The Natural Emergence of Category Effects on Rugged Landscapes

Anthony Vashevko

Abstract

Category theory finds that markets partition producers into categories and that producers who do not fit one specific category—or who span multiple categories—perform worse than their single-category peers. The dominant thread of category theory argues that categorizations stem from the bounded rationality of market audiences, who are forced to impose categorizations and ignore miscategorized producers to efficiently interact with the market. I present an alternative model in which producers in a market segregate into categories and experience an apparent miscategorization penalty not driven by a market audience: In a complex environment, producers imitate successful predecessors. *Ex-post* rationalization identifies clusters as categories. Categories reflect, but do not cause, producer success. This model of exploration of a complex environment accounts for the basic findings of category theory, and it predicts the dynamics of category emergence and change over time. I establish these results in a formal model and simulation.

1 Introduction

Category theory combines a theoretical proposition with an empirical finding. The theory describes how individual cognition processes create a pressure to conform (Zuckerman, 2017; Zhao et al., 2017): Markets are full of ambiguous objects for people to sort through. People simplify their search by categorizing objects. Objects that are hard to categorize are hard to understand, evaluate, and consume (Zuckerman, 1999; Hannan et al., 2007; Hannan, 2019). Evidence that individuals categorize is well-established in cognitive psychology (Rosch et al., 1976; Goldstone, 1994; Murphy, 2004) and marketing (Shocker et al., 1991; Roberts and Lattin, 1991). The theory aims to explain an empirical finding: markets routinely assign categorical labels to objects, and objects with multiple or ambiguous labels face penalties. Restaurants specializing in a single cuisine (Korean or Mexican) earn higher reviews than those serving mixed cuisines (Korean-Mexican, Kovács and Hannan, 2015). Movie audiences prefer single-genre movies (romance or horror) to multi-genre movies (romance-horror, Hsu, 2006). The path from theory to finding seems direct: Miscategorized objects fail in the market because individuals find it hard to understand them. A stream of research measuring the penalty to miscategorized—poorly- and multi-labelled—firms appears to support the story (Hannan, 2010; Durand and Paolella, 2013; Vergne and Wry, 2014).

Various findings complicate this simple aggregation story. The miscategorization penalty sometimes reverses (Smith, 2011; Leung, 2014; Sgourev and Althuizen, 2014; Paolella and Durand, 2016). Markets may harbor multiple audiences with different tastes (Kovács and Liu, 2016) and even audiences with differing preferences for categorical ambiguity (Pontikes, 2012; Goldberg et al., 2016). Further still, findings from cognitive psychology suggest that individuals are categorically flexible, capable of integrating multiple categorical systems (Medin et al., 1997), devising novel categories on the fly (Barsalou, 1983; Durand and Paolella, 2013), and even ignoring categories when they do not clearly apply to the problem at hand (Proffitt et al., 2000). The categories themselves can shift, and markets can appear and disappear (Christensen and Bower, 1996; Durand and Khaire, 2016). Such complexities mire a producer in uncertainty: If they conform, then to whose standard? If they conform now, will their niche persist tomorrow? And if they differentiate, what is the chance they will discover a new market need?

This paper aims to formalize the producer decision process which defines one half of category theory's market interface. It does so by combining two streams of work outside the theory: First, work in psychology and anthropology has recognized that real-world categories tend to align to natural divisions of the environment (Malt, 1995; Brown, 2004) and that independent cultures tend to group the same sets of objects together. Second, a literature on strategic positioning in markets has identified conditions under which producers agglomerate in the marketplace (Hotelling, 1929, Lancaster (1990)), and when such agglomerations result from the natural heterogeneity within markets and audiences (Levinthal, 1997; Adner et al., 2014; Lenox et al., 2007)—from the ruggedness of market landscapes.

This paper proposes that a model of producer entry into rugged landscapes provides a valuable null model for category theory, capable of replicating its core results and accounting for the dynamics of categorical evolution. Producers face vast uncertainty about optimal positioning in markets, including not only uncertainty about categorical boundaries, but also uncertainty about the basic location of consumer niches, or the viability of specific production processes (March et al., 1991; Hannan et al., 2003). Producers must balance this uncertainty against competitive pressures. They end up herding into dense clusters around successful positions in the market identified by prior entrants. In parallel, a passive audience labels these naturally formed clusters, providing a mechanism for communicating about the market without directly affecting outcomes within it.

This paper describes in detail how this landscape model provides an alternative to the cognitive model of market categorization. In doing so, it shows how the landscape model replicates the static findings of category theory, accounts for categorical dynamics and emergence, and identifies several boundary conditions on the existing theory of market categories. In each case, the paper identifies ways to compare the contributions of the two models to the operation of any particular market and to measure the extent to which rugged landscape effects contribute to a specific empirical setting. While the two models are difficult to distinguish, this paper suggests that awareness of both social and material constraints on markets can build a deeper understanding of the mechanisms at work. The paper discusses the theoretical issues at play, and then describes the model and its results. It concludes with strategies for comparing the landscape and cognitive models and for advancing the categorical theory of markets.

2 Theoretical Background

2.1 Constructs in Category Theory

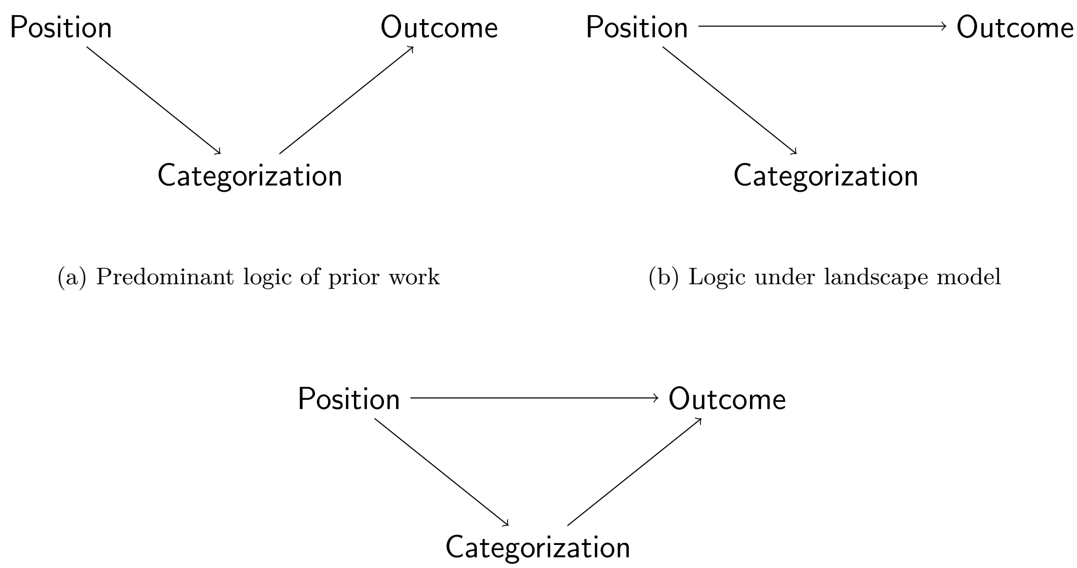
To establish the relationship of this article to category theory, it is helpful to examine the role of three constructs in the theory. First, objects in a market have a *position*, denoting the characteristics of an object which determine how it might appeal to various needs. Second, objects receive a *categorization*, the process by which audience members perceive the object's characteristics and communicate those to one another. Third, objects experience an *outcome*, some measurable degree of success in the market, whether through audience appeal, evaluations, prices, or sales.

The dominant audience-driven perspective in category argues that objects' categorizations mediate the link from their positions to their outcomes: Audiences categorize objects based on their characteristics, but the categorization process determines whether these objects are successful in the market (see Fig. 1a). The seminal works of Zuckerman (1999) and Hannan et al. (2007) set up this argument: Zuckerman (1999) describes how audience categorizations operate in a sorting process that selects against miscategorized objects, establishing the second link of fig. 1a. Hannan et al. (2007; further elaborated in Hannan, 2019) consider objects in an abstract feature space—their position—and describe how audiences assign categorical labels to clusters of similar objects in the space. This establishes how subsequent objects become categorized, the second link of fig. 1a.

Empirical work in category theory relies on the structure provided by these core pieces. The feature space model of Hannan et al. allows for increasingly sophisticated measures of producer categorization (e.g. Pontikes, 2012; Pontikes and Hannan, 2014; Kovács and Hannan, 2015; Hannan, 2019). Zuckerman (1999) establishes the plausible causal link from these categorizations to observable outcomes. Empirical work proceeds by showing an association between market-level measures of categorization and outcomes.

2.2 Categorization and Labels

The explanatory leap from positions to outcomes via categorizations obscures a theoretically critical aggregation. The cognitive foundation of category theory rests on an argument that individual categorization processes exhibit predictable imperfections. Indeed, work in category theory that examines individual categorizations supports this leg of the argument (Negro and Leung, 2013).



(a) Predominant logic of prior work

(b) Logic under landscape model

(c) Categorization as a partial mediator

Figure 1: Differing causal logics in category theory.

The theory then proposes that individual imperfections aggregate into collective flaws, that 'market categories' are the sum of many individual categories.

There are several concerns with this position. First, individuals appear to be categorically flexible. People form *ad hoc* categories to resolve unexpected or infrequent situations (Barsalou, 1983; Durand and Paolella, 2013). People use distinct category systems to classify the same objects, selecting the best categorizations for a given task (Medin et al., 1997), and incorporating non-categorical causal arguments to make predictions about specific objects (Proffitt et al., 2000). Second, collective categorizations do not simply aid individual categorization but serve collective coordination and communication functions. Collective concepts are constrained by the need for efficient, expressive, and socially conformant communication (Goldberg, 2019). Situation-specific languages emerge spontaneously (Weber and Camerer, 2003) and converge on mutually comprehensible terms (Guilbeault et al., 2021). In short, individuals appear to be able to flexibly adapt their categorical schemata to the needs of specific situations.

