped together as an audience" (Durand & Paolella, 2013, p. 1100). Yet, while audiences use categories as interpretive lenses, producers often retain substantial agency in how they categorize themselves and their offerings (Glynn & Navis, 2013)—particularly in category systems with overlapping or loosely defined boundaries (Lo et al., 2020).

This recognition has given rise to a growing literature on strategic categorization, which shifts attention from categories as externally imposed constraints (e.g., Zuckerman, 1999) to categories as cultural resources that producers can actively deploy (e.g., Aversa et al., 2021; Gehman & Grimes, 2017; Granqvist et al., 2013; Kodeih et al., 2019; Pontikes, 2022; Pontikes & Barnett, 2015; Pontikes & Kim, 2017). Although this research has generated important insights into producers' initial category choices and the strategic value of different categories, it has paid surprisingly little attention to when and why producers engage in *category-switching*—that is, producers' deliberate decision to change their offering's categorization. Notable exceptions that take a more dynamic view of producers' category choices (e.g., Kovács et al., 2024; Pontikes, 2022; Pontikes & Barnett, 2015) often implicitly assume that category-switching results from producers' ongoing efforts to proactively optimize their categorical positioning. However, such an assumption likely does not hold for all types of producers, particularly for capacity-constrained producers such as freelancers, solo entrepreneurs, or small businesses, who are highly constrained in their attention and decision-making capacity (Ocasio, 2011; Simon, 1990).

In this paper, we develop a theoretical account of category-switching that more realistically accounts for the limited attention and decision-making capacity of such capacity-constrained producers (Ocasio, 2011). Building on key insights from problemistic search theory (see Posen et al., 2018), we argue that many capacity-constrained producers will rarely revisit their category choices unless prompted by external cues that indicate a potential problem with their current positioning. Specifically, we emphasize the central role of *expectation violation cues*—salient pieces of information that indicate a potential violation of audience expectations. Such cues can trigger a two-stage search process in which producers first aim to diagnose the cause of the expectation violation (problem search) before searching for solutions to resolve the identified problem and prevent further expectation violations (solution search). If producers attribute the expectation violation to a mismatch between their offering and the expectations set by its categorical positioning, category-switching may emerge as a preferred solution.

In the context of online marketplaces, we examine negative customer reviews as a particularly salient type of expectation violation cue. We hypothesize that negative reviews increase the likelihood of category switching, especially when the focal category is characterized by heterogeneous expectations—that is, when there is little consensus about the features expected of its members. In such cases, it becomes harder for producers to resolve the perceived mismatch by aligning their offering more closely with categorical expectations (e.g., by adding features), making category-switching a comparatively more viable solution. Drawing on prior work on category consensus and heterogeneity (Cattani et al., 2008; Lo et al., 2020; Soublière et al., 2024), we predict that category heterogeneity will amplify the positive relationship between negative customer reviews and category-switching.

To test our hypotheses, we assembled a longitudinal dataset of all Airbnb listings in 12 major US markets from 2018 to 2021 ($N = 111{,}088$ listings), sourced through repeated web-scraping surveys at six-month intervals, as well as all customer reviews these listings received during our study period ($N = 2{,}262{,}300$ reviews). Within the Airbnb marketplace, individual producers ("hosts") offer accommodations for short-term rental. During our study period, Airbnb's categorization system consisted of 47 mutually exclusive accommodation categories, and each listing was affiliated with exactly one category.[1] The setting offers several empirical advantages: category-switching is unambiguous, listing features are relatively stable, and customer reviews are validated, time-stamped, and publicly visible.

Our study shows a strongly positive association between Airbnb listings' number of negative reviews and their likelihood of category-switching, particularly when the listing is affiliated with a heterogeneous category. Each negative review received in the prior period increases the predicted probability of category-switching by 27.8% if category heterogeneity is high, but only by 8.7% if category heterogeneity is low. Our supplemental content analysis of all negative customer reviews, leveraging advances in large language models (LLMs), further supports our theoretical arguments by showing that negative reviews most strongly increase listings' category-switching likelihood when they explicitly communicate some form of feature-related

[1]One advantage of this context is that Airbnb's classification system was very stable throughout the study period as Airbnb hosts lack the power to create new categories or significantly influence a category's meaning and associated expectations, and Airbnb—as the marketplace provider—did not add any or remove any categories. Our theorization is thus bounded to contexts in which category-switching does not represent an (inevitable) response to the death of a category or the rapid decline of a category's appeal, and in which producers do not have the power to substantially alter category-specific audience expectations.

expectation violation. Moreover, consistent with prior insights about producers' solution search in decision contexts like ours (Gavetti & Levinthal, 2000), we find that—when category-switching—producers are most likely to switch into a category in close proximity to their original category. Among such proximate categories, producers tend to choose those that exhibit greater heterogeneity because these categories seem particularly accommodating to a wide range of offerings.

Our study makes several contributions to categories research. First, our study problematizes the prior assumption that producers continuously seek to optimize their categorical positioning (e.g., Aversa et al., 2021; Pontikes & Barnett, 2015; Pontikes & Kim, 2017), identifying capacity-constrained producers as a critical boundary condition under which this assumption is likely violated. In the context of such producers, category-switching may represent a more reactive behavior than prior theory would suggest. Second, by introducing the notion of *expectation violation cues* and outlining how such cues can trigger a search process that may result in the decision to category-switch, our theoretical account advances understanding of the specific triggers and events that may lead producers to revisit and potentially revise their prior category choices. Third, our study advances understanding of how category heterogeneity shapes organizational behaviors and outcomes (Haans, 2019; Lo et al., 2020; Soublière et al., 2024) by highlighting how a lack of consensus around a category's expected features can simultaneously push existing members out of the category while attracting non-members to the category. We discuss how our findings can spark important new directions for categories research, including research concerned with the consequences of categorization choices.

2 | THEORETICAL BACKGROUND AND HYPOTHESES

2.1 | Category-switching as a form of strategic categorization

Interest has surged among organization theorists, strategic management, and entrepreneurship scholars regarding the influence of categories on organizational and entrepreneurial behaviors as well as their implications (see Cattani et al., 2017; Durand & Thornton, 2018; Vergne & Wry, 2014). Categories generally allow audiences to recognize an evaluation object (like an organization or offering) as part of a group of similar objects (Grodal et al., 2015). Each category consists of a unique semantic label that is associated with certain information and features (Wry et al., 2011), thereby anchoring audience expectations about the probable features of category members (Durand & Paolella, 2013). Perceiving an offering as a member of a given category allows audiences to infer certain information about the otherwise unknown offering, while influencing their expectations about the offering's features (e.g., Hannan et al., 2019). Alignment with such category-specific expectations is commonly conceived as a major mechanism through which organizations and their offerings gain legitimacy in the eyes of evaluating audiences, and offerings failing to meet category-specific expectations tend to be disregarded or devalued by many evaluators (Zuckerman, 1999).

The growing body of strategic categorization research emphasizes producers' agency in categorizing themselves and their offerings (Aversa et al., 2021; Gehman & Grimes, 2017; Granqvist et al., 2013; Kovács et al., 2024; Pontikes, 2022; Pontikes & Barnett, 2015; Pontikes & Kim, 2017). In line with prior work (e.g., Pontikes & Barnett, 2015), we use the label of *producers* to refer to supply-side actors whose offerings get categorized. Moving away from viewing categories as purely external constraints (for a critique of this view, see Glynn & Navis, 2013), this body of research conceives of categories as cultural resources (Lounsbury & Glynn, 2019) that producers can strategically deploy to their advantage (Pontikes & Kim, 2017; Verhaal & Pontikes, 2022; Zhao et al., 2013).

Category-switching, which refers to a producer's deliberate decision to change the category with which they affiliate themselves or their offerings, stands as a vital yet underexplored facet of strategic categorization. While we recognize that category-switching can also entail substantive changes in the offering or organization, our theoretical development focuses on category-switching as a largely symbolic action. As such, we conceive of category-switching as conceptually distinctive from market exits and entries—that is, actions that are primarily reflected through resource divestments from or investments into a given market domain—which have received much more scholarly attention (e.g., Durand & Vergne, 2015; Pontikes & Barnett, 2017).[2] Rather, category-switching represents a form of cultural entrepreneurship (Lounsbury & Glynn, 2001) in which producers leverage categories to favorably influence perceptions about themselves or their offerings.

The few existing studies that take a dynamic perspective on producers' category choices commonly aim to explain instances of category-switching by attributing them to changes in the relative utility of those categories that a producer could plausibly choose from—for example, in terms of the degree to which each given category provides meaningful information to audiences and/or the degree to which audiences attend to the given category (Pontikes & Barnett, 2015; Pontikes & Kim, 2017). An implicit assumption underlying such theoretical accounts is that producers engage in proactive efforts to optimize their categorical positioning; therefore, monitoring and frequently re-evaluating the utility of both their own and alternative categories. Yet, such an assumption unlikely holds for all producer types—particularly those producers that only consist of one or a few individuals (e.g., solo-entrepreneurs, freelancers, small businesses) because such producers are highly constrained in their attention and decision-making capacity. Given the general limits on any individual's attention and decision-making capacity (March & Simon, 1958), as well as the myriad of issues and decisions that most producers have to attend to, it seems unlikely that such *capacity-constrained producers* would invest their limited cognitive capacities to engage in ongoing proactive efforts to optimize their categorical positioning.

While our main arguments aim to be generalizable, we subsequently contextualize our theoretical account of category-switching by specifically focusing on capacity-constrained producers that compete with their offerings in online marketplaces. Given our conceptualization of category-switching, online marketplaces provide a useful context to explore the factors that lead producers to switch categories.[3] To start, online marketplaces—such as eBay, Amazon, Etsy, Udemy, or Airbnb—tend to represent competitive domains in which a large number of relatively small producers compete with their offerings for consumers' attention. Categories play a central role in online marketplaces as they help consumers to preselect a cognitively manageable number of offerings, provide information about otherwise unfamiliar offerings, and support the formation of realistic expectations about the unfamiliar offering (Piazzai et al., 2024). Indeed, empirical evidence suggests that offerings' categorization within online marketplaces

[2]Due to its largely symbolic nature, we also conceive of category-switching as conceptually distinctive from the notions of strategic change and strategic reorientation, which generally refer to more substantive actions (e.g., Rajagopalan & Spreitzer, 1997) and tend to involve a substantial shift in an organization's overall mission, priorities, and goals (e.g., Fiss & Zajac, 2006).

[3]While some online marketplaces allow producers to affiliate their offering with more than one product category, we specifically focus on online marketplaces in which each product is situated in exactly one category. As such, our context differs from settings in which producers can span multiple categories.

substantially influences their audience perceptions and commercial success (e.g., Cutolo & Ferriani, 2024b; Piazzai et al., 2024). Moreover, online marketplaces provide a useful research setting for our purposes because producers' category choices are fully transparent and publicly observable. With an empirical focus on online marketplaces, our theorizing is therefore bounded to contexts in which the relevant category system tends to be highly transparent.

2.2 | Expectation violation cues as triggers for category-switching

In this section, we propose that capacity-constrained producers rarely revisit their prior category choice unless they are triggered to do so by certain external cues. This proposition is informed by key insights from research on problemistic search, which attends to the cognitive processes underlying individuals' decision-making as they relate to decisions of change, among others (see Posen et al., 2018). This body of research emphasizes that individuals are generally reluctant to change their previous decisions unless they become cognitively aware of problems arising from them (March & Simon, 1958). Given their limited cognitive capacity and the virtually infinite array of potential issues and decisions, individuals are significantly more likely to attend to a particular issue if it becomes cognitively salient (Simon, 1990). An issue typically becomes cognitively salient when a decision-maker encounters a problem related to it (Cyert & March, 1963), prompting a search process in which the decision-maker first seeks to diagnose the cause of the problem (problem search) and then searches for a suitable solution (solution search) (Posen et al., 2018). This process suggests that change decisions most often arise after a potential problem captures the decision-maker's attention—what Cyert and March describe as a "problemistic search … stimulated by a problem [and] de-pressed by a problem solution" (Cyert & March, 1963, p. 121). These insights imply that capacity-constrained producers are, by default, unlikely to continuously monitor or reassess the relative utility of available categories, or to engage in ongoing efforts to proactively optimize their categorical positioning. Drawing on these behavioral insights, we expect that once capacity-constrained producers initially select a category, they tend to devote little attention to that choice unless a potential problem becomes cognitively salient.

