Who Depends on Why

Toward an Endogenous, Purpose-driven Mechanism in Organizations' Reference Selection

Luger, Johannes

Document Version
Final published version

Published in:
Strategic Management Journal

DOI:
10.1002/smj.3486

Publication date:
2023

License
CC BY-NC-ND

Citation for published version (APA):
Luger, J. (2023). Who Depends on Why: Toward an Endogenous, Purpose-driven Mechanism in Organizations' Reference Selection. *Strategic Management Journal*, *44*(8), 2035-2059. https://doi.org/10.1002/smj.3486

Received: 14 November 2021 | Revised: 5 January 2023 | Accepted: 9 January 2023 | Published on: 1 February 2023

DOI: 10.1002/smj.3486

RESEARCH ARTICLE

STRATEGIC MANAGEMENT JOURNAL WILEY

Who depends on why: Toward an endogenous, purpose-driven mechanism in organizations' reference selection

Johannes Luger

Department of Strategy and Innovation, Copenhagen Business School, Frederiksberg, Denmark

Correspondence
Johannes Luger, Department of Strategy and Innovation, Copenhagen Business School, Kilevej 14A, K. 2.70, 2000 Frederiksberg, Denmark.
Email: jlu.si@cbs.dk

Funding information
Schweizerischer Nationalfonds zur Förderung der Wissenschaftlichen Forschung, Grant/Award Number: 100013M 204670

Abstract

Research Summary: This paper investigates how firms select reference organizations, that is, other firms to which they compare themselves. We question the exogenous nature of references (i.e., them being defined via industry-categorizations) but suggest that, via motivations or purposes, firms endogenously select them. We evaluate our findings when analyzing proprietary data on hotels' self-selection of comparison-hotels. In support of our arguments, we find that in situations of increased uncertainty regarding firms' own relative abilities and standing, firms make adjustments to their selected references toward more similar ones. This enables them to obtain more diagnostic information about their relative abilities and this effect holds constant of (exogenous) industry-entry or exit events. Our findings contribute to an updated understanding about the role of comparison organizations in firms' decision-making.

Managerial Summary: Prior work shows that comparisons with other firms (i.e., references) play an important role for our understanding of firms' decision-making. For example, performance comparisons with references can trigger search or a decision-need, ultimately, leading to acquisition-decisions, new-product-introductions, and

the like. When questioning the selection of such references, prior work has typically derived them from (exogenous) industry-categorizations. We review this practice by relying on rare, longitudinal data on firms' reference self-selection. When controlling for industry level effects, we find that firms adapt references as a function of changes in their comparison needs (e.g., self-assessment). This is important because it implies an endogenous reference selection mechanism and shifts the attention from industry-categorizations toward an understanding of comparison needs and their emergence when attempting to understand firms' decision-making.

KEYWORDS

behavioral strategy, performance feedback, problemistic search, reference groups, social comparison theory

Comparisons between organizations play a central role in many management theories focused on senior executives' decision-making. Examples include performance-feedback theory (Greve, 1998), vicarious learning and imitation (Posen, Ross, Wu, Benigni, & Cao, 2022), impression management and reputation transfer (K. H. Kim & Tsai, 2012), and institutional theory (DiMaggio & Powell, 1983). According to these theories, comparisons with others (i.e., references) affect important decisions and actions, such as acquisitions (Greve, 2011), product introductions (Joseph & Gaba, 2015), and numerous others (Baum & Dahlin, 2007).

In the above-mentioned decision-making theories, reference-group comparisons are typically, empirically and theoretically, treated as exogenous, industry-derived. For example, comparisons with competitors based on industry code (e.g., Iyer & Miller, 2008) or geographical distance (e.g., Greve, 2011) are associated with problem identification, which in turn triggers the mentioned actions and decisions.

We suggest that the extant literature's consideration of references comes with an empirical and theoretical complication. Empirically, data on firms' self-selection of reference groups is scarce and difficult to observe, causing much of this work to be assumption-driven, such that reference groups are indirectly derived via competition-based approximations. Greve and Gaba (2017, p. 31) mention that prior work suffers from a key problem in this regard—organizations selected for comparison are not easy to observe and are typically "justified contextually or determined endogenously from the data."

Theoretically, we claim that a competition- or industry-derived selection fails to consider the specific motivation or purpose underlying such selection. Prior work often conceptually mentioned such purposes, such as the goal of self-assessment (or benchmarking) in the case of performance feedback (e.g., Greve, 1998) or the goal of learning (or best-practice sharing) in the case of vicarious learning (e.g., Li, Qian, & Yao, 2015). This, however, also implies that if purposes underlie comparisons, these purposes, regardless of competition or industry affiliation arguments, should also drive reference selection. Eventually, this means that reference groups are not exogenously defined (e.g., via industry affiliations) but endogenously derived based on a given comparison need or purpose.

Given these theoretical and empirical shortcomings, we aim to better understand the mechanisms behind firms' reference self-selection. To do so, we rely on social comparison theory (SCT) (Festinger, 1954), which explains reference selection at the individual level. We apply this theory to the organization or senior managers and suggest that purposes, or determinants of them (Kulik & Ambrose, 1992), drive firms' reference selection. Specifically, we focus on the performance-feedback debate and the associated self-assessment purpose, suggesting that firms compare to others in order to derive an accurate view of their own relative abilities and standing. Research suggests that this purpose is typically triggered in situations of uncertainty regarding one's own relative (cap-)abilities and standing (Greve, 2008; Sedikides & Strube, 1997). Overall, in order to evaluate our theory, we analyze whether variations in purpose (i.e., variations in relative uncertainty) affect subsequent changes in the self-selected references, while controlling for potential exogenous, industry-derived changes in the reference group.

We analyze the STR competitive set database, which contains rare, longitudinal data on firms' self-selection of references. STR is the largest data provider in the hospitality industry. One of its services enables hotel properties[1] to select a group of other hotels (i.e., references), and receive information on the operations and performance of hotels in that group. With this data as well as full population-level data on hotels entering or leaving the market, we can examine whether variation in (endogenous) purposes or (exogenous) industry affiliations drive subsequent adaptations in hotels' self-selected references.

We find evidence that supports our theory. Specifically, for profit-oriented firms, we operationalize uncertainty regarding one's own relative (cap-)abilities and standing using deviations from important performance benchmarks.[2] We find that the higher those deviations, the higher the likelihood of subsequent reference group adaptations and that this association holds regardless of hotels entering or leaving the industry (i.e., exogenous sources of reference changes). Furthermore, while deviations in relative performance might, in themselves, be endogenous, we leverage an Airbnb-related regulatory shock that increased relative uncertainty for one segment of the hospitality industry (i.e., the low end but not high end). A corresponding differences-in-differences analysis provides further support for our theory.

Finally, we find that performance deviations lead to comparisons with more similar reference groups in terms of both class affiliation and profitability level. This effect also serves to document the diagnosis mechanism underlying self-assessment, but not alternative motivations, such as self-enhancement. When uncertainty about one's relative capabilities increases, comparisons with more similar others provide more accurate diagnoses. Conversely, the results of comparisons with dissimilar others may be ambiguous because the resulting performance differences are most likely due to differences in attributes rather than abilities. Altogether, we contribute by providing evidence of an endogenous reference selection mechanism that has far-reaching implications for the above-mentioned theories relying on comparisons.

[1]From here onward, we use the terms "hotel" and "property" interchangeably. They describe individual hotel properties at a specific location. We use the terms "corporation" or "hotel chain" when describing multiple hotels (or properties) associated via a common brand or parent company.

[2]As we explain later in more detail, for profit-oriented organizations, relative performance is one of the most important signals concerning the focal firm's state or relative standing in the market. As firms typically have an incentive to meet their chosen benchmarks (e.g., Yu, 2006), any (unintended) deviation from performance benchmarks signals uncertainty about this firm's state or condition.

1 | ALTERNATIVE PERSPECTIVES ON REFERENCE SELECTION IN ORGANIZATIONS

Prior work provides comprehensive evidence on the importance of inter-organizational comparisons for our understanding of firm actions (Baum & Dahlin, 2007; Kacperczyk, Beckman, & Moliterno, 2015). In many decision-making theories, comparisons regarding, for instance, performance (Greve, 1998), products or services (Baum, Li, & Usher, 2000), or reputation (K. H. Kim & Tsai, 2012) are viewed as important intervening mechanisms. For example, in performance-feedback theory, performance comparisons result in the definition of a problem, which leads to a search for solutions (Greve, 1998). Although these comparisons are important for the theory, they have not been studied in detail. In fact, the problemistic-search process has been criticized as a "black box, examining the correlation between a performance shortfall that triggers the process and the changed outcome" (Posen, Keil, Kim, & Meissner, 2018, p. 209).