At the same time, the apparent stability of market categories appears to reflect the general tendency of cognitive concepts to anchor on material constraints. Despite the flexibility of individual categorizations, categories in the wild obey various regularities (Murphy, 2004). Categories bind objects with some degree of internal similarity and external distinction (Rosch and Mervis, 1975; Goldstone, 1994; Goldstone and Son, 2012). Hannan et al. (2007) likewise argue that categories are most meaningful when they refer to dense, distinct clusters of similar objects. In addition, people tend to rely on a basic level of categorization for any particular object—referring to cats as 'cats,' for example, and not by specific breeds or as 'animals' (Rosch et al., 1976; Murphy, 2004). Across cultures, societies tend to align on similar categorizations of specific objects (Malt, 1995; Brown, 2004; Xu et al., 2013), such as different kinds of plants or animals; in the case of biological categorizations, folk categories frequently map onto scientific taxa. In general, collectively held categories appear to frequently reflect natural kinds, natural clusters and distinctions—there is little reason to suspect that market categories follow some different principle.

A finding that collectively held categorizations predict product outcomes theoretically implicates individual categorization processes only if no other causal mechanism could produce collective categorizations. This article lays out a mechanism by which natural kinds emerge in the market.[1]

2.3 Agglomeration, Environmental Complexity, and Imitation

There is an extensive economic literature on positioning, product variety, and agglomeration (Hotelling, 1929; Biscaia and Mota, 2013; Fujita and Thisse, 1996; Lancaster, 1990). Such work tries to explain how producers position their products and the circumstances under which they will tend to cluster or separate from each other. For this paper, a key distinction among classes of mechanisms involves assumptions about variation in the product 'landscape'—whether there are positions of high or low appeal, customer density, or other local complications. A large class of papers identify mechanisms that generate clustering in the absence of exogenous variation, identifying various strategic implications to agglomeration or efficiencies that emerge as producers collocate. Such research aims to show that agglomerations can emerge even under strong assumptions about the environment—that is, even in homogeneous environments in which no position is favored *ex-ante*. Such models characterize an important set of mechanisms, but they are often sensitive to specific

[1] It must be pointed out that social classification is itself subject to interpretation as a strategic producer decision. Signaling is a well-established market dynamic (Spence, 1973), and producers will adopt labels that aid their performance in the market (Podolny, 2008; Etzion, 2014; Pontikes and Kim, 2017). This paper discusses the emergence of classification schemes in the absence of any market benefit to classification.

behavioral and environmental assumptions (see e.g. Salop, 1979; D'Aspremont et al., 1979).

This article instead assumes that exogenous environmental variation provides a more general description of the world, following a tradition of organizational work on environmental and organizational complexity, and on rugged landscapes. Externally, organizations are subject to environmental opacity and complexity, resulting in unanticipated outcomes across both time (Levinthal, 1991) and position (Levinthal, 1997). They only perceive limited information about their competitors (White, 1981). Internally, organizations have a limited understanding of their own routines and production processes (Nelson and Winter, 1982; Hannan et al., 2003; Bernstein, 2012). Attempts to achieve a particular market position may be hampered by unobservable components of a strategy (Rivkin, 2000) or by the impossibility of generating critical resources (Wernerfelt, 1984; Teece et al., 1997). In general, producers' ability to predict the success of a given position in the market may be very limited.

In such complex environments, producers' positioning is substantially driven by informational availability. Producers entering the marketplace must balance competitive pressure against the informational advantages of close imitation of competitors. In attempting to maximize their outcomes, producers will effectively face a choice between the imitation of successful predecessors and exploratory differentiation, producing a population that is at once exploratory and agglomerative (c.f. Banerjee, 1992; Strang and Macy, 2001; Denrell and March, 2001; Denrell and Le Mens, 2007). As even efforts to conduct market research or pursue lean strategies (Ries, 2011) can at best mitigate but not remove uncertainty, such agglomerative pressures will operate in complex markets. Differentiation carries the unavoidable threat of failure, which will deter those organizations seeking to escape competitive churn.

2.4 A Landscape Model of Category Formation

This article asserts that a rugged landscape model (Levinthal, 1997) of markets contains the minimal elements necessary to produce a marketplace featuring product agglomerations. These agglomerations act as natural kinds to anchor the formation of collective categorizations. Collective categorizations do not affect producers' outcomes in that market. To the extent that categorical effects appear within such markets, those effects are epiphenomenal to producers' efforts to manage the uncertainty of the landscape (fig. 1b).

I examine three questions about this model: first, whether it reproduces the basic patterns of category theory; second, whether it accounts for the problem of categorical dynamics and emergence; and third, whether it generates distinguishing predictions from the cognitive account.

2.4.1 Question 1: Reproduction of the category penalty

Category theory predicts that miscategorized objects will be penalized in markets because of they cannot enter audience members' consideration or they violate categorical rules. Empirically, this manifests in the finding that category spanners—objects that are classified into multiple categories—suffer relative to single-category objects. Conceptually, this argument also covers objects that fail to attract any category label. Under the landscape model, this penalty arises from producers' tendency to cluster around successful positions.

2.4.2 Question 2: Category dynamics and emergence

Category theory tends to consider markets in categorical equilibrium and relegates the problems of categorical emergence and change to ancillary mechanisms. Research has emphasized the role of strategic action by producers or audience members aimed at shifting categories. Some researchers discuss the role of social movements in creating new categories (Lee et al., 2017; Weber et al., 2008; Carroll and Swaminathan, 2000), or the strategic actions of organizations themselves (Kennedy, 2008; Pontikes and Kim, 2017; Grodal, 2018). Others discuss the role of influential vanguards (Koçak et al., 2014; Rao et al., 2005; Ruef, 2000). Scholars are generally pursuing mechanisms of category emergence (Glynn and Navis, 2013; Durand and Khaire, 2016; Lo et al., 2020). Another line of work discusses how categories shift over time, considering the role of successful exemplars (Zhao et al., 2018), of strategic manipulation of categorical boundaries by incumbents (Hsu and Grodal, 2015), or of categorical violations (Rao et al., 2005). These approaches emphasize the intentionality of category emergence and a stress on category definition as a prerequisite for markets.

By allowing producers to react to each other and explore the landscape over time, and by allowing collective categorizations to evolve in reaction to the market, the landscape model directly examines how categories shift, emerge, and disappear over time. Thus, the model can account for both stable and shifting markets using a single set of mechanisms. Moreover, it can establish conditions under which it may be possible to predict categorical change and emergence.

2.4.3 Question 3: Distinguishing landscape effects from audience cognition

Finally, this article aims not only to provide an alternative mechanism for categorization (fig. 1a vs fig. 1b), but also to provide some guide for how the two models might be empirically distinguished or integrated (fig. 1c). In general, distinguishing social constructionist from structuralist accounts of categorization is difficult (c.f. Malt, 1995). Nevertheless, this article explores several avenues, each relying on the ability of a formal model to provide deep insight into the mechanisms that drive specific outcomes. First, the model can characterize the core findings of a spanning penalty (Question 1) as a stochastic outcome with some probability of reversal. Second, the model is able to explore in depth how informational shocks cause predictable shifts in categorical structure. And third, by exploring the space of possible competition structures, the model is able to identify boundary conditions on markets under which categorical effects should be expected to disappear or reverse.

3 Model

This paper models the behavior of entrepreneurial producers sequentially entering a rugged landscape and observing the outcomes of their predecessors. In parallel, a passive audience observes and applies categorical labels to the producers in the market. The paper models the rugged fitness landscape of the market as a Brownian path, a novel machinery for examining complex environments (Callander, 2011; Ganz, 2018; Callander and Matouschek, 2019). I discuss the choice of landscape model, the producer entry decision, and finally the audience categorization process.[2]

3.1 Rugged Landscapes and Brownian Paths

Research on rugged landscapes has typically relied on NK landscapes to model performance environments (Levinthal, 1997; Kauffman and Weinberger, 1989) and organizational positioning problems (Lenox et al., 2007; Adner et al., 2014). Three features of NK landscapes render them unsuitable for the present context: First, they complicate the mapping of the constructs of category theory onto the landscape. In NK models, actors position themselves on discrete vertices of a

[2] Code implementing this model is available in an online repository (https://anonymous.4open.science/r/os-categories-42D7).

hypercube. Short of defining each position as its own category, it is unclear how to assign categorical positions to producers or how to define categorical spanning or partial membership. Second, it is not clear how to identify categorical movement across discrete positions. Third, it is difficult to define optimal search behavior in NK spaces; in practice, researchers study agents that use heuristic rules to explore the landscape. Since this article argues that categorical effects emerge in the absence of cognitive constraints, it seeks a model that does not embed cognitive biases into the search process. While none of these problems are insurmountable in NK, Brownian paths provide a cleaner analytical machinery.[3]

This paper models a market as a fitness landscape defined by a Brownian path. Following the Hotelling tradition, producers locate themselves continuously along the real number line, with each point representing some market position (Hotelling, 1929; Salop, 1979). Positions that are close to one another represent similar products, products that act as close substitutes to each other in the market. The outcome assigned to a given position—that is to say, the inherent appeal or quality of a given position—is given by the height of a Brownian path at that point. Denoting a market position as x, its appeal is A_x. The basic property of the Brownian path is that the difference in appeal between two positions is given by a normal distribution with variance proportional to the distance between the positions and a possible directional drift term μ:

$$A_x - A_y \sim N((x-y)\cdot\mu, |x-y|\cdot\sigma^2)$$

I set $\mu = 0$ to model a homogeneous landscape where any position is *ex-ante* equivalent. Other work on Brownian landscapes considers $\mu \neq 0$ to explore the efficiency of search (e.g. Callander, 2011); research on organizational mortality has used Brownian motion models to represent an organization's internal buffer against failure (Levinthal, 1991; Denrell, 2004; Le Mens et al., 2011). In addition, because σ^2 acts only as an arbitrary scale on landscape distances, I set $\sigma^2 = 1$ for convenience. This reduces the landscape's governing equation to $A_x - A_y \sim N(0, |x-y|)$.