We argue that producers' likelihood of attending to their category choice increases substantially when they encounter *expectation violation cues*—salient pieces of information indicating that their offering appears to violate customers' expectations. As outlined above, categories play a central role in evaluation processes because they anchor customers' expectations about the features of an offering. Cues that these expectations may have been violated can trigger a search process in which the producer first seeks to diagnose the source of the apparent expectation violation (i.e., problem search) and subsequently considers possible solutions to prevent such expectation violations in the future (i.e., solution search). This process, we argue, can result in category-switching if the producer (i) identifies a discrepancy between the expected and actual features of their offering and (ii) considers category-switching as the most viable means to prevent such a perceived discrepancy in the future.

In the context of online marketplaces, we focus specifically on *negative customer reviews* as a particularly relevant form of expectation violation cues. Cues tend to attract decision-makers' attention and prompt action when they provide salient, novel, and vivid information that is strategically relevant (Li et al., 2013). Negative customer reviews fit these criteria well: they are prominently displayed on most online marketplaces (Proserpio & Zervas, 2017), and their free-text format contributes to their novelty and vividness (Pavlou & Dimoka, 2006). Moreover, negative reviews carry strategic weight, as they disproportionately influence other consumers' purchasing decisions (Dellarocas et al., 2007).

For capacity-constrained producers in online marketplaces, a negative customer review can trigger a problemistic search process that may lead to category-switching. Research on consumer behavior suggests that negative customer evaluations generally result from a gap between the customer's a priori expectations and their actual experience with the product or service (Parasuraman et al., 1985; Zeithaml et al., 1990). This gap can stem from various sources—for example, customers may have expected different features or a higher quality level (Reeves & Bednar, 1994). As salient indicators of potential violations of customer expectations, negative reviews are thus likely to prompt producers to engage in problemistic search.

In the first stage of the problemistic search process (problem search), producers attempt to diagnose the source of the seeming expectation violation. Due to the social conventions that shape online reviewing behavior (see Dellarocas, 2003), most customer reviews tend to be very short and lacking in detail. Consequently, many negative reviews merely express dissatisfaction without offering a comprehensive explanation of the underlying issue (Mudambi & Schuff, 2010). This brevity often leaves producers with considerable uncertainty during the problem search stage. One possible outcome is that the producer attributes the gap between customers' expectations and experience to the offering's categorical positioning—recognizing that their category choice may have inadvertently created certain expectations that their offering does not fulfill. While explicit comments about unmet feature expectations are particularly likely to draw producers' attention to their category choice, even negative reviews that do not specify any specific expectations may therefore prompt producers to scrutinize their categorical positioning as a potential source of the expectation violation.

During the solution search stage, producers seek to identify ways to address the assumed source(s) of the expectation violation. Category-switching can represent a viable solution, particularly if the producer has identified a discrepancy between expected and perceived features. By switching categories, producers may attempt to reshape customer expectations in ways that more closely align them with the actual attributes of their offering. In some instances, producers may choose to switch categories even amid uncertainty about the precise source of the problem—viewing category-switching as a relatively low-cost solution compared to more substantive changes, such as redesigning or enhancing the offering itself.

In sum, category-switching represents one possible outcome of a problemistic search process triggered by negative customer reviews. Although not all negative reviews will trigger such a search process, we expect that a greater number of negative customer reviews will increase the likelihood that a producer engages in problemistic search that leads them to reassess their categorical positioning. All else being equal, receiving a greater number of negative reviews should thus heighten the probability that producers attend to—and potentially change—their categorical positioning. Before examining the categorical conditions under which producers will most likely choose category-switching as a solution to address the expectation violation indicated by negative reviews, we propose the following baseline hypothesis:

> **Hypothesis 1.** There is a positive relationship between the number of negative customer reviews an offering receives and its likelihood of category-switching.

2.3 | The moderating role of category heterogeneity

Thus far, we have argued that negative customer reviews—as expectation violation cues—trigger a search process that can draw producers' attention to their categorical positioning as a potential source of the perceived violation. This process may ultimately lead to category-switching as a means of preventing such violations in the future. Extending this argument, we propose that the likelihood of a producer selecting category-switching as a solution during the solution stage depends on the degree of consensus around expected features within the given category. Specifically, we propose that in categories in which audience expectations are diverse and inconsistent (Cattani et al., 2008; Kennedy et al., 2010), producers face greater difficulty in aligning their offerings with shared expectations when aiming to prevent future expectation violations, making category-switching a more likely outcome of the search process.

When faced with a perceived violation of audience expectations, producers typically have a range of strategic responses to choose from (e.g., Bundy et al., 2013; Durand et al., 2019). These may include substantive adjustments to the offering itself to better meet customers' expectations or symbolic actions aimed at better communicating their alignment with audience expectations (Durand et al., 2019). Such responses are more viable when there is a clear and collectively shared understanding of what constitutes appropriate features within the given category. In such cases, if producers diagnose a mismatch between the features of their offering and the features expected in the category, they can reasonably resolve the issue by adjusting their offering's features and/or by better communicating how their offering meets those categorical expectations—that is, address the problem without necessarily switching category. Hence, producers are less likely to resort to category-switching when categorical expectations are clearly defined.

To derive a testable hypothesis for our argument, we focus on *category heterogeneity*, defined as the degree of diversity in the features exhibited by members of a product category (Soublière et al., 2024), as an observable characteristic that reflects the extent to which a consensus about expected features exists (Haans, 2019; Soublière et al., 2024). In highly homogeneous categories, where feature diversity is minimal, there tends to be a well-understood and widely accepted prototype of what category members should offer (Cattani et al., 2008). Producers operating in such categories typically have a clear sense of audience expectations, enabling them to address feature-related mismatches through targeted adjustments or persuasive explanations of why their offering does not provide a commonly expected feature. Accordingly, even when negative reviews highlight expectation violations, the problemistic search process in highly homogeneous categories is more likely to lead to solutions that preserve category membership rather than prompt category-switching.

In contrast, highly heterogeneous categories tend to lack a clearly defined set of expected features, which leads to greater variability in the expectations of customers (Soublière et al., 2024) and can complicate any effort to prevent violation expectations. Expectation violation cues represent an important mechanism through which a producer learns about these diverse customer expectations and the potential challenges they create. For instance, a producer may be confronted with a situation in which the negative reviews they receive convey that different customers have conflicting feature expectations for the given heterogeneous category. When aiming to prevent further expectation violations for a listing in a highly heterogeneous category, producers will therefore find it comparatively difficult—if not impossible—to do so through better alignment of their offering with the given category's diverse and potentially conflicting expectations. Consequently, when negative reviews trigger a problemistic search process, producers operating in more heterogeneous categories are more likely to consider category-switching as the most viable path to prevent future violation expectations. All else equal, we therefore expect that category heterogeneity will increase the likelihood that producers respond to negative customer reviews by switching categories.

Hypothesis 2. Category heterogeneity positively moderates the relationship between negative customer reviews and category-switching, such that offerings situated in more heterogeneous categories are more likely to switch categories in response to negative customer reviews than those in more homogenous categories.

2.4 | Category selection during category-switching

Having established the conditions under which producers are likely to switch categories, we now turn to the question of how producers select a new category once they have decided to switch. If, as we have argued, capacity-constrained producers in online marketplaces primarily engage in category-switching as a corrective response to prevent expectation violations, then they should primarily aim to select a new category with the intention to prevent any future expectation violations. In other words, the goal of producers' category selection in online marketplaces is to identify a category that establishes realistic expectations about their offering's features.

However, due to capacity-constrained producers' bounded rationality (March & Simon, 1958) and the inherent ambiguity surrounding the expected features of many categories, it is often difficult to predict ex ante whether positioning an offering in a given category will indeed prevent future expectation violations. As such, producers may only be able to fully evaluate the degree of alignment between their offering and a new category once they have positioned their offering in that category. This process reflects what prior research refers to as experiential or "on-line" search—a solution search process that requires at least partial implementation of the solution in order to assess its efficacy (Gavetti & Levinthal, 2000). In such cases, producers tend to avoid solutions that are too distant from their current one, as more distant options carry higher perceived risk. As Gavetti and Levinthal (2000, p. 115) explain, the variation that results from changing to a distant solution "may prove fatal because the actor experiences the consequences of each experimental draw." Given their limited attentional capacity, producers typically settle for a solution that appears "good enough" rather than exhaustively evaluating all possible options (e.g., Cyert & March, 1963; Greve, 2003). As a result, producers engaging in "on-line" search processes tend to select solutions that are similar, rather than very different from, their prior choice (Gavetti & Levinthal, 2000).

Applied to category-switching, this means that producers—faced with uncertainty and limited capacity—are more likely to begin their search for a new category by considering categories that are proximate to their current category. Hence, they will likely attend first to those categories that occupy nearby positions in the feature space—those that share similar characteristics with their current category (Hannan et al., 2019). Such proximate categories are not only easier to cognitively access but also more likely to be perceived as low-risk options (Gavetti & Levinthal, 2000). Since capacity-constrained producers are likely to stop searching once they identify a category that appears sufficiently compatible, the likelihood of selecting a category should thus be higher when the category is proximate to, as opposed to distant from, the original one. We thus hypothesize:

Hypothesis 3. Conditional on category-switching, the likelihood of selecting a given category increases with the proximity between the original and the new category.

Earlier, we proposed that producers are more likely to switch out of a category upon receiving negative reviews if that category exhibits high heterogeneity. For producers who positioned their offering in a heterogeneous category, we argued, negative reviews not only represent cues that can lead them to revisit their categorical positioning but also serve as an important mechanism through which producers learn about consumers' diverse and potentially conflicting expectations about the given category. Anchored to our Hypothesis 3, we now consider the role of category heterogeneity in the selection of a new category among a set of proximate categories.

Among the proximate category options, producers will least likely choose to switch into a category if they perceive their offering as clearly incompatible with that category's expected features. In the absence of any experiential insights about specific customer expectations for a given category, capacity-constrained producers may primarily infer information about their offering's compatibility with a category from the offerings that are already positioned in that category. A category that exhibits high homogeneity—that is, in which all offerings share a very narrow set of features and expectations are clearly understood—should lead a producer to more easily recognize any incompatibility between their offering's features and the category's narrow set of expected features. In contrast, a heterogeneous category will seem more accommodating to a wide range of offerings—with some of the diverse offerings potentially resembling the producer's own offering—and the lack of well-defined expectations will make it less likely that producers perceive their offering as clearly incompatible with the category.

While we do not assume that capacity-constrained producers strategically seek to select heterogeneous categories or that heterogeneous categories represent more rational choices, these arguments suggest that producers should more likely perceive a given proximate category as sufficiently compatible if that category exhibits greater heterogeneity. During the previously outlined solution search process—in which producers iteratively consider the most proximate categories until they identify a seemingly compatible category—producers should therefore disproportionately select a proximate category that exhibits high heterogeneity because, prior to any experiential learning about customers' diverse expectations for the category, such a category will more likely be perceived as accommodating and compatible. Therefore, among proximate categories, the likelihood that a producer selects a given category should increase at greater levels of heterogeneity. We thus hypothesize:

> **Hypothesis 4.** The positive relationship between a category's proximity and its likelihood of being selected will be accentuated (i.e., more positive) by greater category heterogeneity, such that the likelihood of selection is highest when a category is both proximate and heterogeneous.