Prior work has adopted various perspectives on reference-group selection or definition. The most established view takes an external perspective and focuses on how industry observers (e.g., researchers, security analysts) aim to structure or categorize industries or markets (e.g., Durand & Paolella, 2013; McGee & Thomas, 1986). From an internal perspective, scholars question how managers define competitors or interpret industry classifications (e.g., Reger & Palmer, 1996). In the following, we briefly summarize the evolution of this literature. We conclude by arguing that the extant work has often combined the two perspectives by, for instance, studying firm-internal decision-making while relying on the external definition of references (e.g., Henisz & Delios, 2001; Iyer & Miller, 2008). We suggest that this combination has given rise to several inaccuracies in the literature, requiring a re-examination of the role of references in firms' decision-making.

Within the fields of economics and management, the most established view on firm comparisons relates to competitor identification (see Gur & Greckhamer, 2019, for a recent literature review). Derived from competitive strategy (Porter, 1980), this perspective highlights that a firm's performance and prosperity are direct functions of its competitive environment (e.g., Chen, 1996). This implies that firms and their senior managers need a sufficient understanding of (or comparison with) their closest competitors.

The competitor-identification argument is rooted in the industrial organization (IO) view, which conceptualizes competition occurring within industry boundaries (Thomas & Pollock, 1999). Work adopting this perspective typically relies on industry codes (e.g., SIC, NAICS) or geographical approximations. Most of the work considering references in firms' decision-making related to acquisitions (Iyer & Miller, 2008), new product introductions (Gaba & Joseph, 2013), and similar issues (Kacperczyk et al., 2015) relies on this perspective.

The strategic-groups debate, which emerged from the IO view, offers a more fine-grained argument regarding competition (Hatten & Schendel, 1977; McGee & Thomas, 1986). In contrast to IO, it suggests that the construct of competition (or competitor identification) is more focused than the industry-wide application. Specifically, it highlights competitive sub-groups (or strategic groups), which are often based on more specific criteria, such as niche membership, competitive strategies, geographical coverage, or environmental factors (Hatten & Schendel, 1977; L. Kim & Lim, 1988; McGee & Thomas, 1986).

A third perspective highlights the importance of customers when defining competition. Related work focuses on brand-switching behavior or direct competition for customers (Dawes, 2014; DeSarbo, Grewal, & Wind, 2006). Notably, none of these three perspectives

determine competitors or reference groups by surveying organizations or senior managers. Instead, they all adopt an external view derived from industry classification.

A fourth perspective, which also originated from the competitor-identification literature, explicitly relies on comparison data collected from senior managers. (e.g., Porac, Thomas, & Baden-Fuller, 1989; Reger & Huff, 1993). This perspective takes a first step toward understanding the direct, idiosyncratic role of individuals in organizations' reference selection. This stream of literature demonstrates that competitor identification originates from managers' "cognitive maps of competitive positioning" (Reger & Palmer, 1996, p.22), thereby leading to "competitive groups as cognitive communities" (Porac et al., 1989). Overall, this perspective highlights the idiosyncratic and independent effects of senior managers in reference selection. However, it does not question the exogenous nature of this selection. In other words, although managers have idiosyncratic understandings of competition, which often substantially differ across individuals (and from an external perspective (Porac et al., 1989)), reference groups still derive from industry classifications (with managers interpreting them) and are, hence, exogenous to their decisions.

We claim that the extant work suffers from theoretical imprecision, mostly because it was originally developed in order to understand and analyze certain industries (typically from the external perspective) and was, without much adjustment, applied in work focused on firm-internal decision-making. To some extent, this application is reasonable—as discussed above, when a focal firm's performance is a function of competition, considering competition in decision-making is warranted. However, several conceptual arguments suggest that the role of references from a firm-internal decision-making perspective is different or more nuanced than general industry categorizations or competitor analyses suggest. Examples include firms comparing themselves to non-industry members when they wish to learn about technologies not yet established in their industry (Srinivasan, Haunschild, & Grewal, 2007) and firms comparing themselves with firms "beyond industry boundaries" (Porac, Wade, & Pollock, 1999, p. 112) in order to enhance their legitimacy. Most of this work suggests that comparisons are driven by specific purposes (e.g., learning or legitimacy in the examples just given).

2 | SCT IN ORGANIZATIONS

When questioning how purposes drive comparison decisions, especially decisions made by individuals or managers, SCT is the most established starting point (Festinger, 1954). This theory suggests that four purposes, or motivations, drive humans' comparison decisions: self-assessment (i.e., the motivation to compare in order to accurately assess oneself); self-improvement (i.e., the motivation to compare in order to learn and improve); self-verification (i.e., the motivation to compare in order to justify oneself); and self-enhancement (i.e., the motivation to compare in order to present oneself in a positive light) (Sedikides, 1993). In the management literature, SCT has mostly been applied to comparisons between employees or managers (Baumann, Eggers, & Stieglitz, 2019).

We claim that SCT is also applicable to senior managers who select references (i.e., comparison organizations) on behalf of their organizations. This is because the main purposes mentioned in the theory also apply to organizations. In simple terms, when organizations aim for prosperity and survival, they compare themselves to others in order to accurately assess their standing relative to the competition, they compare to learn, to present themselves in a positive light, and so on. For example, consider the most established individual-level comparison motive of self-assessment. This logic is reminiscent of the overall argument behind competitor

identification—when senior managers compare their firms to competitors in order to derive an accurate view of their firms' relative standing and abilities in an industry, this represents an organizational-level application of the self-assessment rationale. Similar parallels apply to the other comparison purposes. For example, consider the individual-level self-improvement comparison rationale (i.e., the motivation to compare in order to learn and improve). This is reminiscent of the debates on vicarious learning (Baum & Dahlin, 2007) or best-practice sharing, which suggest that organizations compare with others in order "to learn successful practices from other organizations" (Goh & Richards, 1997, p. 582).

3 | PURPOSE MECHANISMS AS TRIGGERS OF REFERENCE SELECTION

Based on the arguments above, we claim that references are not only exogenously (industry) derived, but also endogenously emerge from certain purposes. In this paper, we focus on the self-assessment rationale, which is highly relevant in organizations (e.g., Greve, 1998, p. 60).

The identification of purposes is difficult.[3] We therefore follow prior work (Kulik & Ambrose, 1992) in identifying comparison purposes via their determinants. In the case of self-assessment, this determinant is typically assumed to be uncertainty concerning one's own relative abilities (Sedikides, 1993). Whenever such uncertainty is high, the desire to obtain "diagnostic information" that may allow for an accurate assessment of oneself relative to others grows (Sedikides & Strube, 1997). Albeit indirectly, a similar argument is presented at the organizational level. Boyd and Fulk (1996) suggest that perceived uncertainty in the competitive environment increases firms' environmental-scanning activities. This implies that the greater a focal firm's uncertainty regarding its own relative (cap-)abilities, the greater the need for diagnosis and, hence, the likelihood of environmental scanning in form of observing other organizations.

When applied to profit-oriented organizations, we suggest that uncertainty concerning firms' relative abilities is best approximated via deviations from performance benchmarks. Performance is one of the most important indicators of the condition of a for-profit firm and any deviation from this yardstick signals uncertainty about the firm's condition. Moreover, in order to limit firm-internal and firm-external uncertainty, firms typically have an incentive to meet such benchmarks (Yu, 2008). According to the self-assessment logic, when aiming to deal with such increased, *relative to others*, uncertainty, comparisons with others or adjustments in the selected reference group are likely coping strategies. This enables the comparing firm to obtain more diagnostic information and, thereby, reduce its relative uncertainty.

The logic used in the debates on problemistic-search and performance feedback (Posen et al., 2018)—debates that explicitly mention the self-assessment rationale (Greve, 1998)—equally points to a positive association between a deviation from performance benchmarks and subsequent adjustments to organizations' reference groups. According to these discussions, a performance shortfall signals the existence of an underlying problem, the identity of which is *uncertain*, leading to problemistic-*search*. As such, uncertainty is a direct consequence of the unintended nature of a shortfall. In other words, for profit-oriented organizations, an accurate understanding of one's own relative abilities (i.e., one's own abilities as well as those of

[3]Inter alia, this is because even when directly surveying managers in organizations, prior work highlighted that managers would not accurately report on multiple purposes, for example, on self-enhancement (Tetlock & Manstead, 1985).

reference others) should not yield a performance deviation in the first place. This also implies that falling short of aspirations reflects uncertainty about a firm's relative (cap-)abilities. Turning to "others"— organizations likely to provide diagnostic information that can reduce this uncertainty—is one likely coping strategy. In formal terms:

> **Hypothesis 1 (H1).** A deviation from performance benchmarks is positively associated with subsequent adaptations to a firm's reference group.