The appeal function A encapsulates and abstracts the producer's costs and the heterogeneous

[3]Other alternatives exist. Multi-arm bandit models suffer from the first two issues, through a lack of continuous positioning. PN landscapes, as recently proposed by Rahmandad (2019), suffer from the third, the difficulty of deriving optimal search behavior given bounded knowledge of the environment. In addition, Brownian landscapes can be shown to be a mathematical limit of infinite-N, low-K NK landscapes. As such, the use of a Brownian landscape here represents an analytically convenient modification of an NK model.

tastes of consumers at a given position, representing the inherent appeal of a given position to the producer. Nearby producers experience similar production issues and similar desirability to the audience; distant producers experience unpredictable differences on both dimensions due to the complexity of the market. Finally, because this landscape has no natural cardinality, I anchor the landscape by fixing the appeal at the origin, $A_0 = 0$.

3.2 Producer Competition and Entry

Producers enter the landscape one at a time over multiple periods, attempting to maximize their outcomes given both uncertainty and competition from prior entrants. An initial entrant enters at the origin as an anchor for the landscape. As producers enter the market in each subsequent period, we can define the set X_t of producers' position in the market at period t. These positions are publicly known, as are the appeals of each position.

Competitive pressure takes the form of a simple penalty to the appeal of a particular position depending on its distance to its nearest neighbors on the left and right. A producer considering market entry in period t at position x would consider their immediate neighbors to the left x_l ($x_l = \max_{x>y\in X_t} y$) and right x_r (analogous). The competitive penalty is then $c(x, X_t) = \frac{1}{x-x_l} + \frac{1}{x_r-x}$. If there is no competitor on a given side, then that component of the penalty is set to 0; if x_r does not exist, for example, the penalty is $c(x, X_t) = \frac{1}{x-x_l}$. Fig. 2 illustrates the penalty. If the producer enters at that position, they would then observe its appeal A_x, and the producer's market outcome will be $A_x - c(x, X_t)$. This penalty offers a greatly simplified version of price- or patent-based competition and is chosen to simplify analysis of the model. In particular, because the penalty approaches infinity at perfect imitation, it ensures that no two producers locate at the same market position. Appendix 2 examines the multinomial logit model of product choice and finds a similar pattern of results (McFadden, 1974, 1984; de Palma et al., 1985; De Palma et al., 1987).

Producers enter at the position that maximizes their expected utility $u(A_x - c(x, X_t))$. The choice of utility function is primarily determined by the behavior of expected outcomes on the extremes of the landscape. In period t, if x_R is the rightmost producer in the market ($x_R \geq x$ for all $x \in X_t$), then $E(A_x) = A_{x_R}$ for all $x > x_R$: any position beyond the most extreme producer has the same expected appeal. Because competitive pressure decreases with distance from the incumbent, risk-neutral producers prefer to locate infinitely far away from the most extreme producers. It is

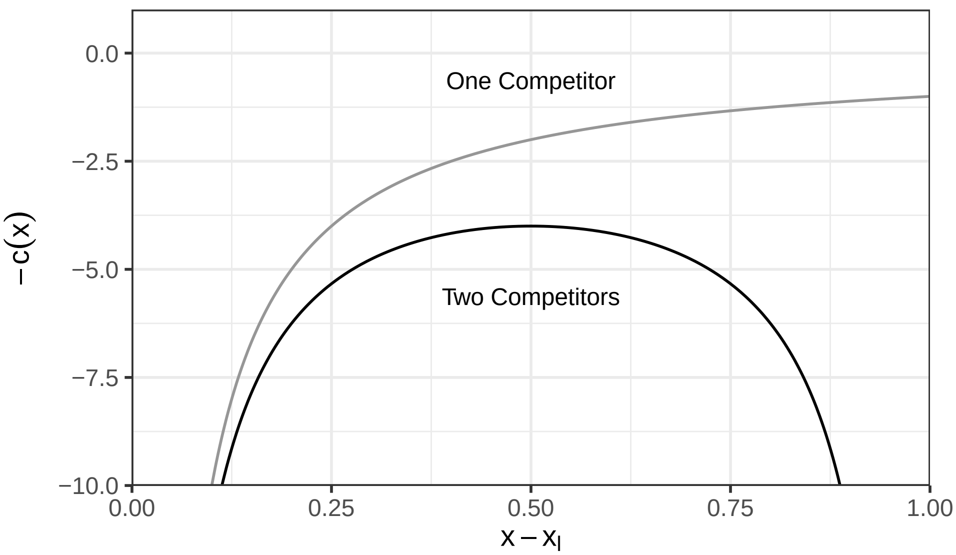


Figure 2: Competitive penalties, with single competitor at $x = 0$ or multiple at $x = 0, 1$

possible to resolve this concern by limiting the span of the environment (e.g., by adopting a circular landscape, as in Salop, 1979) or by limiting the search distance (e.g. through heuristic search rules avoiding extreme positions, or by imposing a search cost increasing in distance). Instead, I assume that producers are risk-averse, and avoid entering at extreme positions due to fear of the uncertainty involved. In particular, u must exhibit decreasing absolute risk aversion (Pratt, 1964; Callander and Matouschek, 2019), and I adopt the specific utility function $u(y) = ay - \exp(-by)$, with $b > 0$. Appendix 2 shows similar results with alternative landscapes and decision rules.

Producers search for entry positions only on the intervals between incumbents (bridge intervals) or on the intervals beyond incumbents (open intervals). Because competition depends only on immediate neighbors, and because the Brownian path is a Markov process, entrants only need to know the positions and appeals of the incumbents on the endpoints of an interval to characterize the optimal entry position on the interval. They select an optimal entry distance Δ from the leftmost competitor (x_l) on bridge intervals (from the nearest competitor on open intervals). For convenience, the distance from their rightmost competitor, at x_r is given by $\bar{\Delta} = x_r - x_l - \Delta$. The expected utility can be decomposed into mean (M) and variance (V) components, taking the form:

$$E(u(\Delta)) = aM(\Delta) - \exp\left(-bM(\Delta) + \frac{1}{2}b^2V(\Delta)\right)$$

$M(\Delta)$ is the expected appeal at Δ, and $V(\Delta)$ is the contribution of variance to the expected utility. On open intervals, these equal:

$$M(\Delta) = A_{X_C} + c(\Delta)$$
$$V(\Delta) = \Delta\sigma^2$$

On bridge intervals, these equal:

$$M(\Delta) = A_{x_l} + \frac{A_{x_r} - A_{x_l}}{x_r - x_l} \cdot \Delta + c(\Delta) + c(\bar{\Delta})$$
$$V(\Delta) = \frac{\Delta \cdot \bar{\Delta}}{x_r - x_l}$$

Entrants enter at the interval and position that maximizes their expected utility.

Fig. 3 illustrates this entry process for three entrants. First, the initial producer (P0) enters at position $x = 0$ with appeal $A = 0$, anchoring the landscape. The full appeal landscape, shown as a gray line, is unobservable to producers. Second, entrant one (P1) evaluates expected appeal over the open intervals left (A) and right (B) of P0 (1 SD of uncertainty in the gray regions), and their expected utility of entering at every point ("Expected Utility"). P1 enters at their position of maximum expected utility, marked by the vertical dashed line, and observes their true appeal, marked by the open circle, which falls below their initial expectation. Third, entrant two (P2) repeats this process, evaluating their expected appeal on the open interval left of P0 (C), the open interval right of P1 (E), and on the bridge interval between them (D). P2 calculates their expected utility over each interval and enters at the maximum, to the left of P0. Note that P1 and P2 have identical expectations over regions A and C, as P1's entry provided no additional information about it.

3.3 The Categorization Process

While the audience plays no role in defining the appeal of different positions on the landscape, it does exist to categorize clusters of similar producers in the landscape. I model this categorization process