3 | DATA AND METHODS

3.1 | Study context and data sources

We test our hypotheses by examining how individuals categorize their accommodation offerings on the Airbnb platform. Airbnb offers individuals and organizations access to a vast consumer base, boasting nearly 400 million bookings worth $63.2 billion in 2022 (Airbnb, 2023). Although Airbnb has diversified its offerings in recent years, its primary focus remains on short-term accommodation rentals. In this setting, individuals (“hosts”) display their accommodation listings in hopes of attracting potential customers ("guests"). Prior research indicates that the vast majority of Airbnb hosts are individuals, family units, or micro-businesses (Fradkin & Holtz, 2023; Zervas et al., 2021)—that is, they closely resemble our conceptualization of capacity-constrained producers. Airbnb also exemplifies a prototypical online marketplace in that individual producers typically lack the ability to change the category system, to shape the collectively understood meaning or desirability of specific categories, or to alter customers' general expectations about a given category.

Airbnb's categorization system comprises 47 accommodation categories (referred to as "property type"), including options such as *Tiny house*, *Loft*, *Townhouse*, *Cottage*, and *Bed and Breakfast*. Hosts are required to assign their accommodation offering to one specific category.[4] Designed to streamline the customer search and evaluation process, these categories convey established meanings and help shape consumer expectations. For instance, a listing categorized as a *Houseboat* signals to potential guests features typically associated with that category. On the listing's profile page, the selected category is prominently displayed. An online post by a San Francisco-based host in an Airbnb forum affirms that hosts have full discretion over their category selection and can change it at any time. The quote below also illustrates hosts' awareness of the importance of aligning their offering with category-specific expectations:

> The topic of how to [select the category] can be confusing. One of the most frustrating aspects of the issue is this: there really isn't any person or entity who evaluates the listings and gives out the correct answer. There really isn't any way [for a host] to find out what this would be considered because there is no one to ask! (Even though we all want to know the answer.) There is a solution though; a way to remember to think about it. The goal of choosing a [category] is to communicate to the guest what they are booking. If the guests who stay with you are happy with what they receive, then you have chosen the correct [category]. If guests complain that what they got is not what they expected, you have chosen the wrong [category].
>
> (Airbnb, 2018)

To empirically examine category-switching behavior, we constructed a longitudinal panel of Airbnb listings using web-scraped data collected between 2018 and 2021. A key advantage of this study period is the stability of Airbnb's classification system: no categories were added or removed during this time, ensuring that observed category switches are not artifacts of changes in the classification scheme. Our dataset includes all Airbnb listings from 12 US cities: Austin (Texas), Boston (Massachusetts), Columbus (Ohio), Denver (Colorado), Minneapolis/Twin Cities (Minnesota), Nashville (Tennessee), New York City (New York), New Orleans (Louisiana), San Diego (California), San Francisco (California), Seattle (Washington), and Washington DC. These markets were selected based on a two-step process: first, we identified cities with varying reputations for favorable or restrictive short-term rental regulations, and second, we ensured geographic diversity by selecting one market from each regional division as defined by the US Census Bureau (2017).

For each market, data were systematically collected at six-month intervals—July 2018, January 2019, July 2019, January 2020, July 2020, and January 2021—using the scraped data

[4]In mid-2022, after the end of our observation period, Airbnb substantially changed its category system, introducing many new categories and allowing listings to be affiliated with more than one category.

from Inside Airbnb, a non-profit initiative that regularly scrapes and publishes Airbnb listing data for selected markets. This approach enabled us to track listings over time and detect any changes in their categorical positioning. By sampling all listings in each selected market, we also control for market-specific dynamics that could otherwise confound our analysis.

We applied a series of exclusion criteria to refine the dataset. Specifically, we removed listings with a minimum stay of 30 days (i.e., long-term rentals), listings by hosts with more than 10 listings (likely professional real estate companies), listings with self-descriptions under 50 characters ($N = 6334$), listings whose categories consisted of fewer than five listings per observation period ($N = 159$), listings with missing data ($N = 92$), and listings with implausible category labels ($N = 23$). We also excluded listings whose price per night exceeded the sample mean by more than 10 times ($N = 8729$). Additionally, due to the nature of our dependent variable, listings with only a single observation were also excluded. After applying these rigorous filters, the final dataset includes 111,401 unique Airbnb listings by 84,667 unique Airbnb hosts, comprising a total of 304,619 listing-period observations.

3.2 | Analyses predicting category-switching

3.2.1 | Measures

Category-switching—the dependent variable to test Hypotheses 1 and 2—is a binary variable that captures whether a listing's category affiliation in the current period differs from its category affiliation in the previous period. When hosts register their accommodation listings on Airbnb, they are required to assign their listing to one of 47 predefined categories (in response to the platform's prompt, "Which is most like your place?"). Airbnb hosts are free to change the categorization of their listing at any time; doing so disaffiliates their listing from the previously chosen category and affiliates the listing with the newly chosen category—hence, a listing was never classified into more than one category during our observation period.[5] Our repeated web-scraping surveys allowed us to accurately track each listing's affiliated accommodation category in every period and to identify category-switching events. As our dataset spans six half-year periods, each listing has up to five potential opportunities to switch categories.

Our key independent variable, *Negative reviews*, counts the number of negative customer reviews a listing received during a given period. Airbnb actively prompts customers to provide a textual review of the booked accommodation following each stay, and these reviews are prominently displayed on the listings' profiles (Zervas et al., 2021).[6] We collected time-stamped

[5]We validated that Airbnb did not eliminate any accommodation categories during our observation period to ensure that hosts deliberately disaffiliated their listing from a category rather had to do so due to a change in the category system.

[6]When writing their public review, guests are also asked to leave a star rating (one to five stars) but unlike many other online marketplaces, Airbnb did not publish individual star-ratings at the time—that is, individual guests' star ratings were neither visible to hosts nor consumers (Fradkin & Holtz, 2023; Zervas et al., 2021), only providing an aggregated rating score across all individual reviews once a listing has attracted more than three individual reviews. Due to its aggregated nature, this metric does not allow us to reliably infer the number of negative reviews in a given period. More recently, Airbnb started to publicly display the star ratings associated with each customer review for all listings across its entire marketplace (https://news.airbnb.com/airbnb-2023-winter-release). Recent research, which deploys the same sentiment analysis approach for a small sample of Airbnb listings, finds that the sentiment-based review classification highly correlates with individual ratings (Fradkin & Holtz, 2023).

review-level data from Inside Airbnb, which allowed us to reliably identify all customer reviews associated with our sample listings between 2018 and the end of the observation period, yielding 2,262,300 individual reviews. We removed reviews that were automatically posted by Airbnb due to host-initiated cancellations ($N = 7639$). On average, guest-written reviews contained 228 characters (median = 164). To identify negative reviews, we applied a sentiment analysis using a fine-tuned version of DistilBERT that was trained on Version 2 of the Stanford Sentiment Treebank (SST-2) (Socher et al., 2013) with the specific purpose of classifying online reviews as positive or negative (Fradkin & Holtz, 2023). This approach classified 3.5% of all online reviews as negative ($N = 79{,}379$)—a rate consistent with prior findings that Airbnb hosts tend to receive more favorable reviews than those on other platforms (Fradkin & Holtz, 2023; Zervas et al., 2021).[7] Because each customer review on Airbnb is time-stamped, we were able to retrospectively match each negative review to the observation period in which it appeared and calculate the number of negative reviews each listing received during each period.

Category heterogeneity captures the degree of diversity in terms of a category members' features, where low levels of category heterogeneity indicate that most listings closely align with a clear set of prototypical features. In the Airbnb context, we focus on listings' set of provided amenities as a relevant and observable set of product features. Within the Airbnb marketplace, hosts need to indicate whether their listing includes any of 54 pre-defined amenity types (e.g., kitchen, washing machine, pool, breakfast). Using time-sensitive data about each listing's offered amenities, we inferred each category's prototypical amenity profile by calculating, for each amenity, the average share of listings in a category that offered it in a given period. This yielded a 54-dimensional vector of amenity probabilities per category-period. We then captured category heterogeneity as:

$$\sum_{A=1}^{54} \sqrt{\frac{1}{N_t} \sum_{i=1}^{N_t} \left(\Theta_{A,i,t} - \overline{\Theta}_{A,c,t}\right)^2}$$

where A indexes amenities, N_t represents the number of listings in the category at time t, $\Theta_{A,i,t}$ refers to the presence of amenity A for listing i at time t, and $\overline{\Theta}_{A,C,t}$ represents the amenity share for amenity A in category c at time t. To facilitate interpretation of effect sizes, we mean-centered the variable to align the sample's medium level of category heterogeneity with zero. In robustness tests, we also tested our hypotheses with alternative measures that capture category heterogeneity in terms of the listings' descriptions or in terms of only the five most commonly occurring amenities in the given category.

To isolate the effects of our independent variables, we included a comprehensive set of control variables at five different levels: listing, host, neighborhood, market (i.e., city), and category. At the listing level, we controlled for reviews in period, prior reviews, amenities count, minimum nights, instant booking, and price premium. *Reviews in period* represents the number of reviews a listing received in the given period to account for the possibility that a negative review might become a somewhat less salient cue if a listing received many reviews during a given period. To prevent multicollinearity issues biasing our main coefficient estimates, we

[7]For robustness tests, we alternatively sentiment-analyze reviews by deploying Python's TextBlob library—which quantifies each review's conveyed sentiment on a continuum from −1.0 to 1.0, where −1.0 indicates a very negative sentiment, 0.0 represents a neutral sentiment, and 1.0 reflects a very positive sentiment—and coded each review with a value lower than 0.0 as "negative."

orthogonalized the measure with respect to negative reviews and used the orthogonalized measure in our main models. *Prior reviews* represents the number of reviews a given listing had received up until the beginning of a given period—a measure that prior research has used as a proxy for the quality of Airbnb listings (Fradkin & Holtz, 2023). This measure is based on all reviews that our sample listings received since being added to the Airbnb marketplace ($N = 3{,}641{,}436$), including in periods before our main study period. These two review-based controls allowed us to account for heterogeneity in the relative demand for listings. *Amenities count* represents the number of offered amenities. *Minimum nights* represent the minimum number of nights required for booking, and *instant booking* is a binary variable reflecting whether guests can instantly book the accommodation without prior communication with the host. *Price premium* represents the accommodation's price per night, standardized (via z-scores) to account for the price level in the given neighborhood.

At the host level, we controlled for whether the host had attained the "superhost" designation at the time (*superhost*), the total number of listings managed by the host (*host listings count*), whether the host's identity was verified at the time (*host identity verified*), as well as the host's most likely *gender* and *race*. Host gender was inferred from the gender probabilities of hosts' first names as classified via Gender-api.com. We used a gender probability threshold of 90% as a cut-off point to classify a host into the categories of "Female" or "Male," and coded all others as "not available." We inferred host race following Sood and Laohaprapanon (2018), who classified US first names using racial distribution data from voter records. Hosts were classified as "Black," "Hispanic," "White," "other," or "not available."

To account for variation in local competitive dynamics, we controlled for *density in neighborhoods*, measured as the natural logarithm (plus 1) of the number of listings in a given listing's neighborhood. At the market level, we included city fixed effects by introducing dummy variables for each of the 12 cities in our sample.

At the category level, we included measures of category density and category distinctiveness to delineate the implications of category heterogeneity (Soublière et al., 2024). *Category density* was measured as the natural logarithm (plus 1) of the number of available listings in a category period. To avoid any multicollinearity concerns between category heterogeneity and category density, we orthogonalized the two measures and used the orthogonalized measure of category density in our models. *Category distinctiveness* captured the degree to which a given accommodation category is distinctive from other accommodation categories in the Airbnb marketplace. We calculated *category distinctiveness* as the median pairwise distance between the focal category and all other categories in the same period (for details on the measurement of inter-category distances, see Section 3.3.2). For consistency, we standardized this measure and anchored its minimum value at zero.