Purpose- or motivation-related studies demonstrate that loss frames have stronger effects on behavior than comparable gain frames (Tversky & Kahneman, 1974). This effect has also been found in organizations. The extant research highlights that, especially in cases of information and attention overload, decision makers typically classify outcomes into simplistic categories (i.e., failures and successes), and that motivation is higher in the more threatening failure category (March & Simon, 1958). In relation to our uncertainty construct (i.e., via performance deviations), the literature shows that negative performance deviations have a stronger effect on subsequent behavior than positive deviations (e.g., Greve, 1998, 2011). Therefore, in our study, loss-framed uncertainty should have a stronger effect on subsequent reference-selection adaptations than success-framed uncertainty.

Similarly, the problemistic-search debate claims that reference-group adaptations should be stronger in the case of negative than positive performance deviations (Posen et al., 2018). In this debate, the search motive (for self-assessment) precedes reference-group selection and is stronger in the case of negative deviations (Greve, 1998). In formal terms:

> **Hypothesis 2 (H2).** A negative deviation from performance benchmarks is more strongly associated with subsequent adaptations to a firm's reference group when compared to a positive deviation.

In studies of the association between relative uncertainty and subsequent adaptations to selected references, the desire to obtain diagnostic information is a central mechanism. In individuals, this has been documented in the selection of individuals with attributes highly similar to the focal, comparing individual (Sedikides & Strube, 1997). Such attribute similarities enable the comparing individual to obtain more diagnostic information, and allow for distinction between comparison objects of high and low ability (Sedikides, 1993, p. 317).

Although organizations are more complex than individuals, similar arguments apply—only a comparison with an organization with similar attributes (in the case of hotels, we suggest that the most important attributes concern class affiliation [e.g., luxury, economy]) can provide diagnostic information. Conversely, when comparing oneself with an organization with unrelated attributes (e.g., a hotel from a different class), the result is likely to be ambiguous because the performance differences tend to reflect the different attributes (i.e., class membership) rather than differences in inherent ability. In short, according to Suls et al. (2002, p. 159), an answer to the question "Am I as good as I ought to be"? requires a comparison to a reference group with similar attributes. In the case of hotels, this implies a need for reference re-selection that favors properties in the same class. We therefore suggest:

> **Hypothesis 3 (H3).** A deviation from performance benchmarks is positively associated with subsequent adaptations in a firm's reference group toward references with more similar characteristics (i.e., a higher percentage of references from the same quality class).

4 | METHODOLOGY

4.1 | Empirical context and data

We test our hypotheses using STR competitive set data. STR is the largest data provider for the US hospitality industry. One of its services allows hotels to request information on its self-selected reference groups' operations and performance.[4] Interviews with senior hotel managers using this service suggest motivations (and associated use-cases) strongly reminiscent of the four purposes discussed above. Most often, the managers we interviewed pointed to benchmarking, followed by learning- or improvement-related motivations.

We started by randomly selecting 300 hotels active in the local markets of New York City Nashville, Tennessee; and Detroit, Michigan. We selected these three markets in order to avoid observing particularities of established, growing, and declining markets. From the 900 hotels, roughly half had made use of STR's competitive set data and, consequently, we could observe their reference selection (data usage was equally distributed across the locations, at 148, 144, and 140, respectively). Hotels subscribing to this service can obtain information on their self-selected references at a fixed price so there are no marginal costs when adding additional references. In order to find changes in hotels' reference-group selection, we made four consecutive annual retrievals (2016–2019). The hotel properties in our sample spanned various classes (e.g., luxury, economy), affiliations (e.g., chain affiliated versus independent), sizes (i.e., number of rooms), and amenities (e.g., spa, golf course, conference venue).

We matched the competitive set data with STR's property data to obtain information on the most important operations- and performance-related data. This includes information on the industry's central performance metrics, occupancy and the average daily rate (ADR). In their multiplication, the two yield in revenue per available room (RevPar), the industry's dominant performance indicator (Canina, Enz, & Harrison, 2005). Furthermore, we matched our data with STR's profit and loss data, which provided detailed information on the sources of income (e.g., spa, food and beverage) and costs (e.g., renovation, human resources).

We believe that this data, especially the industry's focus on RevPar, matches well with our theoretical arguments. The industry heavily relies on standards or categorizations (e.g., hotel classes, star ratings). Moreover, within such standards, RevPar is the core discriminating and most widely used metric for comparison (Baum & Lant, 2003). This is because RevPar is broadly comparable, widely available, and closely monitored (Canina et al., 2005), which implies that it fits well with our theoretical arguments.

4.2 | Dependent variables (DV)

Our main dependent variable, reference-group adaptations, is a dummy variable that captures whether a focal hotel adjusted (i.e., discontinued an existing reference, adopted a new reference) its selected references in t_0 relative to t_{-1}. In addition, we construct a change-count variable. We believe that this variable is particularly informative for a variety of reasons. First, it directly relies on senior managers' self-selected references, which are invisible to the outside

[4]Given STR's dominance on the market, it is almost impossible to find a substitute for this service (i.e., obtain alternative reference information). While hotels subscribing to this service can select individual properties to their reference group, we (as researchers) can only identify these properties via an anonymized indicator.

and, hence, unaffected by impression-management motivations. Second, in contrast to prior work (e.g., Audia, Rousseau, & Brion, 2021), this variable is unaffected by policy or regulation, and therefore provides a more accurate view of senior managers' unconstrained reference self-selection. Third, given the functioning of the database, existing references continue until senior managers make an active decision to change them. All of these factors make our main DV particularly informative.

We construct an alternative DV that details an important directionality in reference-group adaptations. More specifically, in order to evaluate H3, we focus on similarities between references and the selecting hotel. As class categorizations are among the most important attributes in the hospitality industry, we rely on this information. Specifically, *relative same class* measures the ratio of references in the same class as the focal selecting hotel relative to all selected references. For example, if a hotel selects 10 references and 5 of them share the same class as the focal hotel, this ratio amounts to .5. Furthermore, in order to better understand the changes made to this ratio, we created count variables detailing the number of additions and deletions of same or different class references.

4.3 | Independent variables (IV)

In line with prior work, we rely on two yardsticks of firms' performance deviations: a hotel's own prior performance and that of reference others. In the former case, we subtract prior from focal year RevPar and create variable splines (i.e., we differentiate between positive and negative deviations). Both splines are coded positively, such that higher values signal higher (positive or negative) deviations from prior performance.

For comparisons with reference others, we use three alternatives. We subtract a hotel's RevPar from the average RevPar of all hotels in: (a) its self-selected reference group; (b) the same class (e.g., luxury, economy); and (c) the same class and location. We selected class and location because both are central parameters to the industry (Baum & Lant, 2003). The overlap in these groups is limited, such that statistical separation is still possible. For example, only 46% of the firms selected into reference groups were from the same class as the focal selecting hotel. Similar to prior performance, we spline these variables and re-code them to positive values only.

4.4 | Analysis and control variables

Our data are structured as a hotel property-annual panel. Over time, hotel properties are highly stable entities that operate within the previously mentioned fixed industry categories and locations (e.g., the economy segment in New York City). Therefore, property-level fixed effects should account for a large share of the variance in our sample. Apart from sporadic renovations, the most important factors changing over time concern hotels' actual bookings, which are reflected in indicators such as occupancy and ADRs, both of which merge in our main performance comparison metric, RevPar.

As stated in the theory section, our introduction of a novel, purpose-driven selection mechanism that stands free from the extant literature's focus on industry-driven mechanisms requires us to control for the latter. We do so in several ways. First, by obtaining population level data from STR about all hotels entering and leaving our locations, we were able to construct an

industry entry/exit variable. This variable counts all properties entering or leaving any given class in a given city during a given year and, hence, directly captures changes in the competitive landscape. Second, to make more detailed use of this population-level entry/exit data, we adjust our dependent variable (i.e., adaptations to a hotel's reference group) by excluding any addition/deletion of a reference that occurs in parallel to that same property entering or leaving the industry.