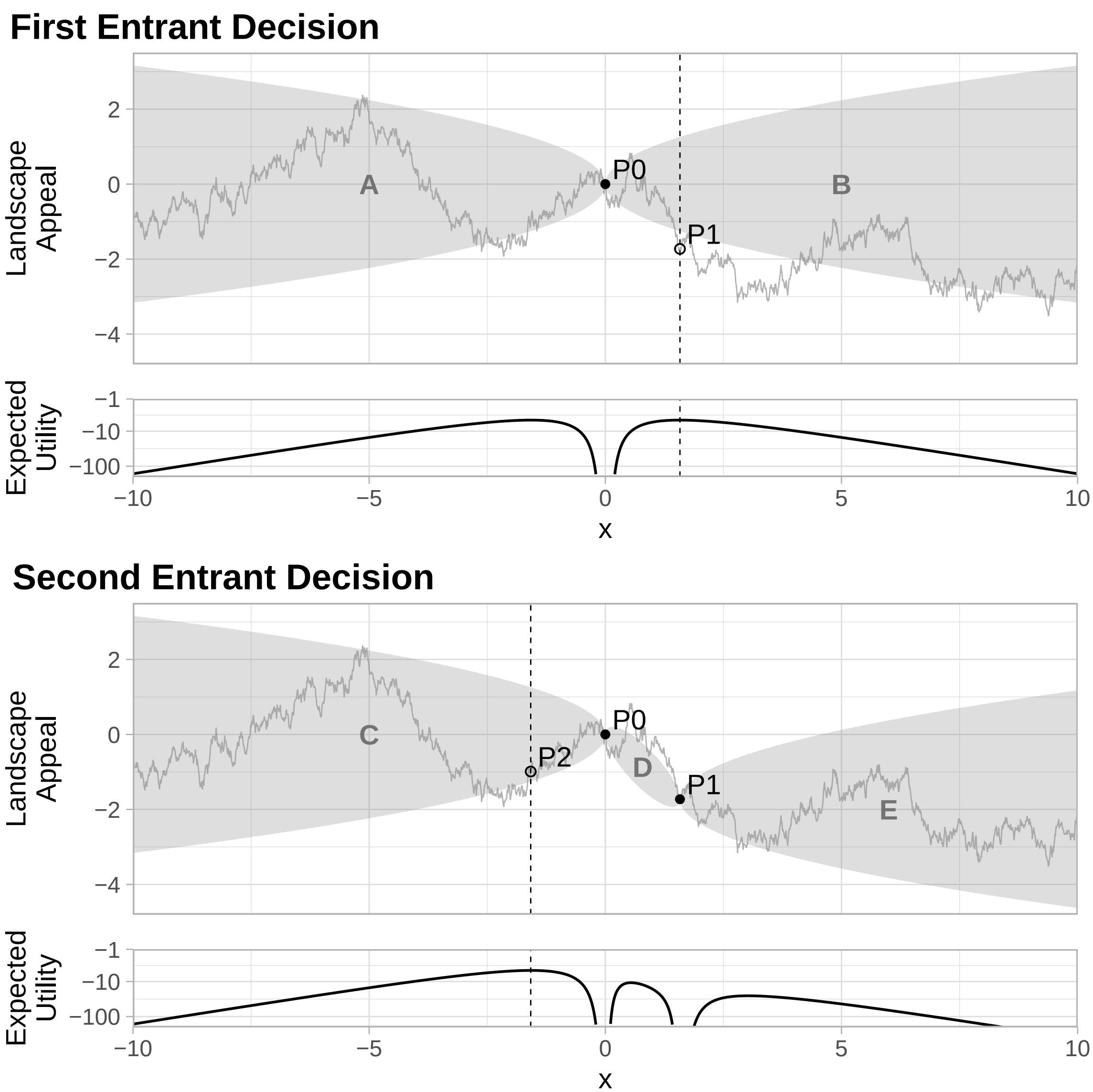


Figure 3: Illustration of the entry process for three entrants.

by fitting a finite Gaussian mixture model to the positions (X_t) of producers in the market in each period (Dempster et al., 1977; Xu and Wunsch II, 2005). This model assumes that the positions of a set of points, $x_1, \ldots x_t$, are given by some set of k normal distributions, $\{N(\mu_1, \sigma_1^2), \ldots, N(\mu_k, \sigma_k^2)\}$, centered at different means and with potentially different variances. Each point x has a probability of belonging to each distribution i, $p_i(x)$. Gaussian mixtures place the centers of clusters at particularly dense parts of the set. This model not only partitions the set of producers into categories but also models each producer's grade of membership in each category—positions close to the center of a cluster receive higher grades of membership. The number of clusters k was selected using the Bayesian Information Criterion (Schwarz, 1978; Steele and Raftery, 2009).

Fig. 4 illustrates one simulated market, the clusters identified in it by the Gaussian mixture model, and the associated measures of category membership described further below. Panel **A** shows the positions of producers and the appeals A of their positions; the bottom of the panel shows the local density of producers across the market. Panel **B** shows the probability density functions of the two clusters identified in the market: the two clusters align with the peaks seen in Panel **A**.

Formally, I define the grade of membership GOM of a point x in cluster i as the logarithm of a point's predicted likelihood of being in the cluster, $p_i(x)$, normalized by the peak predicted likelihood of the entire cluster:

$$\mathrm{GOM}_i(x) = -\left(\max_{z \in \mathbb{R}} [\log p_i(z)] - \log p_i(x)\right)$$

Comparison against the peak-likelihood point controls for differences in cluster width: points in more diffuse clusters have lower likelihoods on average. Panel **C** of fig. 4 shows the grade-of-membership functions associated with each cluster identified in the example: by construction, the most central producers in each clusters have identical grades of membership.

I also construct measures of miscategorization and spanning for each position. For every position, I consider the two clusters in which the point has its highest grades of membership. Taking $p_1(x)$ and $p_2(x)$ to be the highest and second highest predicted likelihoods for the position, I first define a measure of miscategorization and a measure of spanning, as depicted in Panel **D** of fig. 4:

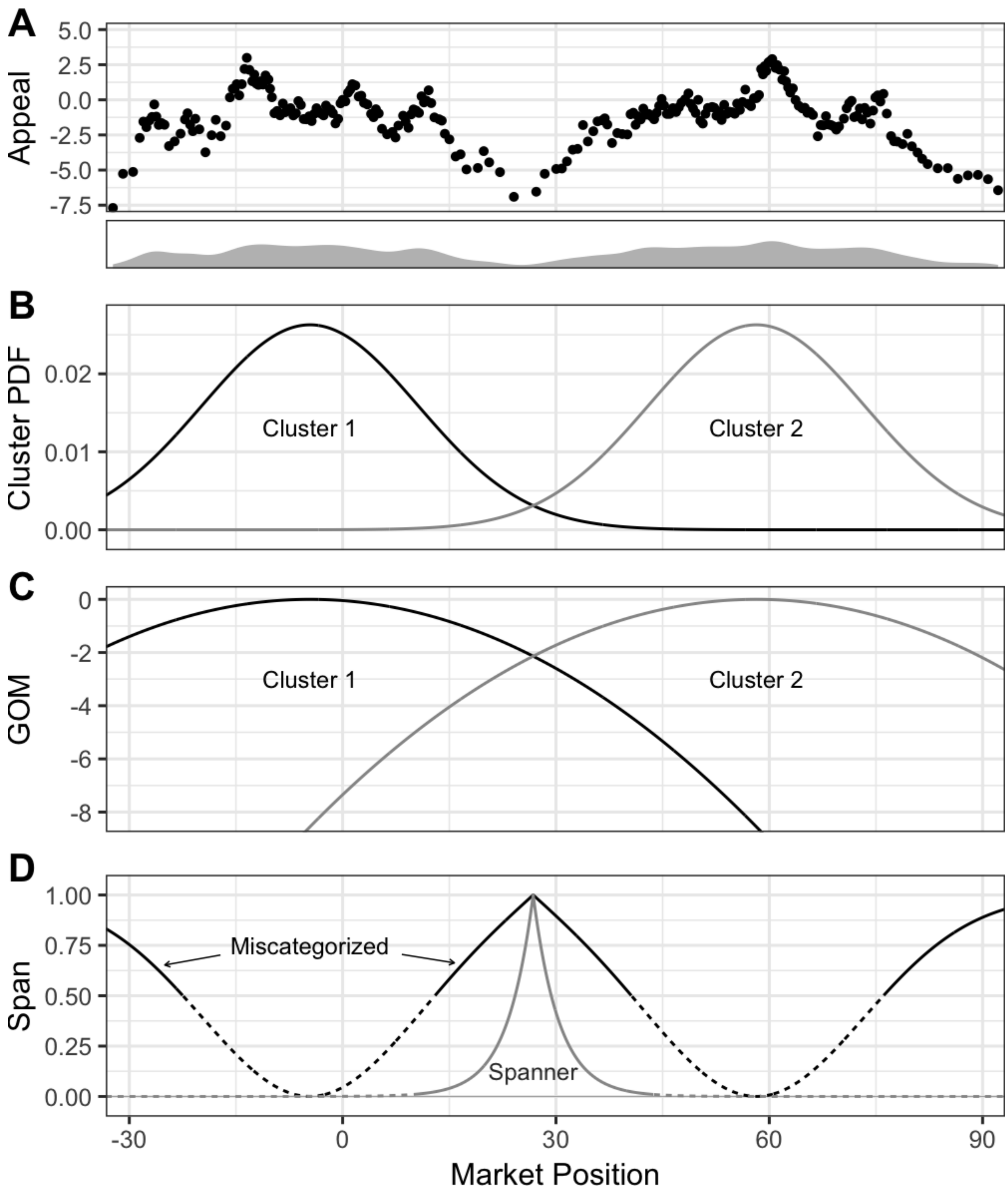


Figure 4: Example of a simulated market, estimated clusters, and measures of category membership

$$\text{Miscat}(x) = 1 - (p_1(x) - p_2(x))$$
$$\text{Span}(x) = \frac{p_2(x)}{p_1(x)}$$

Positions are *miscategorized* either if they are unlikely to belong in any cluster or if they are about equally likely to belong to multiple clusters. Positions are spanners if they are similarly likely to belong in multiple clusters. I reduce these to binary measures, so that a position x is considered miscategorized if $Miscat(x) > 0.5$, and it is considered a spanner if $Span(x) > 0.01$. The particular thresholds were chosen to ensure a balance of producers across classes, but analysis is robust to variation around these specific values. Finally, I construct a binary measure of *non-categorization* if a position is miscategorized and is not a spanner. This allows me to examine the effects of spanning and non-categorization separately.[4]

4 Results

4.1 Reproduction of Miscategorization Effects

I reproduce the core claims of category theory at three levels. First, at the most fundamental level, the miscategorization penalty argues that producers with a higher grade of membership in a category appeal more to members of the audience. Positions that do not fit well into any category should underperform those that do. Second, as a direct replication of category theory, I test whether category spanners are less appealing than single-category specialists. Finally, I examine the direct mechanism by which this effects appears in the landscape model: success generates herding, which generates categorization; conversely failure isolates producers. I measure whether more isolated producers—those with a higher minimum distance to their nearest neighbor—tend to underperform less isolated producers.