Finally, to account for seasonal and temporal trends, we included dummy variables for year and season (summer vs. winter) across the observation period.

3.2.2 | Estimation method for models predicting category-switching

To estimate the likelihood of category-switching, we employed population-averaged logistic regression models, treating listings as the cross-sectional unit and half-years the temporal unit. Given the repeated observations of the same listing over time are unlikely to be independent, it was crucial for us to account for such within-listing correlations. Population-averaged models, a specific form of generalized estimating equations (GEE), are designed to address such correlations and correct for the violated assumption of independence between listing-specific observations (Neuhaus et al., 1991; Wooldridge, 2010). This approach represents a conceptual middle ground between fixed and random effects (Hillman et al., 2007).

We opted for population-averaged models over fixed-effects models due to the limited temporal variation in our dependent variable—most listings did not switch categories during the observation period—which would result in the exclusion of a large share of listings under a fixed effects specification. We specified an exchangeable correlation structure for the models, based on the quasi-likelihood under the independence model criterion (QIC) test statistic (Cui, 2007). This specification provided a better fit to our data than other alternatives. Moreover, GEE models remain consistent even when the correlation structure is misspecified (Liang & Zeger, 1986), and they are particularly robust in the presence of missing observations in unbalanced panels (Mannucci & Yong, 2018).

To further account for within-listing correlation and ensure accurate inference, we calculated significance levels using Huber–White robust standard errors, which are equivalent to clustering standard errors at the listing level (Wooldridge, 2010).[8]

3.3 | Analyses predicting category choice

3.3.1 | Dataset construction and measures

To test Hypotheses 3 and 4, we constructed a dataset that captures the set of alternative categories a producer could plausibly choose at the point of category-switching—a choice set including the category they chose (i.e., the realized outcome) and the categories they did not choose (i.e., non-realized outcomes). Specifically, we identified all listing periods in which a category-switch occurred and paired each of these observations with all categories other than the one the listing originated from, resulting in a dyadic database where each observation reflects a potential choice. This dyadic structure is consistent with common methodological approaches used to study producer decision making among discrete alternatives (e.g., Greve, 2000; Li et al., 2019; Pontikes & Barnett, 2017). This resulting dataset comprises 135,658 dyadic observations across 4015 listing periods and includes 35 categories (i.e., those categories with at least five listings during the period).

While we used the full choice set for robustness tests, our main analyses employed a reduced choice set to focus on more plausible alternatives. Following prior research (e.g., Zhang et al., 2017) and common recommendations suggesting a matching ratio of 1:5 between realized and non-realized outcomes (King & Zeng, 2001), we selected five alternative categories for each listing in addition to the one actually chosen. To identify these five alternatives, we created a measure of feature alignment that captures the degree of similarity between the listing's features and the prototypical features of each category. We calculated this similarity using cosine similarity between the listing's vector of offered amenities and the category's average amenity profile. For each listing, we then selected the five unchosen categories with the highest feature similarity to include in the reduced choice set.

[8]Inspired by prior research (e.g., Pontikes & Barnett, 2015), we also considered estimating category-switching events via survival models, such as Cox proportional hazards. We decided against survival models due to the complexities introduced by the left-censored and right-censored nature of our data.

To test Hypothesis 3, we constructed a measure of *inter-category proximity*, defined as the inverse of the distance between the previously chosen category and each alternative in the choice set. Inter-category distance were calculated as the cosine distance between categories' probability vectors of features: $\sum_{A=1}^{54} \cos\left(\overline{\Theta}_{A,c} - \overline{\Theta}_{A,d}\right)$, where $\overline{\Theta}_{A,c}$ represents the probability of amenity A in category c and $\overline{\Theta}_{A,d}$ refers to the probability of amenity A in category d, summed across all 54 amenities.

In addition to the fixed effects at the listing-period level, we further controlled for category density and category distinctiveness as two categorical characteristics that may influence producers' category choices.

3.3.2 | Estimation method for models predicting category choice

We used conditional logit models (McFadden, 1973) to estimate the conditional probability that a category was chosen from the choice set. Conditional logit models are similar to logit regressions with fixed effects at the level of listing periods, allowing us to effectively control for all characteristics that do not vary within a choice set. This approach allowed us to predict the probability of choosing a given category based on attributes of that category and those of all alternative categories. As such, these models only included covariates that reflect categorical characteristics in the choice set as well as our two hypothesis-testing variables. To further account for the correlations between observations belonging to the same listing, we clustered standard errors at the listing level.

4 | RESULTS

4.1 | GEE models predicting the propensity for category-switching

Table 1 reports descriptive statistics and correlations of our key variables. As expected, category-switching events are rare: we observe category switching in only 4015 listing periods, accounting for approximately 1.3% of all observations. Despite this low rate, our large sample ensures a sufficient absolute number of category-switching events to avoid statistical bias typically associated with extremely low-frequency dependent variables (see Allison, 2012; Williams, 2022).

Model 1 in Table 2 reports baseline results including only control variables. Several control variables show positive associations with category-switching, including reviews in period, prior reviews, amenities count, instant booking, host listings count, and category distinctiveness. By contrast, category density, female and male hosts (in contrast to the base category of "unknown" gender), as well as hosts with verified identity are negatively associated with category-switching.

Hypothesis 1 predicts a positive relationship between negative customer reviews and the likelihood of category-switching. Model 2 in Table 2 confirms this prediction, showing a positive and significant coefficient for negative reviews ($b = .141$; $p < .001$). Model 3, which adds an interaction term between negative reviews and category heterogeneity, also shows a positive and significant coefficient for negative reviews ($b = .172$; $p < .001$). Translating this coefficient into substantive terms, we find that each additional negative review increases the odds of category-switching by 18.8% for listings in categories with average levels of heterogeneity. These findings offer strong support for Hypothesis 1.

TABLE 1 Descriptive statistics and correlations.

Variable	Mean	SD	Min	Max	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Category-switching	0.01	0.11	0.00	1.00	1													
Negative reviews	0.20	0.68	0.00	54.00	0.04	1												
Category heterogeneity	0.00	0.45	−9.06	2.60	−0.01	−0.01	1											
Reviews in period	5.90	11.06	0.00	351.00	0.05	0.00	−0.01	1										
Prior reviews	31.23	61.96	0.00	951.00	0.02	0.17	−0.07	0.35	1									
Amenities count	19.85	7.92	0.00	47.00	0.04	0.11	0.11	0.28	0.25	1								
Minimum nights	4.36	7.13	1.00	3.00	−0.01	−0.10	−0.01	−0.14	−0.05	−0.02	1							
Instant booking	0.40	0.49	0.00	1.00	0.02	0.13	0.04	0.12	0.06	0.15	−0.07	1						
Price premium	−0.05	0.80	−2.37	27.12	−0.01	−0.05	0.06	−0.08	−0.09	0.08	−0.02	−0.02	1					
Superhost	0.34	0.47	0.00	1.00	0.03	0.07	0.00	0.37	0.35	0.43	−0.04	0.07	−0.05	1				
Host listings count	2.64	1.52	1.00	653.00	0.03	0.03	0.01	−0.01	0.00	0.05	0.02	0.03	0.00	0.03	1			
Host identity verified	0.55	0.50	0.00	1.00	−0.01	−0.05	−0.03	0.00	0.11	0.00	0.03	−0.17	−0.02	0.07	−0.01	1		
Density in neighborhood	1016.57	1206.43	2.00	4557.00	−0.03	−0.05	0.06	−0.10	−0.09	−0.14	−0.04	−0.02	−0.03	−0.15	−0.02	−0.01	1	
Category density	0.00	0.97	−6.12	0.75	−0.08	−0.04	0.02	−0.14	−0.12	−0.21	0.00	−0.06	−0.01	−0.18	−0.07	0.01	0.14	1
Category distinctiveness	0.00	1.00	−1.31	18.07	0.04	0.00	−0.72	−0.04	0.02	−0.17	0.04	−0.05	−0.07	−0.08	0.04	0.01	0.03	−0.09

Note: Descriptive statistics for *reviews in period* reflect the unchanged count measure, whereas correlations are represented for the orthogonalized measure.

TABLE 2 GEE models predicting the probability of category-switching.

	Model 1			Model 2			Model 3			Model 4		
	Coef	*p*	SE	*Coef*	*p*	SE	*Coef*	*p*	SE	*Coef*	*p*	SE
Negative reviews				0.141	0.000	0.024	0.172	0.000	0.017	0.171	0.000	0.019
Category heterogeneity				0.267	0.000	0.043	0.237	0.000	0.043	0.238	0.000	0.051
Negative reviews × Category heterogeneity							0.096	0.000	0.020	0.104	0.000	0.025
Category heterogeneity^2										−0.130	0.000	0.032
Negative reviews × Category heterogeneity^2										0.004	0.839	0.020
Reviews in period	0.078	0.000	0.014	0.103	0.000	0.013	0.102	0.000	0.013	0.101	0.000	0.013
Prior reviews	0.001	0.000	0.000	0.001	0.002	0.000	0.001	0.004	0.000	0.001	0.005	0.000
Amenities count	0.020	0.000	0.003	0.019	0.000	0.003	0.019	0.000	0.003	0.019	0.000	0.003
Minimum nights	0.003	0.198	0.003	0.004	0.129	0.003	0.004	0.104	0.003	0.004	0.095	0.003
Instant booking = 1	0.093	0.007	0.034	0.073	0.035	0.035	0.070	0.043	0.035	0.067	0.054	0.035
Price premium	−0.012	0.589	0.022	−0.018	0.415	0.023	−0.014	0.527	0.022	−0.007	0.768	0.022
Superhost = 1	0.055	0.184	0.042	0.086	0.040	0.042	0.093	0.027	0.042	0.085	0.042	0.042
Host listings count	0.005	0.000	0.001	0.004	0.000	0.001	0.004	0.000	0.001	0.004	0.000	0.001
Host identity verified = 1	−0.155	0.000	0.035	−0.136	0.000	0.035	−0.132	0.000	0.035	−0.132	0.000	0.035
Female	−0.236	0.000	0.064	−0.226	0.000	0.065	−0.226	0.000	0.065	−0.227	0.000	0.065
Male	−0.204	0.002	0.065	−0.183	0.005	0.065	−0.180	0.006	0.065	−0.183	0.005	0.065
Race: Black	0.088	0.267	0.080	0.099	0.215	0.080	0.095	0.233	0.080	0.100	0.213	0.080
Race: Hispanic	0.058	0.451	0.076	0.068	0.381	0.077	0.064	0.404	0.077	0.067	0.380	0.077
Race: White	−0.049	0.510	0.074	−0.037	0.618	0.075	−0.041	0.586	0.075	−0.037	0.624	0.075
Race: other	−0.207	0.118	0.132	−0.199	0.131	0.132	−0.201	0.127	0.132	−0.194	0.141	0.132
Density in neighborhood	−0.000	0.144	0.000	−0.000	0.095	0.000	−0.000	0.091	0.000	−0.000	0.070	0.000
Category density	−0.346	0.000	0.015	−0.322	0.000	0.016	−0.321	0.000	0.016	−0.330	0.000	0.016
Category distinctiveness	0.239	0.000	0.019	0.334	0.000	0.026	0.337	0.000	0.026	0.392	0.000	0.030

TABLE 2 (Continued)

	Model 1			Model 2			Model 3			Model 4		
	Coef	***p***	**SE**	***Coef***	***p***	**SE**	***Coef***	***p***	**SE**	***Coef***	***p***	**SE**
Seasonal dummy: Summer	−0.063	0.111	0.040	−0.062	0.119	0.040	−0.057	0.147	0.040	−0.043	0.284	0.040
Year dummy: 2019	−0.927	0.000	0.044	−0.967	0.000	0.044	−0.963	0.000	0.044	−0.931	0.000	0.044
Year dummy: 2020	−1.683	0.000	0.058	−1.651	0.000	0.060	−1.644	0.000	0.059	−1.597	0.000	0.060
Constant	−4.434	0.000	0.107	−4.499	0.000	0.108	−4.508	0.000	0.108	−4.554	0.000	0.109
Observations	305,619			305,619			305,619			305,619		
Wald chi-square	5039.28			5155.37			5221.89			5236.54		
p	.000			.000			.000			.000		

Note: N = 305,619 (111,401 listings). All models additionally include *market dummies* and are estimated with Huber–White robust standard errors.