Given the above explanations, we operationalize our main models as logit panel models with property-level and time fixed effects. While these are our main models, we later also detail changes to the dependent variable, the model structure, the vector of control variables, and the structure of the error term.

$$a_{i,t} = \beta_1 x^{HA}_{i,t} + \beta_2 x^{SA}_{i,t} + \beta_3 \text{Entry/exit}_{j,k,t} + \eta_i + \tau_t + \varepsilon_{i,t}$$

where:

$a_{i,t}$ is a dummy equaling 1 if the reference firms selected into property i's reference.
group in t_1 is different from t_0 and 0 otherwise.
x^{HA} is the difference of a property's RevPar_{t1} from that of the prior year (RevPar_{t0}).
x^{SA} is the difference of a property's RevPar_{t0} from the average of reference firms' RevPar_{t0}.
Entry/$\text{exit}_{(j,k,t)}$ is a count of hotels entering/leaving a class_j in a city_k at a time_t.
η is a property's fixed effect.
τ is a time fixed effect.
ε is the error term.

While we believe that the RevPar variable accounts for large parts of the within-property variation over time, we still want to ensure that potential omitted variables do not interfere with our results. For example, fundamental adjustments to a property's market positioning or strategy (approximated via renovation costs, property and maintenance costs, spa expenses, marketing expenses, and chain or franchise affiliations) may simultaneously affect a property's current profitability and its selected reference group. Accordingly, we include these variables from the matched STR P&L data.

In addition to the logit models, we implement Poisson and negative binomial models to estimate the "change-count" models. In these models, we include property-level fixed effects, year fixed effects, and standard errors clustered at the property level to account for any potential chain-level policies in reference selection. Finally, in order to ensure that overlooking the properties that do not change their references (or that change references in any period) does not affect our results, we run OLS regressions with property-level fixed effects.[5]

5 | RESULTS

Over the observation period (i.e., 4 years), the 432 reporting hotels listed a total of 8,982 reference properties. This amounts to an average of 5.2 references per hotel and year. Descriptive evidence suggests a high degree of heterogeneity within these self-selected reference groups. Only 46% were from the same class, while 88% were from the same class or one class higher/

[5]Another argument in favor of OLS over logit models goes back to various limitations discussed for logistic regression models. For more, see (King & Zeng, 2001).

lower. There is no indication that hotels systemically select upward or downward in terms of class membership, and we also find no systematic evidence of selecting systematically higher- or lower-performing peers (based on RevPar). On average, hotels selected peers generating USD 1.6 less RevPar than they generated themselves. According to a *t*-test, this is not statistically different from zero. However, the distributions of this performance-comparison variable (i.e., a standard deviation of USD 41.4 with a mean of USD 1.6) and the class-comparison variable indicate a high degree of heterogeneity between the focal selecting hotel and the hotels selected into the reference group. Overall, we believe that this heterogeneity signals early support for purpose-driven mechanisms. While selection in accordance with certain competition- or industry-derived metrics would speak in favor of homogeneous reference groups (e.g., most from the same class, similar RevPar levels), the diversity of purposes outlined in the theory, including the required diversity in reference firms' characteristics to serve those purposes, is more aligned with the descriptive evidence.

Another interesting observation pertains to the correlations among different performance-deviation indicators. First, we find low correlation between deviations from historical performance and those of self-selected reference others. However, in the case of deviations from all hotels in the same class (or the same class and location), we find correlations ranging from .25 to .37 (correlations between self-selected references and all same-class/same-location hotels range from .31 to .59). Altogether, this suggests that hotels use deviations from self-selected references as an independent source of information, especially relative to the information obtain from their own historical performance comparisons. Table 1 summarizes the descriptive evidence from the variables used in our models.

> **Hypothesis 1.** Of the 432 hotels in our sample, 231 had at least one change in their reference group during the observation window and 35 of these hotels reported a change in every year. The remaining 196 hotels serve as the basis for our main logit models.[6] As shown in the models in Table 2, the coefficients are consistently positive. These coefficients indicate that any deviation (positive and negative) relative to the firm's prior performance or relative to the firm's self-selected references is positively associated with a subsequent change in the self-selected references. This is consistent evidence that all forms of relative uncertainty (i.e., any deviation from the explained benchmarks) are associated with a higher likelihood of subsequent reference group change. Notably, these results hold firm when controlling for population-level entry or exit of properties in a given city, class, and year.[7]

Table A3 in the Online Appendix replicates the models shown in Table 2 when using an adjusted DV. As mentioned in the variable description, this DV corrected for any reference change that occurred in conjunction with a property newly registering (or de-registering) in the respective time period. In total, the DV was corrected in five instances, leading to models with a higher model fit when compared to Table 2. We view these adapted results as offering more evidence of the existence of a free-standing, purpose-based motivation as the models shown in Table A3 are characterized by lower noise and corrected for potential reference changes going

[6]In our linear probability models, we also account for all sample firms, which enables us to include the counterfactuals of not changing references and changing references each year.

[7]Across our observation period, we observed 67 properties leaving or entering the industry (excluding hotels that were re-branded, retrofitted, etc.).

TABLE 1 Descriptive statistics and correlations.

	Variable name	Mean	S.D.	5%	95%
(1)	Reference group change	0.164	0.370	0	1
(2)	Reference group change corrected	0.161	0.368	0	1
(3)	Count (reference group change)	0.534	1.611	0	4
(4)	Reference addition same class	0.155	0.539	0	1
(5)	Reference deletion same class	0.121	0.457	0	1
(6)	Reference addition different class	0.206	0.661	0	2
(7)	Reference deletion different class	0.193	0.644	0	1
(8)	Relative references same class	0.466	0.271	0	1
(9)	Historical perf. deviation (neg.)	3.607	10.193	0	17.4
(10)	Historical perf. deviation (pos.)	4.011	10.248	0	15.5
(11)	Perf. deviation reference group (neg.)	9.849	22.728	0	42.2
(12)	Perf. deviation reference group (pos.)	11.665	33.289	0	42.7
(13)	Industry entry/exit	0.589	1.046	0	2

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
(2)	0.99											
(3)	0.84	0.85										
(4)	0.64	0.64	0.73									
(5)	0.58	0.59	0.69	0.59								
(6)	0.69	0.70	0.80	0.27	0.39							
(7)	0.66	0.66	0.83	0.47	0.28	0.65						
(8)	0.04	0.05	0.03	0.22	0.1	−0.13	−0.04					
(9)	0.02	0.02	0.02	0	0.03	0.05	−0.01	0.1				
(10)	0.01	0.01	0	0.02	0.02	−0.02	0	0.11	0.16			
(11)	0.08	0.08	0.12	0.13	0.17	0.03	0.07	0.11	0.02	0.29		
(12)	−0.02	−0.02	−0.02	−0.03	−0.02	−0.02	−0.01	0.12	−0.22	0.02	0.15	
(13)	0.1	0.1	0.09	0.11	−0.01	0.05	0.1	−0.02	−0.04	0.06	−0.03	0

back to industry effects. While we believe that these adapted models provide corroborating evidence, we still prefer our original DV for the main models, as registering (or de-registering) does not necessarily exclude purpose-induced selection (e.g., a reference property might have been dropped regardless of whether it left the industry).

We also investigate whether our results are sensitive to the *extent* of reference-group changes. Table A4 in the Online Appendix shows the corresponding results for Poisson and negative binomial models. The respective DV has a mean of .53, whereas properties change an average of 4.1 references conditional on there being a change in the reference group. This combination reveals another interesting observation about firms' reference selection. Firms generally seem reluctant to adjust their reference groups. However, if they do so, they tend to make multiple changes. Subsequent interviews suggested that this is because senior managers value stable time-series reference information. It implies that there is a trade-off in reference selection

TABLE 2 Regression table change in reference group.

	Model 1	Model 2	Model 3
Hist perf. dev. (neg.)	0.031	0.039	0.051
	(.008)	(.003)	(.002)
Hist. perf. dev. (pos.)	0.026	0.028	0.013
	(.072)	(.080)	(.494)
Perf. dev. ref. group (neg.)		0.026	0.021
		(.007)	(.032)
Perf. dev. ref. group (pos.)		0.029	0.032
		(.015)	(.012)
Perf. dev. same class (neg.)			0.035
			(.209)
Perf. dev. same class (pos.)			−0.038
			(.159)
Perf. dev. same class/city (neg.)			−0.019
			(.420)
Perf. dev. same class/city (pos.)			0.003
			(.906)
Industry entry/exit	0.085	0.073	0.072
	(.454)	(.526)	(.536)
Time fixed effects	Included	Included	Included
Property fixed effects	Included	Included	Included
# of observations	581	581	581
# of properties	196	196	196
$\widetilde{\chi}^2$	41.772	55.107	58.528
$\widetilde{\chi}^2(p)$	0.000	0.000	0.000
Pseudo R^2	0.098	0.130	0.138

Note: *p*-values in parentheses.

in terms of balancing the benefits of increased purpose fit obtained via a reference-group change against the downside of interrupting the time series. Finally, with regard to this count DV (and the prior change dummy models), controlling for reference-group size (i.e., the number of references per group) does not affect our findings. Altogether, the regression models concerning this adapted DV provide further evidence in favor of H1.