I test these relationships using cross-sectional linear regressions estimating the effect of each

[4]The regressions I consider below include *Spanner* and *Non-Categorized* as predictors. Using *Spanner* and *Miscategorized* as predictors instead causes the coefficient on spanner to reflect the effect of spanning net of the underlying effect of miscategorization. All else equal, spanning points tend to occur in denser, higher value regions in which two clusters exist close together, which causes *Spanner* to predict a positive effect of spanning relative to the miscategorization alone. Treating spanning and non-categorization as mutually exclusive allows for closer comparison to prior research.

measure of category membership on the appeal of the position. I estimate the effects within each period of each market. Specifically, within each period t and each market i, I estimate three regressions:

$$A_{x,i,t} = \alpha + \beta_{i,t} NGOM_{x,i,t} + \epsilon_{x,i,t}$$

$$A_{x,i,t} = \alpha + \beta_{1,i,t} NoCategory_{x,i,t} + \beta_{2,i,t} Spanner_{x,i,t} + \epsilon_{x,i,t}$$

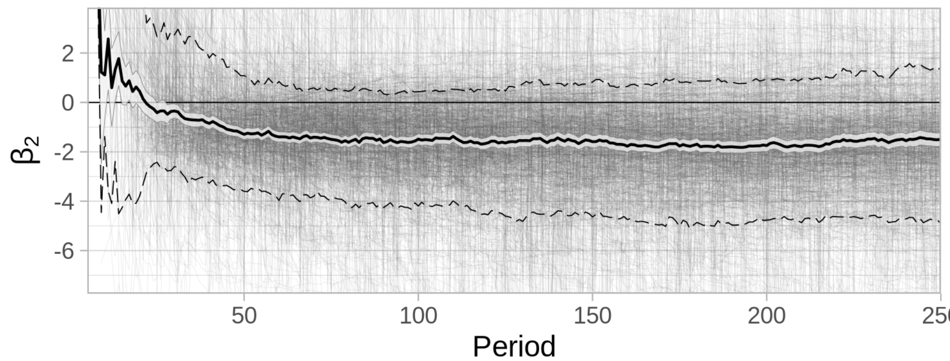
$$A_{x,i,t} = \alpha + \beta_{i,t} Distance_{x,i,t} + \epsilon_{x,i,t}$$

Here, x indexes all producers within a market during the period, and each regression considers one of the measures of category membership: (negative) grade of membership; binary miscategorization; or distance to the nearest neighbor. Thus, in each market and each period, I estimate three separate β coefficients representing how miscategorization predicts producer appeal.

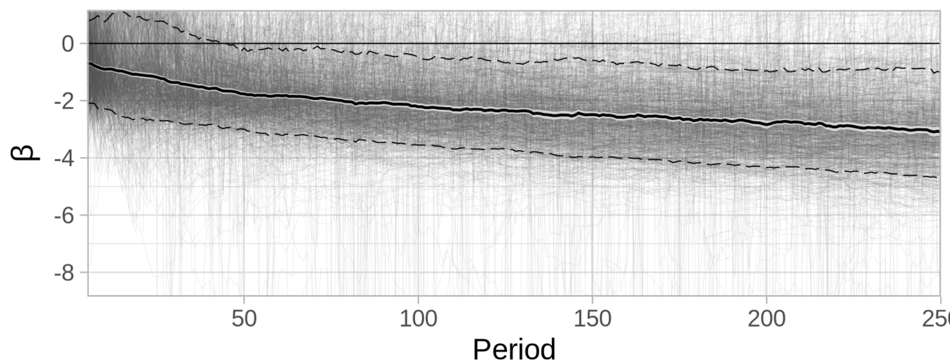
In each case, negative β indicates a miscategorization penalty. Fig. 5 plot these coefficients across all markets over time. Fig. 5a shows the difference between category-spanning and non-category-spanning positions; fig. 5b shows the effect of decreasing grade-of-membership; Fig. 5c shows the effect of an increasing distance to a neighbor on producer appeal. The plots show the average effect (bold), the 10th and 90th percentiles (dashed), and all individual markets (thin gray).

Each of the figures supports the predicted relationships. Category spanners generally have lower appeal—the estimated effect of spanning is generally negative within markets and it is negative when aggregating across all markets. Positions with a high grade of membership in some category have higher appeal. Most fundamentally, isolated positions have lower appeal. These relationships appear to strengthen over time as markets approach an equilibrium state.

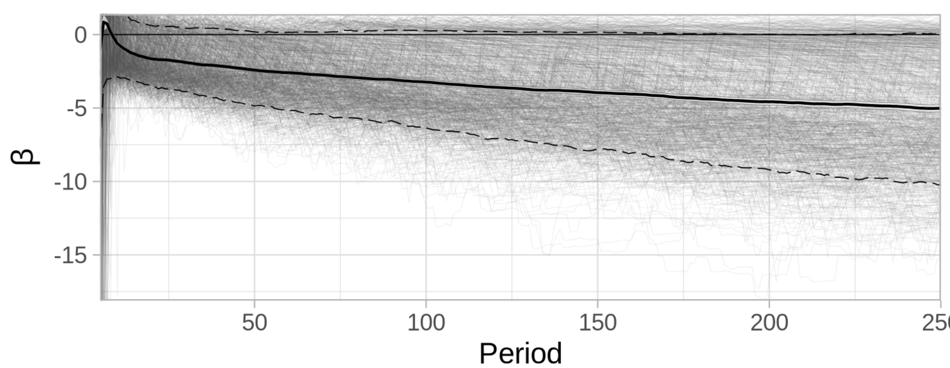
Miscategorization penalties, however, are not absolute. As the thin lines of fig. 5 suggest, many individual markets experience periods when the correlation between clustering and appeal disappears or inverts. In markets at least 50 periods away from initial conditions, about 6 percent of markets will experience an inversion of the penalty in grade of membership; 49 percent of markets will experience such an inversion at some point in their history. With the penalty in category spanner status, 14 percent of markets experience an inversion in any given period, and 70 percent experience it at some point. Inversions occur when the most successful producers in a market are located



(a) Effect of category spanning. $W = \alpha + \beta_1 \cdot NoCategory + \beta_2 \cdot Spanner$



(b) Effect of decreasing grade of membership. $W = \alpha + \beta \cdot NGOM$



(c) Effect of increasing distance to neighbor. $W = \alpha + \beta \cdot distance$

Figure 5: Effect of position characteristics on position appeal, 1000 simulated markets, 250 periods. Average effect in bold, 95% confidence interval in white band. Dashed lines indicate 10th and 90th percentiles of effect across all markets.

outside of major clusters or categories. They occur entirely because the appeal function is unknown: as producers discover highly appealing regions outside of existing clusters, they find themselves simultaneously outside of existing categories and in positions of high appeal. As producers flock to success, enough may wind up outside of existing clusters that the overall predictive power of categories disappears.

4.2 Change and Emergence of Categories over Time

The Gaussian model of categorization adapts each period as producers enter new positions in the landscape. These adaptations take two forms: first, new categories can emerge or disappear as the model fit favors a larger or smaller number of mixture distributions; second, the distributions themselves shift as producers uncover more of the landscape, affecting the grades of membership of existing producers. Both adaptations are predictable.

4.2.1 Category emergence and disappearance

At its core, the miscategorization penalty in this model arises because producers position themselves in regions of high appeal, so that producer density is predictive of local appeal—in effect, producer clusters form a lay theory of where success lies in the landscape. New entrants into the landscape either confirm or refute this theory. They confirm it by succeeding within existing clusters or failing outside those clusters. They refute the theory by finding success outside existing clusters. Refutations weaken the theory and the miscategorization penalty at the same time as they uncover regions of the landscape that may form the core of future categories. Refutations can also reveal that neglected valleys between peaks of the landscape are more appealing than initially presumed, causing producers to fill the gaps between existing clusters, merging two categories together.

If this argument holds, then weakening of the miscategorization penalty (i.e. negative trends in the penalty) will predict changes in the number of categories. I evaluate the penalty in each period of each market as described in the previous section, measuring the effect of decreasing grade-of-membership and the effect of spanning multiple categories. I measure the trend by the overall change in the estimated β coefficient for each effect across the prior 25 periods. A negative trend implies that the miscategorization penalty has weakened; a positive trend, that the penalty grew stronger. I then predict the probability that the number of categories will change in the current

	Model 1	Model 2	Model 3	Model 4	Model 5
(Intercept)	-3.013^{***}	-3.015^{***}	-3.204^{***}	-3.202^{***}	-3.204^{***}
	(0.010)	(0.010)	(0.014)	(0.014)	(0.014)
Penalty Trend (Distance)	-0.637^{***}				
	(0.088)				
Penalty Trend (GOM)		-1.207^{***}			
		(0.116)			
Penalty Trend (No Cat.)			-0.629^{***}		-0.655^{***}
			(0.071)		(0.074)
Penalty Trend (Spanner)				-0.052	0.130
				(0.090)	(0.092)
Num. obs.	221000	221000	138193	138193	138193

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Table 1: Trends in the clustering penalty predict category change

period as a function of recent trend.

Tbl. 1 presents the estimates of these logistic regressions. The table shows that a recent positive trend (strengthening penalty) strongly predicts stability in the number of categories in a given period. A negative trend (weakening penalty) thus predicts an increase or decrease in the number of categories.

Insofar as the miscategorization penalty and its trends are directly observable in markets, these results helps contextualize various empirical findings in the category theory literature. Markets with a weaker or weakening miscategorization penalty indicate the presence of untapped opportunities, with potential for further exploitation. In particular this result presents an alternate interpretation of the market-maker/market-taker distinction drawn by Pontikes (2012). Pontikes argues that different audiences may have different tastes for categorical ambiguity, showing that venture capitalists are more favorable to category spanners than are general audiences. The landscape model suggests that venture capitalists may select for systematically different markets than the general audience. If venture capitalists elect to operate in markets with a weaker penalty, their preference for categorical ambiguity reflects the decreased predictive power of categorical labels in such markets—in such markets the penalty itself is weaker. Similar dynamics may cause the preference for atypical hedge funds identified by Smith (2011).

4.2.2 Category movement

Producers maintain a shared expectation of the value of all positions in the landscape. They enter at the position that maximizes their expectation. What they discover when they enter constitutes not only a shock to their own expectation about the region they entered, but also to the expectations of all subsequent producers. Positive shocks attract subsequent entry, and negative shocks deter entry. Because categories move towards positions of increased producer density, entry shocks cause categories to shift and affect the categorical membership of incumbent producers.