Hypothesis 2 predicts that category heterogeneity positively moderates the relationship between negative reviews and category-switching. Model 3 supports this hypothesis, with a significant and positive coefficient of the interaction term ($b = .096$; $p < .001$). Post-estimations of the predicted category-switching propensity at different levels of negative reviews and different levels of category heterogeneity suggest that this contingency has high practical significance: a one-unit increase in negative reviews (from 0 to 1) increases the predicted category-switching probability by 8.7% if category heterogeneity is two standard deviations below the mean, by 13.2% if category heterogeneity is one standard deviation below the mean, by 22.8% if category heterogeneity is one standard deviation above the mean, and by 27.8% if category heterogeneity is two standard deviations above the mean.

Figure 1 visualizes the predicted category-switching probabilities at one standard deviation below and above the mean level of category heterogeneity. Notably, the positive association between negative reviews and category-switching becomes negligible or even negative in extremely homogeneous categories (more than 4 standard deviations below the mean value of category heterogeneity), though these cases represent only 0.12% of our sample, suggesting limited practical significance.

Beyond the hypothesized interaction, both Models 2 and 3 also reveal a positive association between category heterogeneity and category-switching: listings in more heterogeneous categories are more likely to category-switch. Post-estimations based on Model 2 suggest that a one standard deviation increase in category heterogeneity corresponds to an 18.7% increase in the likelihood of category-switching.

Model 4 in Table 2 further presents results from a model in which we additionally include a squared term of category heterogeneity (category heterogeneity^2) because prior findings about category heterogeneity (Soublière et al., 2024) may imply a non-linear association between category heterogeneity and category-switching. Post-estimating and plotting marginal effects of negative customer reviews at low and high levels of category coherence based on Model 4 (instead

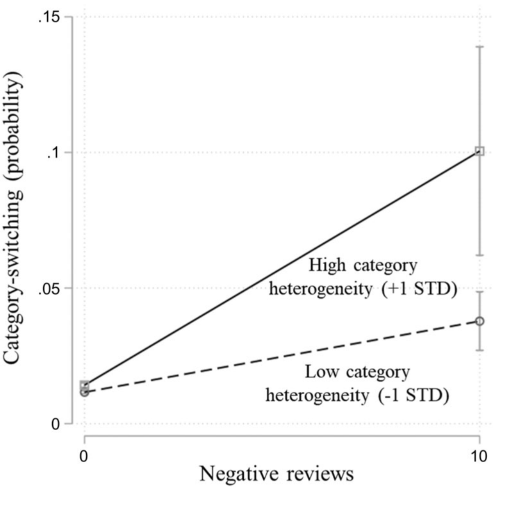

FIGURE 1 Predicted probability of category-switching. The figure plots the predicted propensity of category-switching at varying levels of negative reviews and low and high levels of category heterogeneity. We limit the x-axis to 10 negative reviews because only three observations exhibited more than 10 negative reviews in a given period.

of Model 3) yield very similar results to those presented in Figure 1. Standard errors and p-value of the interaction term between negative reviews and category heterogeneity^2 further reject the alternative hypothesis that category heterogeneity would moderate the relationship between negative reviews and category-switching in a non-linear way. Moreover, plotting the marginal effect of category heterogeneity on category-switching suggests that category heterogeneity has a strictly positive association with category-switching, although at a slightly decreasing rate (i.e., less positive at higher levels of category heterogeneity). A formal test, using STATA's *utest* command (Lind & Mehlum, 2007) confirms that the relationship does not follow an (inverted) U-shape.

4.2 | Conditional logit models predicting category choice

Table 3 reports descriptive statistics and correlations for our dyad-level data used in the category choice analysis. As expected for choice set designs, moderate correlations exist between several variables due to the way the dataset is constructed (e.g., Li et al., 2019). Variance inflation factors remain below 10 for all variables, suggesting that multicollinearity is not a problem in our models (Chatterjee & Hadi, 2012).

Hypothesis 3 proposes that the likelihood of selecting a category increases with its proximity to the previously chosen category. Model 6 in Table 4 supports this hypothesis, with a significant positive coefficient for inter-category proximity ($b = .652$; $p < .001$). To interpret coefficients from conditional logit models, we followed previous research (e.g., Li et al., 2019) and calculated the average elasticity of the probability of category choice with respect to each independent variable, adjusted by the factor $\frac{C-1}{C}$, where C is the total number of choices, leading to an adjustment factor of 0.83 for a choice set of six categories. The adjusted average elasticity of category choice probability suggests that a standard deviation increase in inter-category proximity between the current category and given category increases the likelihood of choosing that given category by 30.3%.

Hypothesis 4 proposes that the positive effect of inter-category proximity is amplified under higher levels of category heterogeneity. Model 7 in Table 4 confirms this interaction effect, showing a positive and significant coefficient ($b = 0.118$; $p < .001$). As illustrated in Figure 2, a one standard deviation increase in inter-category proximity raises the likelihood of selecting a category by 94.8% (from 0.18 to 0.35) if category heterogeneity is at one standard deviation above the mean, compared to 25.2% (from 0.12 to 0.15) if category heterogeneity is one standard

TABLE 3 Descriptive statistics and correlations for the dyad-level dataset.

		Mean	SD	Min	Max	1	2	3	4	5
1	Category choice	0.18	0.38	0.00	1.00	1.00				
2	Inter-category proximity	0.46	0.70	−6.34	1.34	0.23	1.00			
3	Category heterogeneity	0.13	1.49	−18.13	3.22	0.14	0.52	1.00		
4	Category density	−2.81	1.94	−6.28	0.75	0.48	0.42	0.37	1.00	
5	Category distinctiveness	1.79	1.46	0.00	20.45	−0.23	−0.69	−0.68	−0.46	1.00

Note: The table represents descriptive statistics and correlations for the dyad-level dataset, consisting of 21,329 observations across 4015 listing periods.

TABLE 4 Conditional logit models predicting category choice.

	Model 5			Model 6			Model 7			Model 8		
	Coef	***p***	**SE**	***Coef***	***p***	**SE**	***Coef***	***p***	**SE**	***Coef***	***p***	**SE**
Inter-category proximity				0.652	0.000	0.084	0.623	0.000	0.083	0.524	0.000	0.085
Category heterogeneity				0.462	0.000	0.060	0.420	0.000	0.052	0.250	0.003	0.085
Inter-category proximity × Category heterogeneity							0.118	0.000	0.009	0.379	0.000	0.060
Category heterogeneity^2										−0.030	0.462	0.041
Inter-category proximity × Category heterogeneity^2										0.025	0.005	0.009
Listing-period fixed effects	Included			Included			Included			Included		
Category density	0.858	0.000	0.018	1.097	0.000	0.026	1.095	0.000	0.025	1.088	0.000	0.025
Category distinctiveness	−0.160	0.000	0.037	0.282	0.000	0.046	0.257	0.000	0.044	0.212	0.000	0.045
Observations	21,329			21,329			21,329			21,329		
Wald chi-square	3094.54			2496.20			2520.22			2531.02		
R^2	0.427			0.534			0.536			0.538		
p	0.000			0.000			0.000			0.000		

Note: N = 21,329 dyads (4015 listing-periods). Standard errors are clustered by listing.

deviation below the mean. In addition, both Models 6 and 7 show that category heterogeneity is positively associated with the likelihood of category selection.

Model 8 adds a squared term for category heterogeneity (category heterogeneity^2) to test for non-linear effects. Post-estimations suggest that the association between category heterogeneity and the selection likelihood is strictly positive, although at a slightly increasing rate. This finding adds further weight to our argument that categories with higher heterogeneity will more likely be perceived as viable options by non-members. Model 8 also confirms the hypothesized interaction effect when accounting for a potential non-linearity in the association between category heterogeneity and category selection. A replication of Figure 2 based on Model 8 shows that the interaction effect is even more pronounced when accounting for this potential non-linearity.

4.3 | Robustness tests

We conducted multiple robustness checks to ensure the validity of our findings. First, we verified that the results for Hypotheses 1 and 2 hold under alternative measures of negative reviews and category heterogeneity. Our findings are consistent when the number of negative reviews (plus 1) is log-transformed to correct for the skewed distribution of this measure. Our findings also replicate if we infer the sentiment of online reviews by deploying Python's TextBlob library, including for various sentiment thresholds.[9] For category heterogeneity, we alternatively inferred categories' prototypical amenities from the category's median value for the given amenity (i.e., whether the majority of listings in the category offer the amenity or not), and then followed our previously described approach to measure category heterogeneity based on amenity vectors with binary values. Our sample exhibits very little variance in terms of certain amenities (e.g., almost all listings offer internet), and we therefore also constructed an alternative measure that constructs category heterogeneity only based on those amenities for which we observe at least a certain amount of variation in the sample (using a sample-wide standard deviation of 0.1 as a minimum degree of variation, which disregards 12 amenities). We further developed alternative measures based on listings' textual descriptions rather than their associated amenities. Hence, these measures reflect categories' heterogeneity in a semantic space rather than a feature space. We followed previous research (Haans, 2019; Taeuscher et al., 2022) and used a topic modeling approach to infer common topics used across all textual descriptions (using a model with 25 topics representing a 25-dimensional semantic space). We then inferred each listing's position within this semantic space via the degree to which its self-description draws on each of these topics (i.e., the respective topic weights of the given description), and calculated a vector of average topic probabilities for each category-period to represent the category's prototypical content at the time. These topic probability vectors allowed us to

[9]This approach to sentiment analysis quantifies each review's conveyed sentiment on a continuum from −1.0 to 1.0, where −1.0 indicates a very negative sentiment, 0.0 represents a neutral sentiment, and 1.0 reflects a very positive sentiment. We coded each review with a value lower than 0.0 as "negative"—this approach classified 1.1% of online reviews as negative ($N = 25{,}272$). When using this alternative measure of negative reviews, the coefficient of our baseline relationship becomes more pronounced (0.15). Using a more restrictive threshold for classifying negative reviews (e.g., a sentiment score of −0.1) further reduces the total number of negative reviews in the sample. Results using this measure fully support Hypotheses 1 and 2, further augmenting the effect size of the hypothesized relationships.

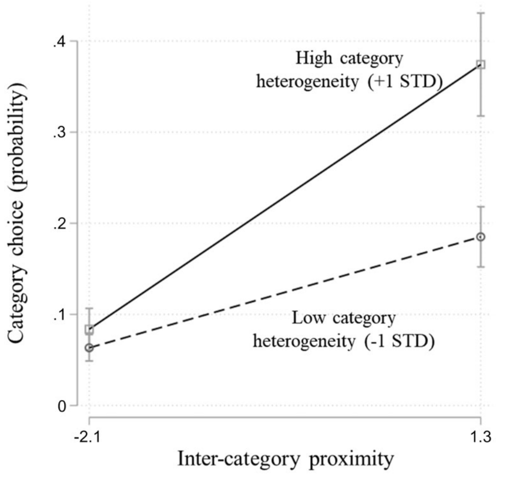

FIGURE 2 Predicted probability of category choice. The figure plots the predicted propensity of category choice at various empirically observed levels of inter-category proximity and at one standard deviation below and above the mean level of category heterogeneity.

create our measure of category heterogeneity as described above. Our results remain qualitatively unaltered when using any of these alternative measures.