Logit models limit our observations to properties initiating changes to their references in one or two periods. The models shown in Table A5 illustrate the linear probability models for all properties reporting references (16 firms were excluded due to missing RevPar information). Similar to the previous models, these models provide support for H1.

In Model 3 of Table A5, we test whether our results are affected by omitted variable bias. While we maintain our initial claim that, apart from occupancy and the associated ADRs, most variance in our models should come from across-property variation, we still wish to analyze whether within-property variation over time affects our results. In the context of individual

properties, the most likely sources of longitudinal variation include sporadic renovations, property upgrades (captured via property and maintenance expenses), marketing expenses, or other changes in the composition of a hotel's main business (identified based on the sources of its revenues [e.g., spa] and main fees [e.g., franchise fees]). As indicated in Model 3, the adapted models also provide support for H1, albeit at low levels of statistical power with regard to the operationalization of relative uncertainty via peer-group performance deviations. This low statistical power may either reflect the strong reduction in observations caused by the matching or the existence of omitted variable bias. In order to further evaluate this issue, we ran an additional model using all observations in Model 3, but excluding all time-variant IVs (i.e., those previously included to evaluate the potential for an omitted variable bias problem). The model is highly comparable to Model 3 (i.e., it has a low statistical power). This suggests that the reduction in observations, rather than the inclusion of time-variant control variables, caused the reduction in statistical power, alleviating concerns regard potential omitted variable bias.

Hypothesis 2. In order to evaluate H2, we look for evidence of gain- or loss-framed asymmetrical effects. We compare coefficient magnitudes between negative and positive deviations from historical or peer performance across Tables 2 and 3 via a two-

TABLE 3 Regression table: Direction of ref. group change.

	Model 1 Add same class	Model 2 Del same class	Model 3 Add diff class	Model 4 Del diff class	Model 5 Rel. same class
Hist. perf. dev. (neg.)	0.022	0.032	0.039	0.024	0.000
	(.087)	(.014)	(.009)	(.083)	(.313)
Hist. perf. dev. (pos.)	0.010	0.007	0.019	0.023	0.000
	(.580)	(.711)	(.342)	(.173)	(.261)
Perf. dev. ref. group (neg.)	0.036	0.031	0.007	0.013	0.000
	(.002)	(.008)	(.225)	(.012)	(.017)
Perf. dev. ref. group (pos.)	0.031	0.028	0.017	0.019	0.000
	(.008)	(.076)	(.124)	(.131)	(.000)
Industry entry/exit	0.187	−0.101	−0.147	0.124	0.002
	(.136)	(.495)	(.264)	(.318)	(.000)
Rel. same class (lag)					0.987
					(.000)
Constant					0.002
					(.000)
Time fixed effects	Included	Included	Included	Included	Included
Property fixed effects	Included	Included	Included	Included	Included
# of observations	384	341	434	440	1,187
# of properties	129	115	146	148	403
$\widetilde{\chi}^2$	26.851	11.840	30.822	23.248	3.80 e+06
$\widetilde{\chi}^2$ (p)	0.000	0.000	0.000	0.000	0.000

Note: p-values in parentheses.

sided Wald test. Table A8 (Online Appendix), which summarizes the results of these tests, shows no evidence of significantly higher magnitudes for negative when compared to positive deviations.

These results show no support for H2. However, we believe there is a context-specific explanation—the asymmetry effects documented in motivation-related studies typically go back to "selective attention under conditions of information overflow" effects (Lavie, 1995). In other words, due to information-processing limitations, humans adopt simplified categorizations of "gains" and "losses" (Van Knippenberg, Dahlander, Haas, & George, 2015), which enable them to direct their scarce attention to the more threatening loss category. This directly applies to prior work's typical implementation of industry-wide performance comparisons, implying comparisons with a high number (sometimes thousands) of references. In our context, however, senior managers (pre-) selected a small group of references (r/a five) for which they, via a user-optimized interface, could monitor any performance deviations. Together with the importance of reference monitoring in our context—the hotels in our data pay to obtain reference information—we, therefore, believe that the crucial attention-scarcity explanations do not apply. In simple terms, when actively paying for reference information and pre-selecting few references, the selective-attention argument underlying attention asymmetry effects may no longer be relevant.

Hypothesis 3. H3 focuses on the direction of reference-group adaptations. Specifically, in line with the underlying self-assessment motivation or the need for diagnostic information, it suggests that hotels adapt their selection toward more similar (i.e., same class) references as a result of increased relative uncertainty. When evaluating this statement, we must account for the high levels of stability in reference groups. We therefore include a lagged dependent variable in this model, although doing so may carry the risk of autocorrelation. Therefore, we estimate a generalized least squares (GLS) model with a lagged DV in the case of Model 5 (Table 3). We find that as a consequence of increasing positive and negative deviations from the self-selected references, hotels select relatively more same-class references. While this confirms H3, we aim to further understand the underlying detailed adaptations. Models 1 through 4 examine the detailed deletions and additions of same and different class references. Interestingly, these models suggest that the overall increase in more similar class references is mostly driven by the deletion of different class references. While performance deviations from the reference group impact additions and deletions of the same class references (see Models 1 and 2), there is only evidence that such deviations increase deletions (and no corresponding additions) of different class references (see Models 3 and 4).

Later, below, in a post hoc analysis section, we analyze our models regarding an alternative DV, that is, hotels' peer group adaptations toward higher- or lower-performing peers. Although we present the details of this analysis later, we believe that its outcome further supports H3. Specifically, we find that as hotels perform better or worse than their self-selected references, they adjust their peer group toward similarly performing peers (i.e., toward better or worse performing peers, that is, toward peers in a similar profitability category). In the post hoc analysis section, we discuss why these results support a self-assessment-informed peer-selection logic, further providing evidence for H3.

5.1 | Robustness test

The core of our argument—endogeneity in reference selection—may also be a source of criticism to our empirical operationalization. Therefore, we focused on ensuring more exogenous variation in our IV (i.e., changes to hotels' perceived relative uncertainty). In so doing, we benefited from the strong growth of an atypical competitor at the time—Airbnb. Its strong growth created considerable negative externalities for cities, which led to comprehensive local regulatory changes focused on the low end of the hotel industry (i.e., lodging) (Nieuwland & Van Melik, 2020). More specifically, the rise of Airbnb triggered comprehensive lobbying activities and multiple regulatory efforts by local authorities, which were often appealed by private companies. According to industry observers, this caused substantive uncertainty in the low end of the market (Zervas, Proserpio, & Byers, 2017), which is why we believe it could represent an exogenous source of variation in hotels' relative abilities and standing in the market (i.e., our IV).

We leverage two major regulatory changes in Tennessee (i.e., the Tennessee Short Term Rental Unit Act of 2018; (Renzino, 2018)) and New York City (i.e., the NY Multiple Dwellings or "Airbnb Law" of 2019; (d'Auguste, 2019)), both of which were surrounded by controversy and uncertainty. These two shocks enable us to implement a quasi-experimental, differences-in-differences estimation to compare pre- and post-event reference selection between treated (i.e., low-end hotels affected by the policy change) and untreated hotels (i.e., high-end hotels unaffected by the policy change). The specific implementation is detailed in the Online Appendix. As illustrated in Tables A1 and A2 of the Online Appendix as well as Figure A1, we find consistent evidence that exogenous, policy-induced uncertainty led to greater changes in subsequent reference selection in the low end of the hospitality industry when compared to the control group (i.e., the high end of the hospitality sector).