The landscape model provides direct insight into producer expectations. I construct a z-scored measure of the difference between realized and expected appeal:

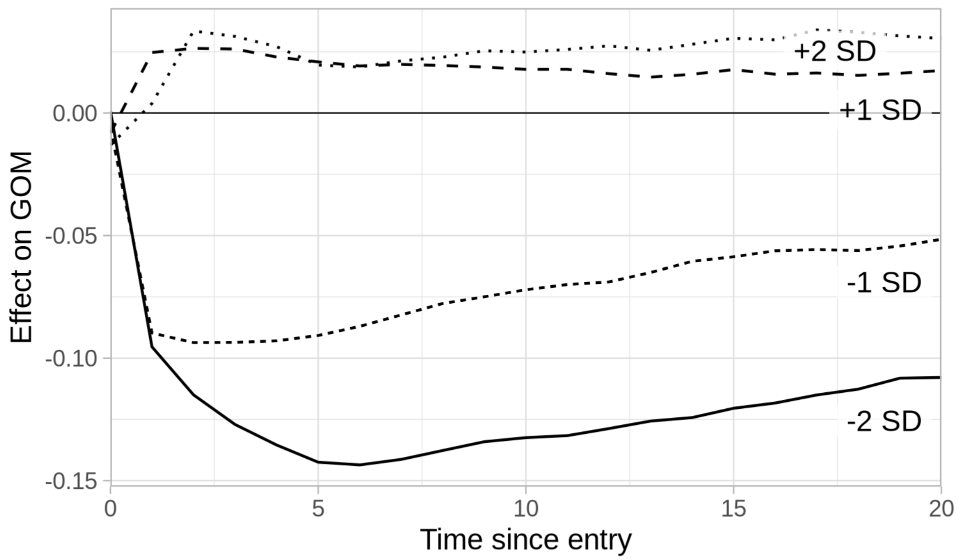
$$z_x = \frac{A_x - E(A_x)}{\sqrt{V(x)}}$$

Here, V is the variance of the Brownian walk at x conditional on the positions of its nearest neighbors.

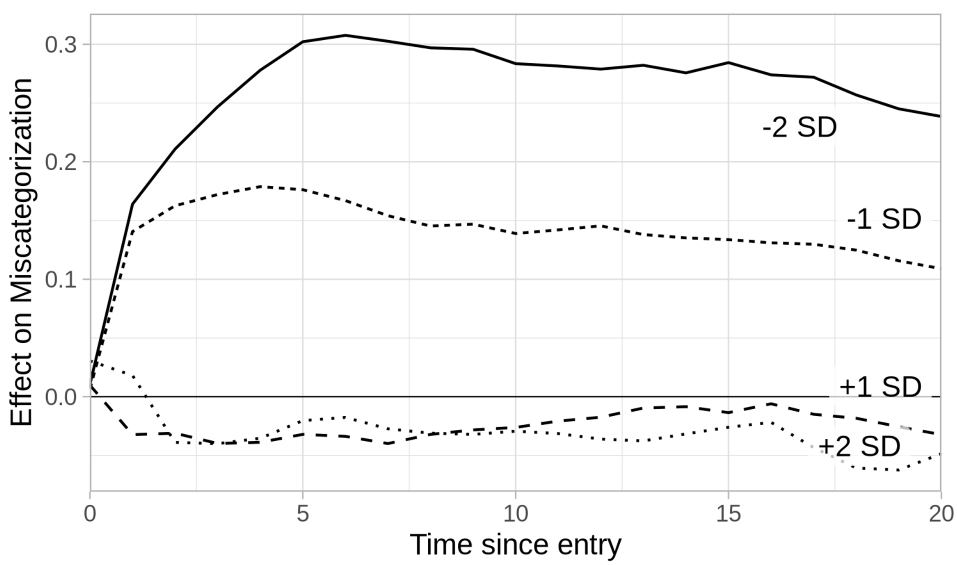
I focus on the entrant themselves and track how the shock z_x affects their own grade of membership and likelihood of miscategorization in subsequent periods.

Fig. 6 plots the effects of entry surprise on categorical centrality over time, highlighting the effect of positive ($z_x = 1, 2$) and negative surprises ($z_x = -1, -2$). The plot confirms that positive shocks increase the categorical centrality of the entrant over time, while negative shocks reduce it. There is a notable asymmetry between positive and negative shocks: Because entrants are already entering at the position of highest expected utility across the landscape, positive shocks have only a limited ability to draw in subsequent entrants. Negative shocks, on the other hand, do serve to dissuade further entry, especially if they push the expected appeal of the nearby region below the expected appeal of the next-best alternative.

As with category emergence, this result directly suggests alternative explanations for prior findings in the category literature. In particular, Zhao et al. (2018) show how hit video games establish proto-categories that develop into established categories as they attract imitators. The landscape model describes how unexpected hits provide the seeds for novel categories, as well as the pace at which market audiences come to recognize such developments.



(a) Effect of entry surprise on grade of membership



(b) Effect of entry surprise on likelihood of miscategorization

Figure 6: Effect of entry appeal surprise on categorical centrality.

4.3 Competitive Regimes in the Landscape Model

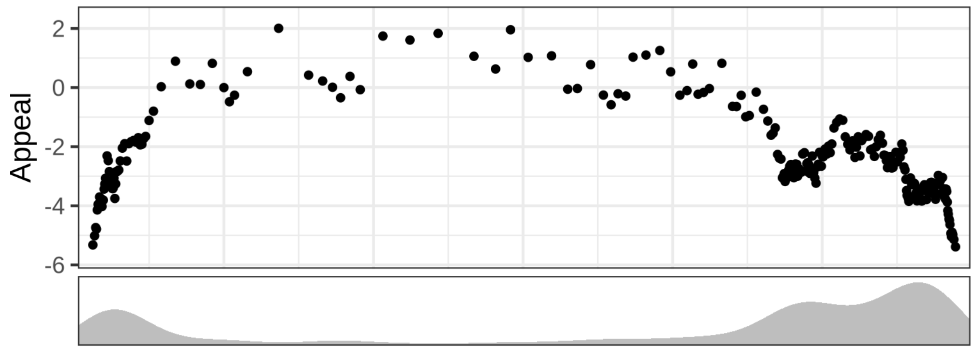
In the landscape model described here, reproduction of the miscategorization penalty relies on the interaction of two mechanisms: (1) the tendency of producers to cluster in regions of high appeal; and (2) the tendency of audiences to categorize dense clusters of producers. The first mechanism depends on the specification of competition in the market. As described above, the model assumes that all producers exert equivalent competitive pressure on each other. If instead high-appeal producers exert greater competitive pressure than low-appeal producers, they may deter and exclude potential competitors, restricting new entrants to a low-appeal periphery (c.f. resource-partitioning, Carroll, 1985). In this case, the audience would assign categories to low-appeal positions, reversing the miscategorization penalty.

I model this variant by making the competition function exponential in producer appeal rather than constant, $c(x, X_t) = \frac{A_{x_l}}{x - x_l} + \frac{A_{x_r}}{x_r - x}$. Fig. 7a plots one such simulated market, showing how entrants avoid high-appeal competitors. Fig. 7b shows that the miscategorization penalty (in grade of membership) reverses in such situations, becoming a miscategorization bonus instead.

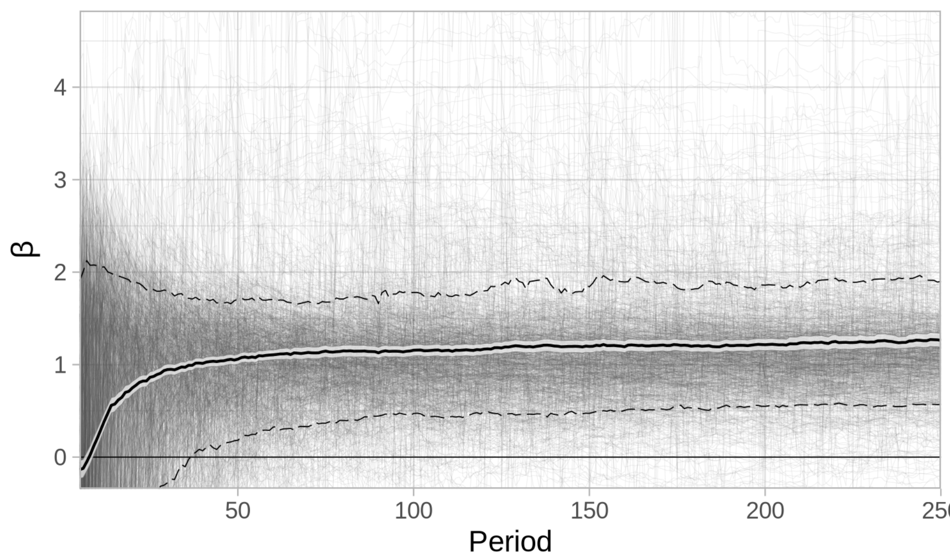
This result primarily suggests that the miscategorization penalty is particularly sensitive to specific interactions between categorization processes and competitive structure. In markets that resemble the situation described here, it should be possible to empirically observe an analogous reversal of the penalty. If instead audience categorizations follow categorical exemplars rather than similarity-based prototypes, the miscategorization penalty may persist (Murphy, 2004; Foster-Hanson and Rhodes, 2019).

5 Discussion & Conclusions

This paper proposes a landscape model as an alternative mechanism for the appearance of miscategorization effects in markets. Category theory has predominantly relied on a mechanism in which market audiences, possessed of categorical rules for market membership, punish or ignore deviant producers and create an empirically observable penalty against category-spanning and miscategorized products. This account struggles not only to explain how cognitively flexible individuals (Durand and Paolella, 2013; Guilbeault et al., 2021) come together into a rigid market collective but also to account for how collective categories shift over time. The landscape model accounts both for the



(a) Producer appeal and positioning in a simulated market.



(b) Effect of decreasing grade of membership on appeal. $W = \alpha + \beta \cdot NGOM$

Figure 7: Category effects in partitioned markets.

cross-sectional appearance of collective penalties and for the dynamics of categories in markets: As producers compete in an uncertain, rugged landscape, they uncover the appeal of different positions and agglomerate in regions of high appeal. As audiences label clusters of similar producers, they will find that these labels correlate with appeal, while low-appeal producers find themselves isolated and unclassified. As producers explore the environment, the categories themselves will emerge, disappear, and shift in predictable ways.