Second, to address the potential endogeneity concern that category heterogeneity does not only affect the relationship between negative reviews and category-switching but also systematically influences the likelihood that a listing receives negative reviews in the first place, we implemented an instrumental variable approach. Specifically, we instrumented negative reviews using the average star rating across all listings in the focal listing's neighborhood. Listings in a neighborhood with low average ratings should receive, on average, a higher number of negative reviews, but a neighborhood's average rating should not directly affect hosts' propensity to category-switch. In the first stage, we regressed a listing's number of negative reviews on this instrument and all other covariates, using STATA's xtnbreg command. In the second stage, we used the predicted number of negative reviews and covariates to predict category-switching. First-stage regression confirms the relevance and strength of the instrument ($b = -2.19$; $p < .001$; $F = 296.97$). The instrument also did not correlate significantly with the error term of category-switching—the second criterion for effective instruments (Wooldridge, 2010). A Hausman test indicated no significant differences between instrumented and non-instrumented models, suggesting that endogeneity is not a major concern in our model. Results from models using a control function approach—which adds residuals from the first stage model to the second stage models (Wooldridge, 2015)—are also consistent with those of non-instrumented models, thereby further supporting the assumption that our main models likely do not suffer from endogeneity issues. Moreover, we also generated very consistent results when specifying random-effects logistic models (using STATA's *xtlogit* command), either with or without fixed effects for categories, and when estimating mixed-effects logistic models (using STATA's *melogit* command) with random intercepts at the listing and category levels.

Third, our results are robust to the inclusion of additional controls that moderately correlate with at least one of our included variables. Examples include the number of bedrooms and bathrooms (proxies for accommodation size) and categorical variables reflecting listings' cancellation policy and room type (entire unit, private room, shared unit, or shared room). Most importantly, we validated that our findings are consistent if we further controlled for a listing's average rating as publicly displayed on the listing's profile at the time. Airbnb only publicly displays a listing's average rating once it has been rated by at least three customers. Including the measure would therefore lead to a significant loss of observations, and we thus did not include the measure in our main models. Models based on the sub-sample for which an average rating is available ($N = 253{,}058$) fully confirmed our hypothesis-testing findings. Across these models, we find that a listing's average rating has no statistically significant association with category-switching. This finding supports our assumption that receiving a negative review represents a meaningful event in that it can trigger a search process—not merely reflecting a listing's low quality. In further models, we transformed average star ratings into a categorical variable with four levels (none, low, moderate, and high) based on ratings' percentile distribution to account for the fact that Airbnb ratings tend to be heavily right-skewed (the median average star rating across our sample is 4.9 out of 5). Models that control for this categorical variable fully confirmed our findings.

For the conditional logit models, results remain robust when clustering standard errors at the listing-period level or using standard errors double-clustered by both listing and period. Findings are also consistent when using the full choice set (i.e., considering all other 34 categories as potential choices) or when randomly sampling five alternative categories in addition to the chosen category. In alternative models, we also included a measure of the mean category choice probability across all observations of a given listing (except for the given dyad observation), as advocated by Lincoln (1984) and sometimes used to further account for the non-independence of observations within dyads (e.g., Pontikes & Barnett, 2017). Our findings are also qualitatively unaltered when we additionally controlled for a category's average price premium and average number of accumulated reviews across all listings in that category, accounting for potential differences in the demand for different categories. We also confirmed that our findings do not change if we only included the first category-switch for the small number of listings ($N = 200$) that experienced more than one category-switch during our observation period. Jointly, these analyses add further confidence in our main results.

4.4 | Supplemental analyses

Our empirical setting and data allow us to further probe our theorized mechanism and rule out alternative explanations. We report these supplemental analyses in the Supporting Information.

First, we examined variations in the content of negative reviews to further probe our assumptions about the role of negative reviews as expectation violation cues. Specifically, we aimed to probe our assumptions that negative reviews should generally enhance producers' propensity to category-switch, while this effect should be strongest for negative reviews that explicitly articulate feature-related expectation violations. Appendix A1 reports our methodological approach, which leverages the large language model GPT-4.1 nano to identify such negative reviews, and Appendix A2 reports findings from these analyses. The findings suggest that the effect of negative reviews on category-switching is 65.0% larger for negative reviews that explicitly indicate feature-related expectation violations compared to other types of negative reviews (e.g., those indicating quality issues), with formal test statistics suggesting that these effect size differences are statistically meaningful ($\chi^2 = 6.58$, $p = .010$). Yet, this analysis also confirms that different types of negative reviews, including those that do not explicitly mention any feature-related expectation violations, are positively associated with category-switching and positively interact with category heterogeneity—thus offering further support for our Hypotheses 1 and 2.

Second, we also aimed to probe the potential impact of non-negative reviews—that is, customer reviews that were not classified as negative in our sentiment analysis—on category-switching. These analyses, reported in Appendix A3, provide further support for our theoretical assumptions about expectation violation cues. These models show a slightly positive association between non-negative reviews and category-switching, but postestimations suggest that the effect size of negative reviews is 9.2 times larger (820% stronger) than that of non-negative reviews. The effect size of feature-related negative reviews (see Appendix A1) is 15.8 times larger (1480% stronger) than that of non-negative reviews. A slightly positive association between non-negative reviews and category-switching could indicate that even some non-negative reviews can act as expectation violation cues—for example, if a review has an overall positive sentiment but mentions some mismatch between what the guest expected and experienced (e.g., "We loved everything about Tony's place even if it isn't really a tiny house")—but it is also possible that non-negative reviews may trigger a slightly different mechanism leading to category-switching. However, additional models presented in Appendix A4 suggest that the positive association between non-negative reviews and category-switching disappears when excluding listing periods with zero reviews from our analysis. Jointly, these supplemental analyses strongly support our theoretical account of expectation violation cues as the primary driver of category-switching.

Appendix A5 further reports on supplemental analyses that attend to the heterogeneity of customers—specifically as it relates to customers' experiences as Airbnb guests prior to leaving the given review. Appendix A6 reports supplemental analyses that probe alternative explanations for category-switching as indicated in prior research. These analyses add further empirical support for our theoretical account and help rule out alternative explanations for category-switching in our context.

5 | DISCUSSION

Our study explored how negative customer reviews, as a form of expectation violation cue, influence producers' propensity for category-switching. Drawing on data from the Airbnb marketplace, we found strong evidence that negative customer reviews significantly increase the likelihood of category-switching and that this relationship is moderated by the category's heterogeneity—that is, producers are particularly prone to switch categories in response to negative reviews when their offerings are situated in heterogeneous categories. These findings, combined with the low baseline propensity of category-switching, suggest that capacity-constrained producers rarely attend to their categorical positioning unless prompted by such cues. Further analyses revealed that when producers do switch categories, they tend to select new categories based on their proximity to the originally chosen category and are more likely to select heterogeneous than homogenous categories. These findings have important implications for research on strategic categorization and the dynamics of category membership.

5.1 | Contributions to theory

Our study primarily advances categorization research as it relates to individuals' and organizations' self-categorization efforts. Prior studies have primarily focused on how organizations initially select categories (Aversa et al., 2021; Gehman & Grimes, 2017; Granqvist et al., 2013) or how categorical characteristics influence market outcomes (Pontikes & Barnett, 2015; Pontikes & Kim, 2017; Soublière et al., 2024; Zunino et al., 2019). Our study extends the scope of strategic categorization research by highlighting category-switching as a distinct and theoretically important phenomenon whose underlying mechanisms likely differ from those that inform producers' initial category choice.

Our theoretical account expands the notion of strategic categorization by showing that capacity-constrained producers often change the categorization of their offerings reactively in response to cues indicating a violation of audience expectations. While this account recognizes producers' agency in their categorization, it challenges the prevailing assumption that category-switching and other forms of strategic categorization necessarily result from proactive optimization efforts. Taking seriously that decision-makers have limited attentional capacities (March, 1994), our theorization suggests that categorization-switching may represent a more reactive behavior than previous theory would suggest, especially for individual entrepreneurs and small businesses with limited capacities. We thus introduce the notion of capacity-constrained producers as a boundary condition under which the assumption of continuous, proactive optimization may not hold, thus opening new avenues for theorizing about producer behavior in categorization contexts.

By introducing the notion of *expectation violation cues*, our theoretical account advances understanding about the triggers and events that may cause producers to revisit and potentially revise their category choices. Drawing on key insights from problematic search, our theoretical account highlights the important role of expectation violation cues by articulating how such salient pieces of information can draw attention to the otherwise unattended issue of categorization. Violation expectation cues, we argued, can trigger a search process that may result in a category switch if the producer identifies their categorical positioning as a potential source of the seeming expectation violation (during the problem search stage) and identifies category-switching as the preferred solution to prevent this problem in the future (during the solution search stage). By highlighting the important role of attention triggers and outlining the search processes that may lead producers to category-switch, our theoretical account thus complements prior theoretical accounts that primarily attended to the economic mechanisms that may explain category-switching.

Our study also contributes to understanding the consequences of category heterogeneity (Haans, 2019; Lo et al., 2020; Soublière et al., 2024) by uncovering a paradoxical dynamic: while heterogeneity makes a category seem more inviting to non-members, it may simultaneously increase the likelihood that existing category members will switch out of the category upon encountering expectation violation cues. This apparent paradox reflects differences in the evaluative aspects of these categories that might be most cognitively salient to non-members versus those category members who have gained a refined understanding of customer expectations through expectation violation cues and other meaningful feedback. That is, being confronted with varied forms of customer expectation violations might gradually shift producers' attention from a heterogeneous category's seeming accommodativeness (inferred from diverse offerings) toward its expectation complexity (experientially learned through violation expectation cues). As members of a heterogeneous category become aware of customers' diverse and potentially conflicting expectations about the category, they will also find it more difficult to prevent future expectation violations through closer alignment with customer expectations and will therefore be more likely to choose category-switching.

Our study findings suggest that experiential learning about a heterogeneous category's evaluative complexity increases producers' likelihood of switching out of a category, but that such category-specific learning does not necessarily lead producers to make inferences about heterogeneous categories in general. As boundedly rational actors, capacity-constrained producers may therefore end up selecting a category that equally exhibits relatively high heterogeneity—even if such a choice may not seem entirely rational. This perspective complements and extends prior work, which has primarily emphasized how category heterogeneity shapes categories' valuation and classification benefits (Soublière et al., 2024) or differentiation opportunities (Haans, 2019).

Our insights about category heterogeneity also question the conventional wisdom that members of heterogeneous categories have an easier time conforming to audience demands than those in homogenous categories (e.g., Haans, 2019). We suggest that the lack of clear categorical expectations may, in fact, present novel challenges for producers—complicating their efforts to prevent further expectation violations. As such, our study questions the characterization of heterogeneous categories as domains in which members easily conform to audience expectations due to the lack of consensus around these expectations. Similarly, our underlying arguments could also inform future research on other categorical characteristics—such as categorical contrast (e.g., Carnabuci et al., 2015; Kovács & Hannan, 2010)—that influence and/or reflect consensus about categorical expectations.

Moreover, although our setting involved producers affiliating their offering with a single category, our findings offer important insights relevant to multi-categorization contexts. Prior studies have shown that category leniency (Pontikes & Barnett, 2015)—the degree to which members of a given category tend to also affiliate with other categories—or increased competition (Pontikes & Kim, 2017) can prompt exit from certain categories. While producers in our context cannot span multiple categories, our findings complement prior research on category exits by identifying a qualitatively different exit mechanism: perceived expectation violations. Moreover, our findings problematize the assumptions underlying prior arguments in questioning whether all producer types would gain awareness of subtle shifts in categories' relative level of competition and similar economic dynamics. By going beyond category-level conditions, our study contributes to a more nuanced understanding of whether and when producers decide to disaffiliate with a category, even in multi-categorization contexts.