5.2 | Post hoc analysis-alternative comparison purposes

As already mentioned in the theory section, comparison purposes alternative to self-assessment may underlie comparison between organizations (e.g., self-enhancement or self-improvement). Therefore, as an extension of (H3), we constructed an alternative DV to compare the profitability of all changed references to the profitability of the focal selecting hotel. Specifically, we counted the net number of added and deleted peers depending on whether these added/deleted properties showed higher/lower RevPar when compared to the focal selecting hotel. We calculated the DV in Table A7 as below. The resulting variable can take positive or negative integer values. Positive values indicate that a hotel added a net higher number of peers with a RevPar higher than its own (and vice versa).

$$\begin{aligned}\text{Net added peers with higher RevPar} = {} & \#\text{of added peers}_{(\text{higher RevPar})} - \\ & \#\text{of deleted peers}_{(\text{higher RevPar})} - \\ & \#\text{of added peers}_{(\text{lower RevPar})} + \\ & \#\text{of deleted peers}_{(\text{lower RevPar})}\end{aligned}$$

The results in Model 1 (Table A7) show that when hotels fall short of the performance of their previously selected peer group, they adjust that group toward lower-performing peers. The opposite holds true when hotels exceed the performance of their peers. These symmetrical

results speak in favor of self-assessment rather than self-enhancement rationales. This is because in self-assessment, firms should select more similar references when they perceive higher levels of uncertainty (operationalized as performance deviations). They select more similar, higher-performing hotels when their own performance is higher and vice versa. In case of self-enhancement, we should only document asymmetrical downward adjustment results.

The results explained above match the characteristics of the database, which contain reference selection information invisible to the outside. For self-enhancement, however, such visibility (or the presence of an audience) is a crucial criterion. That said, the database may still work for engaging in "internal impression management" when using reference selection to positively present oneself to a firm-internal audience. We investigated this explanation by exploiting a particularity of our sample. More specifically, we differentiated between hotels operated by third-party managers that report to a principal and owner-operated hotels. If internal impression-management mechanisms apply, we should only see a corresponding downward-adjusting reference-selection behavior for third-party managers. As illustrated in Models 2 (third-party managers) and Model 3 (owner), we did not find evidence of such differences. This further suggests that self-assessment is the dominant

TABLE 4 Overview of results.

Hypothesis (H1): A deviation from performance benchmarks is positively associated with subsequent adaptations to firms' reference group	
Supported	Any deviation from performance benchmarks, be it positive or negative, from own prior performance or from those selected into the references group, is positively associated with subsequent adaptations to the reference group.
Hypothesis (H2): A negative deviation from performance benchmarks is more strongly associated with subsequent adaptations to firms' reference group when compare to a positive deviation.	
Not supported	We do not find evidence that negative deviations from the performance of selected references (and of own prior performance) shows a higher effect on subsequent reference group adaptations when compared to positive deviations. We are confident in this result because we can replicate prior work's findings about asymmetries when considering all (but not self-selected) references (see Table A6 in the Online Appendix). This difference supports earlier arguments suggesting that such attention asymmetries emerge as information overflow (when considering all but not selected references) encourages managers to engage in selective information processing (i.e., thinking in simplified "gain/loss" categories).
Hypothesis (H3): A deviation from performance benchmarks is positively associated with subsequent adaptations to firms' reference group toward references with more similar characteristics to the focal selecting firm.	
Supported	Positive and negative deviations from the performance of those selected into the reference group are associated with the selection of references that have a higher similarity (in terms of class membership) to the focal selecting firm. One part of this effect stems from a rotation within similar firms (both, deletion and addition) while the higher similarity mostly stems from the deletion of dissimilar references (with no corresponding addition of the same). We further find that this result holds when studying the profitability levels of references selected. As firms perform higher, they equally adapt their references toward higher performing references (and vice versa). In short, hotels selected references in similar "profitability" categories (see post hoc analysis and Table A7). Our analyses further show that these effects are symmetrical. This suggests that self-assessment, but not self-enhancement is the underlying mechanism.

comparison purpose observed in our data. To conclude, these additional findings allow us to present a more robust case for our central claim that purposes (specifically self-assessment) affect reference selection. Table 4 provides an overview of our findings regarding the three hypotheses.

6 | DISCUSSION

We analyze rare, longitudinal data on firms' reference self-selection in order to derive an improved understanding about the mechanisms underlying firms' reference selection. Our main results are that: (a) managers change references in response to performance deviations or regulatory changes in their environment; (b) these changes lead to the selection of more similar references; and (c) these effects hold constant regardless of hotels entering or leaving the industry (i.e., exogenous variation in firms' competitive environment).

6.1 | Toward an endogenous, purpose-driven mechanism in reference selection

From a theoretical perspective, these results suggest that comparison needs or purposes drive firms' reference-selection decisions and that this effect is independent of changes in firms' competitive environments. We refer to this as an "endogenous" selection process because our results suggest that references are not predefined by industry affiliation (so that changes at the "industry affiliation" level drive changes in self-selected references). Instead, they are actively selected because they serve a specific decision-need. Most likely, in our application, performance deviations or regulatory changes prompt uncertainty concerning a focal hotel's relative abilities and standing, thereby triggering a selection of more similar references with the eventual aim of reducing this uncertainty. The comparison with references, therefore, reflects a "diagnosis" or "self-assessment" mechanism in which comparisons with similar references help a firm better assess its own abilities and standing in the market. Prior work conceptually mentions such self-assessment mechanisms (Greve, 1998) as well as other purposes in association with reference selection (e.g., Audia et al., 2021; Porac et al., 1999).

Our findings have numerous implications in relation to prior work directly studying reference selection (e.g., Porac et al., 1989) or studying the implications of such selection (e.g., Greve, 1998). From an empirical perspective, substantial identification problems are associated with reference data. As mentioned earlier, this has prompted researchers to contextually justify their reference comparison data or to endogenously derive this information from the data (Greve & Gaba, 2017). In relation to such observational problems, our findings both confirm and qualify prior work. We confirm extant research by providing evidence that the dominate proxies used for reference selection in this work (e.g., industry (Bromiley & Harris, 2014) or geography proxies (Henisz & Delios, 2001)) are also reflected in our reference self-selection data. However, our finding of substantial selection effects qualifies the extant research. Due to the mentioned observational problems, prior work has typically assumed that firms compare themselves to an average competitor (e.g., Iyer & Miller, 2008). In contrast, our data provides evidence of strong selection effects, which implies that firms actually pick a small subset of references or competitors (an average of only 5.2 references in our case). This raises important questions about the underlying selection logic, which brings us to the next, more theoretical implication.

While recent work highlights the importance of purpose for reference selection (e.g., Audia et al., 2021), most research still assumes that comparisons are ultimately derived from competition or industry membership (e.g., Gur & Greckhamer, 2019; Kilduff, Elfenbein, & Staw, 2010; Porac et al., 1999). Consider, for example, the debate on managerial identification of references (Porac et al., 1989; Reger & Palmer, 1996)—a debate that is particularly precise in identifying managers' idiosyncratic selection of references. According to this debate, managers develop different interpretations of industries due to, for instance, differences in perception (Reger & Huff, 1993), more or less flexible cognitive maps (Reger & Palmer, 1996), mental models (Porac et al., 1989), or political considerations (Porac et al., 1999). While this stream of research indicates that managers rely on their personal views when selecting references, this stream does still not question the exogenous origin of references. In other words, this perspective still assumes that references emerge from exogenous industry classifications, the difference being that individuals (who differ, for example, in their perceptions), interpret industries and their boundaries in different ways. These individual-level differences in the interpretation of industries, in turn, explain variations in the selected references.

Our findings point to a different mechanism. A purpose view focuses on the original need for comparisons and, instead of centering on differences between firms or industries (or individuals' views on them), derives references directly from that initial need. Therefore, the purpose view is more comprehensive and can, for example, explain out-of-industry reference selection (e.g., Srinivasan et al., 2007). Altogether, a purpose driven selection may also be termed self-serving reference selection—the selection of references to serve different firm or senior manager needs.[8]

The industry and purpose views raise substantially different questions regarding reference selection. In the industry view, critical questions pertain to aspects surrounding the characterization of industries. Given the comprehensive literature on strategic groups (McGee & Thomas, 1986), industry classifications or categories (Durand & Paolella, 2013), perceptions of rivalry (Gur & Greckhamer, 2019; Kilduff et al., 2010), and similar ones (Thomas & Venkatraman, 1988), work on this matter is extensive. In the purpose view, critical questions relate to the triggers of alternative purposes, the strength of these triggers, and the translation of such triggers into specifically selected references. While this paper represents a first attempt to answer some of these questions in relation to self-assessment needs, more work is required to develop a detailed understanding of this alternative perspective on reference selection.