The paper suggests various means to identify whether the landscape model operates in a particular market. In cross-sectional analyses, the results above suggest that markets should experience a certain frequency of reversal of the miscategorization penalty, which declines with market maturity (c.f. Paolella and Durand, 2016). The results also suggest that miscategorized producers should, on average, be more isolated and distinct than well-categorized producers. These results can be tested in meta-analysis or through novel research projects. In dynamic analyses, the results above suggest that a weakening miscategorization penalty predicts categorical emergence. In addition, the results show that negative entry shocks cause categories to migrate away from entrants, suggesting that such shocks may cause the disappearance of subcategories within markets. Both effects can be measured; to my knowledge, neither has been.

Beyond these results, the landscape model makes assumptions about the categorization process that lead to several testable implications. The model assumes that categories preferentially attach to clusters of dense, similar producers—an assumption shared by category theory (Hannan, 2019). This leads directly to the prediction that miscategorization penalties should systematically reverse in resource partitioned markets. Evidence to the contrary would suggest either that the landscape model is incorrect, or that the audience categorization process operates according to a different rule—either result would advance the understanding of market categorizations. More broadly though, categorization processes that violate this rule can be used to test the applicability of category theory—category systems that label features orthogonal to producer success (e.g. number of players required for a board game) should be immune to category effects under the landscape model.

Finally, the most fundamental comparison of the landscape and cognitive models would test whether actors can pierce categorical labels: First, category theory would predict that actors engaged in some optimization task should avoid efficient but categorically disfavored strategies, avoid deviating from inefficient but categorically favored strategies, and avoid imitating successful categorical deviations of others. Second, insofar as the landscape model suggests that categories serve to facilitate terse communication, compression of the communication channel from producers to audiences should magnify the strength of cognitive categorical effects relative to landscape effects insofar as audiences become more reliant on categorical imputations (Leung and Sharkey, 2013, suggests that such effects exist).

Beyond these direct comparisons, this paper provides two theoretical advances for category theory: First, it emphasizes a perspective in which categories guide but do not constrain behavior. Categories emerge to facilitate communication about markets, but they do not prevent producers or consumers from pulling back the categorical veil and comparing producers directly. Miscategorization effects appear because categorical labels preferentially attach to the most active parts of the marketplace—to those product configurations that producers consider most viable. Second, this paper emphasizes the role of producer agency and knowledge in shaping collective understandings of the market: Audience-side categorizations reflect discoveries made by producers' attempts to expand the market. While the categorical structure empirically manifests through an audience's categorical descriptions of it, the structure itself is built, maintained, and transformed through producer discoveries.

The model presented here is by no means the final word on landscape models of categorization. This model presents a narrow view of categorization processes that does not accommodate for category nesting (Cudennec and Durand, 2022), or fully model the process of categorical continuity and disappearance. More importantly, the model examines only the producer entry process, and does not consider the possibility that producers can exit the market or embrace learning strategies within it in order to enable cheaper search (c.f. Ries, 2011). Neither does the model measure market participants' differential ability to remember the outcomes of past forays into the market. Each modification offers an avenue for further elaboration of category dynamics within markets. The model also makes specific assumptions about producer risk aversion and their search behavior in novel regions of the marketplace. While the exact form of these assumptions is not critical to the results of the paper, further research on producer search behavior can contribute both to category theory and to the literature on organizational exploration.

Finally, the model takes its place alongside the production-side theory of miscategorization effects described by Hsu (2006) and Hsu et al. (2009). There the principle of allocation argument proposes that miscategorization penalties arise from producers' inability to operate at multiple positions, and this diversification imposes an efficiency penalty relative to specialists. Insofar as the landscape model considers producers only operating at a single point, an extension of the model to accommodate the differences between specialist and generalist production would allow for a closer examination of how miscategorization effects differentially affect specialization strategies.

In summary, the findings of category theory are consistent with a world in which categories inform but do not bind market behavior. Categories inform because they are intended to inform—they serve to usefully encapsulate rich understandings of objects and markets and to convey those understandings among the actors in those markets. They guide because actors know to trust a crowd with much broader knowledge than their own. But they do not bind because market actors know the limits of terse language and collective wisdom. The participants working on a rugged landscape do not know everything, and their shared knowledge has gaps. As they crowd together in the search for safe profits and promising fads, they leave room for entrepreneurs to try their luck in those empty spaces. And in a rugged landscape, some will get lucky.

6 References

References

Adner, R., Csaszar, F. A., and Zemsky, P. B. (2014). Positioning on a Multiattribute Landscape. *Management Science*, 60(11):2794–2815.

Banerjee, A. V. (1992). A Simple Model of Herd Behavior. *Quarterly Journal of Economics*, 107(3):797–817.

Barsalou, L. W. (1983). Ad hoc categories. *Memory & Cognition*, 11(3):211–227.

Bernstein, E. S. (2012). The Transparency Paradox: A Role for Privacy in Organizational Learning and Operational Control. *Administrative Science Quarterly*, 57(2):181–216.

Biscaia, R. and Mota, I. (2013). Models of spatial competition: A critical review. *Papers in Regional Science*, 92(4):851–871.

Brown, D. E. (2004). Human universals, human nature & human culture. *Daedalus*, 133(4):47–54.

Callander, S. (2011). Searching and learning by trial and error. *American Economic Review*, 101(October):2277–2308.

Callander, S. and Matouschek, N. (2019). The Risk of Failure: Trial and Error Learning and Long-Run Performance. *American Economic Journal: Microeconomics*, 11(1):44–78.

Carroll, G. R. (1985). Concentration and specialization: Dynamics of niche width in populations of organizations. *American Journal of Sociology*, 90(6):1262–1283.

Carroll, G. R. and Swaminathan, A. (2000). Why the Microbrewery Movement? Organizational Dynamics of Resource Partitioning in the US Brewing Industry. *American Journal of Sociology*, 106(3):715–762.

Christensen, C. M. and Bower, J. L. (1996). Customer Power, Strategic Investment, and the Failure of Leading Firms. *Strategic Management Journal*, 17(3):197–218.

Cudennec, A. and Durand, R. (2022). Valuing Spanners: Why Category Nesting and Expertise Matter. *Academy of Management Journal*, page amj.2020.0042.

D'Aspremont, C., Gabszewicz, J. J., and Thisse, J.-F. (1979). On Hotelling' s "Stability in Competition". *Econometrica*, 47(5):1145–1150.

de Palma, A., Ginsburgh, V., Papageorgiou, Y. Y., and Thisse, J.-F. (1985). The Principle of Minimum Differentiation Holds under Sufficient Heterogeneity. *Econometrica*, 53(4):767.

De Palma, A., Ginsburgh, V., and Thisse, J.-F. (1987). On Existence of Location Equilibria in the 3-Firm Hotelling Problem. *The Journal of Industrial Economics*, 36(2):245.

Dempster, A., Laird, N., and Rubin, D. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society. Series B (Methodology)*, 39(1):1–38.

Denrell, J. (2004). Random Walks and Sustained Competitive Advantage. *Management Science*, 50(7):922–934.

Denrell, J. and Le Mens, G. (2007). Interdependent Sampling and Social Influence. *Psychological Review*, 114(2):398–422.

Denrell, J. and March, J. G. (2001). Adaptation as Information Restriction: The Hot Stove Effect. *Organization Science*, 12(5):523–538.

Durand, R. and Khaire, M. (2016). Where Do Market Categories Come From and How? Distinguishing Category Creation From Category Emergence. *Journal of Management*, 43(1):87–110.

Durand, R. and Paolella, L. (2013). Category stretching: Reorienting research on categories in strategy, entrepreneurship, and organization theory. *Journal of Management Studies*, 50(6):1100–1123.

Etzion, D. (2014). Diffusion as Classification. *Organization Science*, 25(2):420–437.

Foster-Hanson, E. and Rhodes, M. (2019). Is the most representative skunk the average or the stinkiest? Developmental changes in representations of biological categories. *Cognitive Psychology*, 110:1–15.

Fujita, M. and Thisse, J.-F. (1996). Economics of Agglomeration. *Journal of the Japanese and International Economies*, 10(4):339–378.

Ganz, S. C. (2018). Hyperopic Experimentation: Organizations Learning About Managers Learning About Strategies. Preprint, SocArXiv.

Glynn, M. A. and Navis, C. (2013). Categories, identities, and cultural classification: Moving beyond a model of categorical constraint. *Journal of Management Studies*, 50(6):1124–1137.

Goldberg, A., Hannan, M. T., and Kovács, B. (2016). What Does It Mean to Span Cultural Boundaries? Variety and Atypicality in Cultural Consumption. *American Sociological Review*, 81(2):215–241.

Goldberg, A. E. (2019). *Explain Me This: Creativity, Competition, and the Partial Productivity of Constructions*. Princeton University Press, Princeton, New Jersey.

Goldstone, R. L. (1994). The role of similarity in categorization: Providing a groundwork. *Cognition*, 52(2):125–157.

Goldstone, R. L. and Son, J. Y. (2012). *Similarity*. Oxford University Press.

Grodal, S. (2018). Field Expansion and Contraction: How Communities Shape Social and Symbolic Boundaries. *Administrative Science Quarterly*, 63(4):783–818.

Guilbeault, D., Baronchelli, A., and Centola, D. (2021). Experimental evidence for scale-induced category convergence across populations. *Nature Communications*, 12(1):327.

Hannan, M. T. (2010). Partiality of Memberships in Categories and Audiences. *Annual Review of Sociology*, 36(1):159–181.

Hannan, M. T. (2019). *Concepts and Categories: Foundations for Sociological and Cultural Analysis*. Columbia University Press, New York.

Hannan, M. T., Pólos, L., and Carroll, G. R. (2003). The Fog of Change: Opacity and Asperity in Organizations. *Administrative Science Quarterly*, 48(3):399–432.

Hannan, M. T., Pólos, L., and Carroll, G. R. (2007). *Logics of Organization Theory: Audiences, Codes, and Ecologies*. Princeton University Press, Princeton.

Hotelling, H. (1929). Stability in Competition. *The Economic Journal*, 39(153):41–57.