Our theoretical account and central findings can also advance research concerned with the performance implications of producers' varied categorical trajectories over time. Leung (2014), for example, theorized that audiences infer commitment and competence from producers' history of category choices and would therefore respond most favorably to producers that exhibit a moderate degree of category-switching over those that never or frequently switched categories. Our perspective offers a complementary explanation in that a trajectory of several category-switches may also be indicative of a series of problemistic search processes, where offerings that underwent several category-switches may, on average, exhibit a higher degree of categorical alignment than those that never switched categories. By offering insights into the events and causes that may have led producers to switch categories in the past, our study thus also helps scholars to better understand why certain categorization trajectories are associated with more favorable outcomes than others.

More broadly, our theoretical account can inform research on category-level dynamics (e.g., Grodal et al., 2015; Lo et al., 2020; Soublière et al., 2024) as it illuminates micro-level processes that likely contribute to categories' growth or decline over time. Our finding that producers are significantly less likely to switch out of very homogeneous categories can help to explain why such categories tend to be relatively persistent over time, even as homogeneous categories might attract fewer new entrants. More research is required to further uncover how different categorical conditions shape members' propensity to category-switch and how such member-level decisions can explain outcomes of category growth, decline, and demise.

Finally, our study contributes to research on organizational responses to violated audience expectations (e.g., Durand & Vergne, 2015; Dutton & Dukerich, 1991; Elsbach & Kramer, 1996; Westphal & Graebner, 2010) by situating category-switching as one such response. In doing so, we highlight the agency producers possess in managing categorical fit—not just through symbolic actions or feature changes that convey conformity with the given category, but also by changing the categorical positioning itself. As such, we believe that our study can bring categorization more strongly to the fore and pave the way for novel research avenues at the intersection of these literatures. Among others, we believe that future research on the varied tactics through which organizations convey their categorical conformity to gain legitimacy (e.g., Glynn & Abzug, 2002; Zott & Huy, 2007), research that explores the tensions between categorical conformity and distinctiveness (e.g., Navis & Glynn, 2011; Taeuscher et al., 2021; Taeuscher et al., 2022; Zhao, 2022; Zhao et al., 2017), as well as research related to organizational (a)typicality (see Cutolo & Ferriani, 2024a) stand to gain from acknowledging the agency and flexibility organizations have in responding to categorical expectations. For instance, the insight that organizations can change the categorization of their offerings to address a seeming expectation violation may present an intriguing complication for research on optimal distinctiveness, which typically assumes organizations' category membership to be exogenously given. Future research on categorization, identity, and strategic positioning may benefit from recognizing category-switching as a dynamic, intentional strategy that organizations can strategically deploy.

5.2 | Generalizability and limitations

Our arguments and findings are most applicable to contexts characterized by relatively stable categorization systems, where producers affiliate with one single category and where those affiliations are both visible and meaningful to audiences. Our findings may generalize less to settings where categories are highly fluid, rapidly fading, or experiencing constant shifts in audience attention and value. Under such conditions, category-switching may also commonly be driven by the emergence of a hot new category (Pontikes & Barnett, 2017) or the domain's convergence toward one dominant category (Grodal et al., 2015). Acknowledging the context-specific nature of categorization processes, we hope to encourage future research to build on our arguments and explore the causes of category-switching in other settings.

An important strength of our empirical context is that it allows for clean observation of category-switching, since producers must affiliate with one and only one category. However, a corresponding limitation is our inability to fully disentangle the exit decision (to leave the current category) from the entry decision (to select a new category). Our qualitative insights and category choice analyses suggest that category-switching in the Airbnb marketplace is primarily driven by the decision to exit the current category rather than the decision to enter a new category, but we cannot completely isolate these decisions empirically. Similarly, while the Airbnb marketplace provides a strong setting for studying categorization, we are unable to observe how categorization shapes consumer expectations about a given accommodation listing, which remains an inferential assumption in our framework. In our study setting, no listing switched category more than three times—thus, our study should most precisely reflect producers' initial category-switching decisions but might not fully capture longer-term learning processes that may ultimately lead producers to strategically avoid heterogeneous categories after having experienced the complexities of several such categories.

In addition, while negative customer reviews are a salient and measurable type of expectation violation cues in online marketplaces, other types of cues may play a similar role in other domains. We encourage future research to test our theory in different empirical settings, including those with more complex or informal feedback mechanisms. Finally, while our study provides compelling empirical evidence in support of the hypothesized relationships, our data do not allow us to make definitive causal claims or directly observe the micro-level mechanisms underlying producers' decisions. We encourage future studies to complement our large-scale longitudinal study with qualitative or experimental methods that more directly capture the cognitive processes that shape producers' interpretation of expectation violation cues and categorization choices.

ACKNOWLEDGMENTS

We are very grateful for the highly constructive feedback we received at the Cultural Entrepreneurship track at EGOS 2024, including from Rudy Durand, Eunice Rhee, Christian Hampel, Jean-François Soublière, and Dev Jennings, among others. We also would like to thank Gino Cattani for his excellent editorial guidance.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

Research data are not shared.

ORCID

Karl Taeuscher https://orcid.org/0000-0003-0997-324X
Eric Yanfei Zhao https://orcid.org/0000-0002-3012-5926
Michael Lounsbury https://orcid.org/0000-0002-1234-0972

REFERENCES

Airbnb. (2018, February 2). *Airbnb community center*. https://community.withairbnb.com

Airbnb. (2023). *Form 10-K 2022*. https://www.sec.gov/Archives/edgar/data/1559720/000155972021000010/airbnb-10k.htm

Allison, P. (2012). *Logistic regression for rare events*. Statistical Horizons.

Aversa, P., Huyghe, A., & Bonadio, G. (2021). First impressions stick: Market entry strategies and category priming in the digital domain. *Journal of Management Studies*, *58*(7), 1721–1760. https://doi.org/10.1111/joms.12712

Bundy, J., Shropshire, C., & Buchholtz, A. K. (2013). Strategic cognition and issue salience: Toward an explanation of firm responsiveness to stakeholder concerns. *Academy of Management Review*, *38*(3), 352–376. https://doi.org/10.5465/amr.2011.0179

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.

Cattani, G., Ferriani, S., Negro, G., & Perretti, F. (2008). The structure of consensus: Network ties, legitimation, and exit rates of U.S. feature film producer organizations. *Administrative Science Quarterly*, *53*(1), 145–182. https://doi.org/10.2189/asqu.53.1.145

Cattani, G., Porac, J. F., & Thomas, H. (2017). Categories and competition. *Strategic Management Journal*, *38*(1), 64–92. https://doi.org/10.1002/smj.2591

Chatterjee, S., & Hadi, A. S. (2012). *Regression analysis by example* (5th ed.). Wiley.
Cui, J. (2007). QIC program and model selection in GEE analyses. *The Stata Journal*, *7*(2), 209–220. https://doi.org/10.22004/ag.Econ.119269
Cutolo, D., & Ferriani, S. (2024a). Atypicality: Toward an integrative framework in organizational and market settings. *The Academy of Management Annals*, *18*(1), 157–209. https://doi.org/10.5465/annals.2022.0005
Cutolo, D., & Ferriani, S. (2024b). Now it makes more sense: How narratives can help atypical actors increase market appeal. *Journal of Management*, *50*(5), 1599–1642. https://doi.org/10.1177/01492063231151637
Cyert, R. M., & March, J. G. (1963). *A behavioral theory of the firm*. Prentice-Hall.
Dellarocas, C. (2003). The digitization of word of mouth: Promise and challenges of online feedback mechanisms. *Management Science*, *49*(10), 1407–1424. https://doi.org/10.1287/mnsc.49.10.1407.17308
Dellarocas, C., Zhang, X., & Awad, N. F. (2007). Exploring the value of online product reviews in forecasting sales: The case of motion pictures. *Journal of Interactive Marketing*, *21*(4), 23–45. https://doi.org/10.1002/dir.20087
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. https://doi.org/10.1111/j.1467-6486.2011.01039.x
Durand, R., & Thornton, P. H. (2018). Categorizing institutional logics, institutionalizing categories: A review of two literatures. *Academy of Management Annals*, *12*(2), 631–658. https://doi.org/10.5465/annals.2016.0089
Durand, R., & Vergne, J.-P. (2015). Asset divestment as a response to media attacks in stigmatized industries. *Strategic Management Journal*, *36*(8), 1205–1223. https://doi.org/10.1002/smj.2280
Durand, R., Hawn, O., & Ioannou, I. (2019). Willing and able: A general model of organizational responses to normative pressures. *Academy of Management Review*, *44*(2), 299–320. https://doi.org/10.5465/amr.2016.0107
Dutton, J. E., & Dukerich, J. M. (1991). Keeping an eye on the mirror: Image and identity in organizational adaptation. *Academy of Management Journal*, *34*(3), 517–554. https://doi.org/10.5465/256405
Elsbach, K. D., & Kramer, R. M. (1996). Members' responses to organizational identity threats: Encountering and countering the business week rankings. *Administrative Science Quarterly*, *41*(3), 442–476.
Fiss, P. C., & Zajac, E. J. (2006). The symbolic management of strategic change: Sensegiving via framing and decoupling. *Academy of Management Journal*, *49*(6), 1173–1193. https://doi.org/10.5465/AMJ.2006.23478255
Fradkin, A., & Holtz, D. (2023). Do incentives to review help the market? Evidence from a field experiment on Airbnb. *Marketing Science*, *42*(5), 853–865. https://doi.org/10.1287/mksc.2023.1439
Gavetti, G., & Levinthal, D. (2000). Looking forward and looking backward: Cognitive and experiential search. *Administrative Science Quarterly*, *45*(1), 113–137. https://doi.org/10.2307/2666981
Gehman, J., & Grimes, M. (2017). Hidden badge of honor: How contextual distinctiveness affects category promotion among certified B corporations. *Academy of Management Journal*, *60*(6), 2294–2320. https://doi.org/10.5465/amj.2015.0416
Glynn, M. A., & Abzug, R. (2002). Institutionalizing identity: Symbolic isomorphism and organizational names. *Academy of Management Journal*, *45*(1), 267–280. https://doi.org/10.2307/3069296
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. https://doi.org/10.1111/joms.12023
Granqvist, N., Grodal, S., & Woolley, J. L. (2013). Hedging your bets: Explaining executives' market labeling strategies in nanotechnology. *Organization Science*, *24*(2), 395–413. https://doi.org/10.1287/orsc.1120.0748
Greve, H. R. (2000). Marketing niche entry decisions: Competition, learning, and strategy in Tokyo banking, 1894–1936. *Academy of Management Journal*, *43*(5), 816–836. https://doi.org/10.5465/1556412
Greve, H. R. (2003). A behavioral theory of R&D expenditures and innovations: Evidence from shipbuilding. *Academy of Management Journal*, *46*(6), 685–702. https://doi.org/10.5465/30040661
Grodal, S., Gotsopoulos, A., & Suarez, F. F. (2015). The coevolution of technologies and categories during industry emergence. *Academy of Management Review*, *40*(3), 423–445. https://doi.org/10.5465/amr.2013.0359
Haans, R. F. J. (2019). What's the value of being different when everyone is? The effects of distinctiveness on performance in homogeneous versus heterogeneous categories. *Strategic Management Journal*, *40*(1), 3–27. https://doi.org/10.1002/smj.2978
Hannan, M. T., Le Mens, G., Hsu, G., Kovács, B., Negro, G., Pólos, L., & Sharkey, A. J. (2019). *Concepts and categories: Foundations for sociological and cultural analysis*. Columbia University Press.

Hillman, A. J., Shropshire, C., & Cannella, A. A. (2007). Organizational predictors of women on corporate boards. *Academy of Management Journal*, *50*(4), 941–952. https://doi.org/10.5465/amj.2007.26279222

Kennedy, M. T., Lo, J., & Lounsbury, M. (2010). Category currency: The changing value of conformity as a function of ongoing meaning construction. In G. Hsu, G. Negro, & Ö. Koçak (Eds.), *Categories in markets* (1st ed., pp. 369–397). Emerald Group.