6.2 | Implications to the behavioral theory of the firm (BTOF)

Reference comparisons matter for debates related to the BTOF, including debates on imitation (Posen & Martignoni, 2018), vicarious learning (Baum & Dahlin, 2007), and change and risk-taking in general (Kacperczyk et al., 2015). Within the BTOF, our focus on self-assessment is most closely associated with the debates on problemistic-search (Posen et al., 2018) and

[8]While we suggest that the term "self-serving" can be used as an alternative to "purpose-driven" reference selection, we prefer the latter for two reasons. First, prior work focused on the specific purpose of self-enhancement used the term "self serving" (Audia et al., 2022, p. 849). This implies that the term is simultaneously used for the overall mechanism (i.e., "purpose-driven selection" - which includes all four mentioned purposes) and one form of that mechanism (i.e., self-enhancement), which may cause confusion. Second, in prior work, the term "self-serving" is often associated with agency-related debating misaligned interest or information asymmetry problems between firms and managers. However, these are not major mechanisms debated in our paper, which may also cause confusion regarding the term's meaning.

performance feedback (Greve, 1998). Therefore, we focus our subsequent elaborations on these debates.

A central claim in problemistic-search is that comparisons with exogenous reference groups trigger the definition of a problem, causing subsequent search (Posen et al., 2018) or action (Greve, 1998). For the above-mentioned endogeneity reason, our results suggest that this mechanism requires updating. More specifically, comparisons with others and potential associated shortfalls are not necessarily the starting point of search or risk-taking but rather an intermediate mechanism. Instead of being a two-step process, in which a shortfall relative to exogenous references (i.e., "the trigger"—step one) is directly followed by search or action (step two, (Posen et al., 2018)), our findings suggest a more nuanced three-step process: relative uncertainty (i.e., "the trigger"—step one) is followed by the assessment of the existence of a problem (through self-assessment or comparisons with relevant references—step two), which eventually triggers search or action (step three). As we show, this intermediate step may come with additional variation (e.g., how to select references for self-assessment) and, therefore, has the potential to explain the inconsistencies observed in prior work that only associated shortfalls with subsequent actions (e.g., Boyle & Shapira, 2012; Iyer & Miller, 2008).[9] As such, our findings suggest that these inconsistencies relate to a deeper issue than the problems associated with imprecise reference group identification (Greve & Gaba, 2017). Instead, reference group comparisons may not be the starting point for search but rather describe an intermediary self-assessment process, which may or may not result in subsequent search.

Our finding on H3 is not in line with problemistic-search arguments, as a performance shortfall is associated with more "local" (or similar) but not dissimilar reference selection. Upon reflection, however, we believe that our findings suit the above described adapted logic for reference comparisons in the problemistic-search process. This is because problemistic-search is context or problem oriented. This implies that when a performance shortfall signals the existence of a problem (e.g., concerning a certain product or market), the subsequent search is directly related to that problem (e.g., distant search for product variation). In other words, firms engage in risk-taking to find a solution to the uncovered problem. In contrast, our self-assessment mechanism describes a preceding or intermediate mechanism aimed at evaluating the existence of a problem in the first place. It does not focus on "risk-taking" or "distant search" but rather on diagnosis. As proposed in H3, selecting more similar (but not "dis-similar" or "distant") references is more suitable for such diagnosis.

Another contribution relates to the lack of a finding regarding asymmetries in loss- or gain-framed uncertainty (H2). We deem this result particularly reliable, as we were able to reproduce asymmetrical effects when replicating prior work's focus on industry-wide (i.e., not self-selected) reference-group comparisons. In contrast to the extant work, our findings imply that the respective selection effects can be attributed to attention effects or managerial information overload (Van Knippenberg et al., 2015). In short, when reference pre-selection alleviates overload effects, we no longer observe asymmetrical attention effects. Instead, managers consider the entire performance deviation continuum with equal attention.

Altogether, the above findings provide more detail on the intermediate mechanism in problemistic-search and performance feedback, theories that have often been criticized for resembling a "black box" (Posen et al., 2018). Our work shows that when there is variation in the self-assessment need, firms adjust their references or comparison groups (H1) toward more

[9]In recent work, these frequent inconsistencies have even induced scholars abandoning investigating the "social aspiration" mechanism altogether (Ref & Shapira, 2017).

similar members (H3), thereby enabling the selecting firms to develop a more detailed understanding of their relative abilities. In this regard, we document an important intermediate mechanism in problemistic-search and performance feedback.

6.3 | Implications for SCT

Our study entails the direct application of SCT arguments to an organizational setting. We conclude that SCT mechanisms are applicable to both private and professional individuals (or economic actors) who engage in reference selection *on behalf of* their organizations. Thereby, our paper is among the first to provide evidence of the existence of self-assessment motives in firms' reference selection. Prior work has found support for self-enhancement motives (Audia et al., 2021; Porac et al., 1999), which may reflect that work's reliance on publicly visible reference-selection information. In contrast, we show that organizations use reference groups not only as a means of signaling or impression management but also as a way to derive an accurate view of themselves. In this regard, our research is comparable to the work on (exogenous) industry categorizations. Industry categorizations are not an end in themselves but serve a need (i.e., the need to structure, understand, or simplify complex industries and environments), as is also evident in the debate on goal-directed categorizations (Durand & Paolella, 2013). This implies that the main difference between our perspective and the traditional industry or categorization perspective goes back to differences in needs—the general need to better understand a complex competitive environment (as in the industry perspective) and the specific need to assess one's own capabilities (as in our self-assessment-derived perspective).

6.4 | Managerial implications

Our results indicate that when considering reference groups in relation to their actions and decisions, managers should be more aware and critical of the origins of those groups. Such groups play an important role in many actions and decisions (e.g., Iyer & Miller, 2008), and our results suggest that they are formed to serve certain needs. This implies that reference groups should not be considered separate from context or the need that originally created them. In simple terms, when comparisons with a reference group affect managers' decisions, we encourage managers to question the origin and actual relevance of that group for the focal matter or decision.

6.5 | Limitations and future research

Our findings call for research that investigates the purpose-driven reference-selection mechanism in more detail. Most importantly, such work should focus on identifying and isolating other purposes (e.g., self-improvement, self-verification), investigating their triggers, and studying how they translate into actual reference selection. As we mentioned above, this might be challenging given the difficulties associated with identifying purpose. Nevertheless, we encourage future work along these lines, for example, conducting field or lab experiments or studying alternative determinants to purpose (Kulik & Ambrose, 1992). The ultimate objective of all such

research should be to develop a more precise understanding of the important phenomenon of reference selection.

A second limitation pertains to our empirical context of the hospitality industry. This industry encompasses a variety of characteristics likely to affect reference-selection behavior and the identification of purpose. For instance, this industry is asset intensive and rather stable (i.e., hotel properties are typically locked into a location and a quality category), well structured, and transparent (in terms of, e.g., class affiliations and amenities). This is not the case in other industries, such as technology-oriented industries, which are typically more challenging to structure, observe, and compare. Such factors may lead to reference-selection behavior and purpose activation different from those observed in our sample. In short, we question the ecological validity of our work and encourage future work that examines the applicability of our findings to other industries.

Third, we encourage examinations of the consequences of the reference-selection mechanism observed in this paper. As reference selection is endogenous, we question whether the updated logic behind problemistic-search or performance feedback yields similar predictions regarding firms' subsequent actions or decisions. We mentioned that our findings can explain the many inconsistencies observed in these theories. Future research that adopts our updated mechanism, should review this possibility in relation to the actual search outcome. This also relates to the fact that we find no asymmetrical effects. Specifically, as there are no asymmetrical effects in our intermediate process (H2), one might wonder whether this translates into no asymmetrical effects on subsequent firm actions (e.g., Greve, 1998).

A final area for future research relates to agency problems. Prior work provided evidence of such problems. For example, senior managers act according to their own comparison purposes and not according to the comparison purposes of their firms (Porac et al., 1999). In this regard, we highlight two possible mechanisms. First, as mentioned above, senior managers may put their own interests above those of the firm. Second, imprecision may arise as managers select references they *believe* will (but in fact do not) best serve their organization's comparison purpose. Future research should more carefully aim to separate comparison purposes at the managerial and firm levels and, thereby, improve our overall understanding of SCT at an organizational level. This also relates to the micro-foundations of such mechanisms.

7 | CONCLUSION

Reference groups are pivotal in organizations' actions and decision-making. Our findings that these reference groups emerge from purpose and not only from industry affiliation are important because they change our view on how comparisons with reference groups manifest in organizational actions. According to our findings, these groups may not have an independent, exogenous effect but rather work as an endogenous intermediate mechanism. A more precise understanding of this intermediate mechanism also provides a more detailed theoretical understanding of subsequent actions and decisions. Although this paper represents a first step toward an updated understanding, much more work is required if we are to fully comprehend the phenomenon of reference selection, including its determinants and consequences

ACKNOWLEDGEMENTS

This work would have not been possible without the help of many. In terms of data, I am grateful to the STR SHARE Center, especially toward Duane Vinson and Stephen Hood.