Hsu, G. (2006). Jacks of all trades and masters of none: Audiences' reactions to spanning genres in feature film production. *Administrative Science Quarterly*, 51(3):420–450.

Hsu, G. and Grodal, S. (2015). Category Taken-for-Grantedness as a Strategic Opportunity: The Case of Light Cigarettes, 1964 to 1993. *American Sociological Review*, 80(1):28–62.

Hsu, G., Hannan, M. T., and Koçak, Ö. (2009). Multiple Category Memberships in Markets: An Integrative Theory and Two Empirical Tests. *American Sociological Review*, 74(1):150–169.

Kauffman, S. A. and Weinberger, E. D. (1989). The NK model of rugged fitness landscapes and its application to maturation of the immune response. *Journal of Theoretical Biology*, 141(2):211–245.

Kennedy, M. T. (2008). Getting Counted: Markets, Media, and Reality. *American Sociological Review*, 73(2):270–295.

Koçak, Ö., Hannan, M. T., and Hsu, G. (2014). Emergence of Market Orders: Audience Interaction and Vanguard Influence. *Organization Studies*, 35(5):765–790.

Kovács, B. and Hannan, M. T. (2015). The Space of Categories and the Consequences of Category Spanning. *Sociological Science*, 2:252–286.

Kovács, B. and Liu, C. (2016). Audience structure and status multiplicity. *Social Networks*, 44:36–49.

Lancaster, K. (1990). The Economics of Product Variety: A Survey. *Marketing Science*, 9(3):189–206.

Le Mens, G., Hannan, M. T., and Pólos, L. (2011). Founding Conditions, Learning, and Organizational Life Chances: Age Dependence Revisited. *Administrative Science Quarterly*, 56(1):95–126.

Lee, B. H., Hiatt, S. R., and Lounsbury, M. (2017). Market Mediators and the Tradeoffs of Legitimacy—Seeking Behaviors in a Nascent Category. *Organization Science*, (forthcoming):1–57.

Lenox, M. J., Rockart, S. F., and Lewin, A. Y. (2007). Interdependency, Competition, and Industry Dynamics. *Management Science*, 53(4):599–615.

Leung, M. D. (2014). Dilettante or Renaissance Person? How the Order of Job Experiences Affects Hiring in an External Labor Market. *American Sociological Review*, 79(1):136–158.

Leung, M. D. and Sharkey, A. J. (2013). Out of Sight, Out of Mind? Evidence of Perceptual Factors in the Multiple-Category Discount. *Organization Science*, forthcomin.

Levinthal, D. A. (1991). Random Walks and Organizational Mortality. *Administrative Science Quarterly*, 36(3):397–420.

Levinthal, D. A. (1997). Adaptation on Rugged Landscapes. *Management Science*, 43(7):934–950.

Lo, J. Y., Fiss, P. C., Rhee, E. Y., and Kennedy, M. T. (2020). Category Viability: Balanced Levels of Coherence and Distinctiveness. *Academy of Management Review*, 45(1):85–108.

Malt, B. (1995). Category Coherence in Cross-Cultural Perspective. *Cognitive Psychology*, 29(2):85–148.

March, J. G., Sproull, L. S., and Tamuz, M. (1991). Learning from Samples of One or Fewer. *Organization Science*, 2(1):1–13.

McFadden, D. (1974). The measurement of urban travel demand. *Journal of Public Economics*, 3(4):303–328.

McFadden, D. L. (1984). Chapter 24 Econometric analysis of qualitative response models. In *Handbook of Econometrics*, volume 2, pages 1395–1457. Elsevier.

Medin, D. L., Lynch, E. B., Coley, J. D., and Atran, S. (1997). Categorization and Reasoning among Tree Experts: Do All Roads Lead to Rome? *Cognitive Psychology*, 32(1):49–96.

Murphy, G. L. (2004). *The Big Book of Concepts*. A Bradford Book. MIT Press, Cambridge, Mass., 1. mit press paperback ed edition.

Negro, G. and Leung, M. D. (2013). "Actual" and Perceptual Effects of Category Spanning. *Organization Science*, 24(3):684–696.

Nelson, R. R. and Winter, S. G. (1982). Organizational Capabilities and Behavior. In *An Evolutionary of Economic Change Theory*.

Paolella, L. and Durand, R. (2016). Category Spanning, Evaluation, and Performance: Revised Theory and Test on the Corporate Law Market. *Academy of Management Journal*, 59(1):330–351.

Podolny, J. M. (2008). *Status Signals: A Sociological Study of Market Competition.* Princeton University Press, Princeton.

Pontikes, E. and Hannan, M. (2014). An Ecology of Social Categories. *Sociological Science*, 1:311–343.

Pontikes, E. G. (2012). Two Sides of the Same Coin: How Ambiguous Classification Affects Multiple Audiences' Evaluations. *Administrative Science Quarterly*, 57(1):81–118.

Pontikes, E. G. and Kim, R. (2017). Strategic Categorization. In Durand, R., Granqvist, N., and Tyllström, A., editors, *From Categories to Categorization: Studies in Sociology, Organizations and Strategy at the Crossroads Research (Research in the Sociology of Organizations)*, volume 51, pages 71–111. Emerald Publishing Limited.

Pratt, J. W. (1964). Risk Aversion in the Small and in the Large. *Econometrica*, 32(1):122–136.

Proffitt, J. B., Coley, J. D., and Medin, D. L. (2000). Expertise and category-based induction. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26(4):811–828.

Rahmandad, H. (2019). Interdependence, Complementarity, and Ruggedness of Performance Landscapes. *Strategy Science*, 4(3):234–249.

Rao, H., Monin, P., and Durand, R. (2005). Border Crossing: Bricolage and the Erosion of Categorical Boundaries in French Gastronomy. *American Sociological Review*, 70(6):968–991.

Ries, E. (2011). *The Lean Startup: How Today's Entrepreneurs Use Continuous Innovation to Create Radically Successful Businesses.* Crown Business, New York, 1st ed edition.

Rivkin, J. W. (2000). Imitation of Complex Strategies. *Management Science*, 46(6):824–844.

Roberts, J. H. and Lattin, J. M. (1991). Development and Testing of a Model of Consideration Set Composition. *Journal of Marketing Research*, 28(4):429–440.

Rosch, E. and Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology*, 7(4):573–605.

Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., and Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology*, 8(3):382–439.

Ruef, M. (2000). The Emergence of Organizational Forms: A Community Ecology Approach. *American Journal of Sociology*, 106(3):658–714.

Salop, S. C. (1979). Monopolistic Competition with Outside Goods. *The Bell Journal of Economics*, 10(1):141–156.

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, 6(2):461–464.

Sgourev, S. V. and Althuizen, N. (2014). "Notable" or "Not Able": When Are Acts of Inconsistency Rewarded? *American Sociological Review*, 79(2):282–302.

Shocker, A. D., Ben-Akiva, M., Boccara, B., and Nedungadi, P. (1991). Consideration set influences on consumer decision-making and choice: Issues, models, and suggestions. *Marketing Letters*, 2(3):181–197.

Smith, E. B. (2011). Identities as Lenses: How Organizational Identity Affects Audiences' Evaluation of Organizational Performance. *Administrative Science Quarterly*, 56(1):61–94.

Spence, M. (1973). Job Market Signaling. *The Quarterly Journal of Economics*, 87(3):355.

Steele, R. J. and Raftery, A. E. (2009). Performance of Bayesian Model Selection Criteria for Gaussian Mixture Models. Technical report.

Strang, D. and Macy, M. W. (2001). In Search of Excellence: Fads, Success Stories, and Communication Bias. *American Journal of Sociology*, 107(1):147–82.

Teece, D. J., Pisano, G., and Shuen, A. (1997). Dynamic Capabilities and Strategic Management. *Strategic Management Journal*, 18(March):509–533.

Vergne, J. P. and Wry, T. (2014). Categorizing Categorization Research: Review, Integration, and Future Directions. *Journal of Management Studies*, 51(1):56–94.

Weber, K., Heinze, K. L., and DeSoucey, M. (2008). Forage for Thought: Mobilizing Codes in the Movement for Grass-fed Meat and Dairy Products. *Administrative Science Quarterly*, 53:529–567.

Weber, R. A. and Camerer, C. F. (2003). Cultural Conflict and Merger Failure: An Experimental Approach. *Management Science*, 49(4):400–415.

Wernerfelt, B. (1984). A resource-based view of the firm. *Strategic Management Journal*, 5(2):171–180.

White, H. C. (1981). Where Do Markets Come From? *American Journal of Sociology*, 87(3):517–547.

Xu, J., Dowman, M., and Griffiths, T. L. (2013). Cultural transmission results in convergence towards colour term universals. *Proceedings of the Royal Society B: Biological Sciences*, 280(1758):20123073.

Xu, R. and Wunsch II, D. (2005). Survey of clustering algorithms for MANET. *IEEE Transactions on Neural Networks*, 16(3):645–678.

Zhao, E. Y., Fisher, G., Lounsbury, M., and Miller, D. (2017). Optimal Distinctiveness: Broadening the Interface between Institutional Theory and Strategic Management. *Strategic Management Journal*, 38(1):93–113.

Zhao, E. Y., Ishihara, M., Jennings, P. D., and Lounsbury, M. (2018). Optimal Distinctiveness in the Console Video Game Industry: An Exemplar-Based Model of Proto-Category Evolution Optimal Distinctiveness in the Console Video Game Industry: An Exemplar-Based Model of Proto-Category Evolution. Organization Science. *Organization Science*, 29(4).

Zuckerman, E. W. (1999). The Categorical Imperative: Securities Analysts and the Illegitimacy Discount. *American Journal of Sociology*, 104(5):1398–1438.

Zuckerman, E. W. (2017). The Categorical Imperative Revisited: Implications of Categorization as a Theoretical Tool. In Durand, R., Granqvist, N., and Tyllström, A., editors, *From Categories to Categorization: Studies in Sociology, Organizations and Strategy at the Crossroads Research (Research in the Sociology of Organizations)*, volume 51, pages 31–68. Emerald Publishing Limited.