King, G., & Zeng, L. (2001). Logistic regression in rare events data. *Political Analysis*, *9*(2), 137–163. https://doi.org/10.1093/oxfordjournals.pan.a004868

Kodeih, F., Bouchikhi, H., & Gauthier, V. (2019). Competing through categorization: Product- and audience-centric strategies in an evolving categorical structure. *Organization Studies*, *40*(7), 995–1023. https://doi.org/10.1177/0170840618772597

Kovács, B., & Hannan, M. T. (2010). The consequences of category spanning depend on contrast. In G. Hsu, G. Negro, & Ö. Koçak (Eds.), *Categories in markets: Origins and evolution. Categories in markets: Origins and evolution* (Vol. 31, pp. 175–201). Emerald Group. https://doi.org/10.1108/S0733-558X(2010)0000031008

Kovács, B., Hsu, G., & Sharkey, A. (2024). The stickiness of category labels: Audience perception and evaluation of producer repositioning in creative markets. *Management Science*, *70*(9), 5627–6482. https://doi.org/10.1287/mnsc.2021.02070

Leung, M. D. (2014). Dilettante or renaissance person? How the order of job experiences affects hiring in an external labor market1. *American Sociological Review*, *79*(1), 136–158. https://doi.org/10.1177/0003122413518638

Li, Q., Maggitti, P. G., Smith, K. G., Tesluk, P. E., & Katila, R. (2013). Top management attention to innovation: The role of search selection and intensity in new product introductions. *Academy of Management Journal*, *56*(3), 893–916. https://doi.org/10.5465/amj.2010.0844

Li, Y., Hernandez, E., & Gwon, S. (2019). When do ethnic communities affect foreign location choice? Dual entry strategies of Korean banks in China. *Academy of Management Journal*, *62*(1), 172–195. https://doi.org/10.5465/amj.2017.0275

Liang, K.-Y., & Zeger, S. L. (1986). Longitudinal data analysis using generalized linear models. *Biometrika*, *73*(1), 13–22. https://doi.org/10.2307/2336267

Lincoln, J. R. (1984). Analyzing relations in dyads: Problems, models, and an application to interorganizational research. *Sociological Methods & Research*, *13*(1), 45–76. https://doi.org/10.1177/0049124184013001002

Lind, J. T., & Mehlum, T. (2007). *UTEST: Stata module to test for a U-shaped relationship. Statistical software components*. Boston College Department of Economics.

Lo, J. Y.-C., Fiss, P. C., Rhee, E. Y., & Kennedy, M. T. (2020). Category viability: Balanced levels of coherence and distinctiveness. *Academy of Management Review*, *45*(1), 85–108. https://doi.org/10.5465/amr.2017.0011

Lounsbury, M., & Glynn, M. A. (2001). Cultural entrepreneurship: Stories, legitimacy, and the acquisition of resources. *Strategic Management Journal*, *22*(6-7), 545–564. https://doi.org/10.1002/smj.188

Lounsbury, M., & Glynn, M. A. (2019). *Cultural entrepreneurship: A new agenda for the study of entrepreneurial processes and possibilities*. Cambridge University Press.

Mannucci, P. V., & Yong, K. (2018). The differential impact of knowledge depth and knowledge breadth on creativity over individual careers. *Academy of Management Journal*, *61*(5), 1741–1763. https://doi.org/10.5465/amj.2016.0529

March, J. G. (1994). *A primer on decision making: How decisions happen*. Free Press.

March, J. G., & Simon, H. A. (1958). *Organizations*. Wiley.

McFadden, D. L. (1973). Conditional logit analysis of qualitative choice behavior. In P. Zarembka (Ed.), *Frontiers in econometrics* (pp. 105–142). Academic Press.

Mudambi, S. M., & Schuff, D. (2010). Research note: What makes a helpful online review? A study of customer reviews on Amazon.com. *MIS Quarterly*, *34*(1), 185–200.

Navis, C., & Glynn, M. A. (2011). Legitimate distinctiveness and the entrepreneurial identity: Influence on investor judgments of new venture plausibility. *Academy of Management Review*, *36*(3), 479–499. https://doi.org/10.5465/amr.2008.0361

Neuhaus, J. M., Kalbfleisch, J. D., & Hauck, W. W. (1991). A comparison of cluster-specific and population-averaged approaches for analyzing correlated binary data. *International Statistical Review*, *59*(1), 25. https://doi.org/10.2307/1403572

Ocasio, W. (2011). Attention to attention. *Organization Science*, *22*(5), 1286–1296. https://doi.org/10.1287/orsc.1100.0602

Parasuraman, A., Zeithaml, V. A., & Berry, L. L. (1985). A conceptual model of service quality and its implications for future research. *Journal of Marketing*, *49*(4), 41–50. https://doi.org/10.2307/1251430

Pavlou, P. A., & Dimoka, A. (2006). The nature and role of feedback text comments in online marketplaces: Implications for trust building, price premiums and seller differentiation. *Information Systems Research*, *17*(4), 392–414. https://doi.org/10.1287/isre.1060.0106

Piazzai, M., Liu, M., & Montauti, M. (2024). Cognitive economy and product categorization. *Organization Science*, *35*, 1866–1889. https://doi.org/10.1287/orsc.2021.15751

Pontikes, E. G. (2022). Category innovation in the software industry: 1990–2002. *Strategic Management Journal*, *43*(9), 1697–1727. https://doi.org/10.1002/smj.3383

Pontikes, E. G., & Barnett, W. P. (2015). The persistence of lenient market categories. *Organization Science*, *26*(5), 1415–1431. https://doi.org/10.1287/orsc.2015.0973

Pontikes, E. G., & Barnett, W. P. (2017). The non-consensus entrepreneur: Organizational responses to vital events. *Administrative Science Quarterly*, *62*(1), 140–178. https://doi.org/10.1177/0001839216661150

Pontikes, E. G., & Kim, R. (2017). Strategic categorization. In *From categories to categorization: Studies in sociology, organizations and strategy at the crossroads* (Vol. 51, pp. 71–111). Emerald Publishing Limited.

Posen, H. E., Keil, T., Kim, S., & Meissner, F. D. (2018). Renewing research on problemistic search—A review and research agenda. *The Academy of Management Annals*, *12*(1), 208–251. https://doi.org/10.5465/annals.2016.0018

Proserpio, D., & Zervas, G. (2017). Online reputation management: Estimating the impact of management responses on consumer reviews. *Marketing Science*, *36*(5), 645–665. https://doi.org/10.1287/mksc.2017.1043

Rajagopalan, N., & Spreitzer, G. M. (1997). Toward a theory of strategic change: A multi-lens perspective and integrative framework. *The Academy of Management Review*, *22*(1), 48. https://doi.org/10.2307/259224

Reeves, C. A., & Bednar, D. A. (1994). Defining quality: Alternatives and implications. *The Academy of Management Review*, *19*(3), 419–445.

Simon, H. A. (1990). Invariants of human behavior. *Annual Review of Psychology*, *41*(1), 1–19. https://doi.org/10.1146/annurev.ps.41.020190.000245

Socher, R., Perelygin, A., Wu, J., Chuang, J., Manning, C. D., Ng, A. Y., & Potts, C. (2013). *Recursive deep models for semantic compositionality over a sentiment treebank*. Paper presented at the Empirical Methods in Natural Language Processing.

Sood, G., & Laohaprapanon, S. (2018). Predicting race and ethnicity from the sequence of characters in a name. *arXiv* 1805.02109.

Soublière, J.-F., Lo, J. Y., & Rhee, E. Y. (2024). Coherence within and across categories: The dynamic viability of product categories on Kickstarter. *Academy of Management Journal*, *67*(1), 61–91. https://doi.org/10.5465/amj.2022.0105

Taeuscher, K., Bouncken, R. B., & Pesch, R. (2021). Gaining legitimacy by being different: Optimal distinctiveness in crowdfunding platforms. *Academy of Management Journal*, *46*(1), 149–179. https://doi.org/10.5465/amj.2018.0620

Taeuscher, K., Zhao, E. Y., & Lounsbury, M. (2022). Categories and narratives as sources of distinctiveness: Cultural entrepreneurship within and across categories. *Strategic Management Journal*, *43*(10), 2101–2134. https://doi.org/10.1002/smj.3391

US Census Bureau. (2017). *Census regions and divisions of the United States*. https://www2.census.gov/geo/pdfs/maps-data/maps/reference/us_regdiv.pdf

Vergne, J.-P., & Wry, T. (2014). Categorizing categorization research: Review, integration, and future directions. *Journal of Management Studies*, *51*(1), 56–94. https://doi.org/10.1111/joms.12044

Verhaal, J. C., & Pontikes, E. G. (2022). Shaping cultural meanings in markets with category strategy and optimal distinctiveness: An agency-based perspective. In J. F. Soublière & C. Lockwood (Eds.), *Research in the sociology of organizations* (Vol. 80, pp. 179–191). Emerald Publishing Limited.

Westphal, J. D., & Graebner, M. E. (2010). A matter of appearances: How corporate leaders manage the impressions of financial analysts about the conduct of their boards. *Academy of Management Journal*, *53*(1), 15–44. https://doi.org/10.5465/amj.2010.48036721

Williams, R. (2022). *Analyzing rare events with logistic regression*. https://www3.nd.edu/~rwilliam/stats3/rareevents.pdf

Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data* (2nd ed.). MIT Press.

Wooldridge, J. M. (2015). Control function methods in applied econometrics. *The Journal of Human Resources*, *50*(2), 420–445. https://doi.org/10.3368/jhr.50.2.420

Wry, T., Lounsbury, M., & Glynn, M. A. (2011). Legitimating nascent collective identities: Coordinating cultural entrepreneurship. *Organization Science*, *22*(2), 449–463. https://doi.org/10.1287/orsc.1100.0613

Zeithaml, V. A., Parasuraman, A., & Berry, L. L. (1990). *Delivering quality service: Balancing customer perceptions and expectations*. Free Press.

Zervas, G., Proserpio, D., & Byers, J. W. (2021). A first look at online reputation on Airbnb, where every stay is above average. *Marketing Letters*, *32*(1), 1–16. https://doi.org/10.1007/s11002-020-09546-4

Zhang, L., Gupta, A. K., & Hallen, B. L. (2017). The conditional importance of prior ties: A group-level analysis of venture capital syndication. *Academy of Management Journal*, *60*(4), 1360–1386. https://doi.org/10.5465/amj.2013.1088

Zhao, E. Y. (2022). *Optimal distinctiveness: A new approach to the competitive positioning of organizations and markets*. Cambridge University Press.

Zhao, E. Y., Fisher, G., Lounsbury, M., & Miller, D. (2017). Optimal distinctiveness: Broadening the interface between institutional theory and strategic management. *Strategic Management Journal*, *38*(1), 93–113. https://doi.org/10.1002/smj.2589

Zhao, E. Y., Ishihara, M., & Lounsbury, M. (2013). Overcoming the illegitimacy discount: Cultural entrepreneurship in the US feature film industry. *Organization Studies*, *34*(12), 1747–1776. https://doi.org/10.1177/0170840613485844

Zott, C., & Huy, Q. N. (2007). How entrepreneurs use symbolic management to acquire resources. *Administrative Science Quarterly*, *52*(1), 70–105. https://doi.org/10.2189/asqu.52.1.70

Zuckerman, E. W. (1999). The categorical imperative: Securities analysts and the illegitimacy discount. *American Journal of Sociology*, *104*(5), 1398–1438. https://doi.org/10.1086/210178

Zunino, D., Suarez, F. F., & Grodal, S. (2019). Familiarity, creativity, and the adoption of category labels in technology industries. *Organization Science*, *30*(1), 169–190. https://doi.org/10.1287/orsc.2018.1238

SUPPORTING INFORMATION

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

How to cite this article: Taeuscher, K., Zhao, E. Y., & Lounsbury, M. (2025). Right on cue? Category-switching in online marketplaces. *Strategic Management Journal*, 1–34. https://doi.org/10.1002/smj.70018