Furthermore, I am grateful to Achim Schmitt and the EHL Hospitality Business School. I thank the Associate Editor, Andrew Shipilov, and two anonymous reviewers for their valuable feedback and guidance throughout the review process. I also thank Julia Bodner, Kristian Dahlin, J. P. Eggers, Nicolai Foss, Henrich Greve, Paul Huenermund, Thomas Keil, Aleksey Korniychuk, Xu Li, Dirk Martignoni, Ali Mohammadi, and Hart Posen for their feedback on earlier versions of this manuscript. This research received support from the Swiss National Science Foundation (SNSF grant # 204670).

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from Smith Travel Research (STR Inc.). Restrictions apply to the availability of these data, which were used under license for this study. Data are available from the authors with the permission of Smith Travel Research.

ORCID

Johannes Luger https://orcid.org/0000-0001-5877-5448

REFERENCES

Audia, P. G., Rousseau, H. E., & Brion, S. (2021). CEO power and nonconforming reference group selection. *Organization Science*, *33*(2), 831–853.

Baum, J. A., & Dahlin, K. B. (2007). Aspiration performance and railroads' patterns of learning from train wrecks and crashes. *Organization Science*, *18*(3), 368–385.

Baum, J. A., & Lant, T. K. (2003). Hits and misses: Managers'(mis) categorization of competitors in the Manhattan hotel industry. *Advances in Strategic Management*, *20*, 119–158.

Baum, J. A., Li, S. X., & Usher, J. M. (2000). Making the next move: How experiential and vicarious learning shape the locations of chains' acquisitions. *Administrative Science Quarterly*, *45*(4), 766–801.

Baumann, O., Eggers, J., & Stieglitz, N. (2019). Colleagues and competitors: How internal social comparisons shape organizational search and adaptation. *Administrative Science Quarterly*, *64*(2), 275–309.

Boyd, B. K., & Fulk, J. (1996). Executive scanning and perceived uncertainty: A multidimensional model. *Journal of Management*, *22*(1), 1–21.

Boyle, E., & Shapira, Z. (2012). The liability of leading: Battling aspiration and survival goals in the jeopardy! Tournament of champions. *Organization Science*, *23*(4), 1100–1113.

Bromiley, P., & Harris, J. D. (2014). A comparison of alternative measures of organizational aspirations. *Strategic Management Journal*, *35*(3), 338–357.

Canina, L., Enz, C. A., & Harrison, J. S. (2005). Agglomeration efects and strategic orientations: Evidence from the us lodging industry. *Academy of Management Journal*, *48*(4), 565–581.

Chen, M.-J. (1996). Competitor analysis and interfirm rivalry: Toward a theoretical integration. *Academy of Management Review*, *21*(1), 100–134.

d'Auguste, J. (2019). City of New York v. airbnb, Inc. Retrieved from https://casetext.com/case/city-of-ny-v-airbnb-inc

Dawes, J. G. (2014). Patterns in competitive structure among retail financial services brands. *European Journal of Marketing.*, *48*, 924–942.

DeSarbo, W. S., Grewal, R., & Wind, J. (2006). Who competes with whom? A demandbased perspective for identifying and representing asymmetric competition. *Strategic Management Journal*, *27*(2), 101–129.

DiMaggio, P. J., & Powell, W. W. (1983). The iron cage revisited: Instutitional isomorphism and collective rationality in organizational fields. *Americal Sociological Review*, *48*, 147–160.

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.

Festinger, L. (1954). A theory of social comparison processes. *Human Relations*, *7*(2), 117–140.

Gaba, V., & Joseph, J. (2013). Corporate structure and performance feedback: Aspirations and adaptation in m-form firms. *Organization Science*, *24*(4), 1102–1119.

Goh, S., & Richards, G. (1997). Benchmarking the learning capability of organizations. *European Management Journal*, *15*(5), 575–583.

Greve, H. R. (1998). Performance, aspirations, and risky organizational change. *Administrative Science Quarterly*, *43*(1), 58–86.

Greve, H. R. (2008). A behavioral theory of firm growth: Sequential attention to size and performance goals. *Academy of Management Journal*, *51*(3), 476–494.

Greve, H. R. (2011). Positional rigidity: Low performance and resource acquisition in large and small firms. *Strategic Management Journal*, *32*(1), 103–114.

Greve, H. R., & Gaba, V. (2017). *Performance feedback in organizations and groups: Common themes*. New York, NY: Oxford University Press.

Gur, F. A., & Greckhamer, T. (2019). Know thy enemy: A review and agenda for research on competitor identification. *Journal of Management*, *45*(5), 2072–2100.

Hatten, K. J., & Schendel, D. E. (1977). Heterogeneity within an industry: Firm conduct in the US brewing industry, 1952-71. *Journal of Industrial Economics*, *26*, 97–113.

Henisz, W. J., & Delios, A. (2001). Uncertainty, imitation, and plant location: Japanese multinational corporations, 1990-1996. *Administrative Science Quarterly*, *46*(3), 443–475.

Iyer, D. N., & Miller, K. D. (2008). Performance feedback, slack, and the timing of acquisitions. *Academy of Management Journal*, *51*(4), 808–822.

Joseph, J., & Gaba, V. (2015). The fog of feedback: Ambiguity and firm responses to multiple aspiration levels. *Strategic Management Journal*, *36*(13), 1960–1978.

Kacperczyk, A., Beckman, C. M., & Moliterno, T. P. (2015). Disentangling risk and change internal and external social comparison in the mutual fund industry. *Administrative Science Quarterly*, *60*(2), 228–262.

Kilduff, G. J., Elfenbein, H. A., & Staw, B. M. (2010). The psychology of rivalry: A relationally dependent analysis of competition. *Academy of Management Journal*, *53*(5), 943–969.

Kim, K. H., & Tsai, W. (2012). Social comparison among competing firms. *Strategic Management Journal*, *33*(2), 115–136.

Kim, L., & Lim, Y. (1988). Environment, generic strategies, and performance in a rapidly developing country: A taxonomic approach. *Academy of Management Journal*, *31*(4), 802–827.

King, G., & Zeng, L. (2001). Logistic regression in rare events data. *Political Analysis*, *9*(2), 137–163.

Kulik, C. T., & Ambrose, M. L. (1992). Personal and situational determinants of referent choice. *Academy of Management Review*, *17*(2), 212–237.

Lavie, N. (1995). Perceptual load as a necessary condition for selective attention. *Journal of Experimental Psychology: Human Perception and Performance*, *21*(3), 451–468.

Li, J., Qian, C., & Yao, F. K. (2015). Confidence in learning: Inter-and intraorganizational learning in foreign market entry decisions. *Strategic Management Journal*, *36*(6), 918–929.

March, J. G., & Simon, H. A. (1958). *Organizations*. New York: John Wiley and Sons.

McGee, J., & Thomas, H. (1986). Strategic groups: Theory, research and taxonomy. *Strategic Management Journal*, *7*(2), 141–160.

Nieuwland, S., & Van Melik, R. (2020). Regulating airbnb: How cities deal with perceived negative externalities of short-term rentals. *Current Issues in Tourism*, *23*(7), 811–825.

Porac, J. F., Thomas, H., & Baden-Fuller, C. (1989). Competitive groups as cognitive communities: The case of scottish knitwear manufacturers. *Journal of Management Studies*, *26*(4), 397–416.

Porac, J. F., Wade, J. B., & Pollock, T. G. (1999). Industry categories and the politics of the comparable firm in CEO compensation. *Administrative Science Quarterly*, *44*(1), 112–144.

Porter, M. (1980). *Competitive strategy*. New York: Free Press.

Posen, H. E., Keil, T., Kim, S., & Meissner, F. D. (2018). Renewing research on problemistic search|a review and research agenda. *Academy of Management Annals*, *12*(1), 208–251.

Posen, H. E., & Martignoni, D. (2018). Revisiting the imitation assumption: Why imitation may increase, rather than decrease, performance heterogeneity. *Strategic Management Journal*, *39*(5), 1350–1369.

Posen, H. E., Ross, J.-M., Wu, X., Benigni, S., & Cao, Z. (2022). Reconceptualizing imitation: Implications for dynamic capabilities, innovation, and competitive advantage. *Academy of Management Annals*. Forthcoming.

Ref, O., & Shapira, Z. (2017). Entering new markets: The effect of performance feedback near aspiration and well below and above it. *Strategic Management Journal*, *38*(7), 1416–1434.

Reger, R. K., & Huff, A. S. (1993). Strategic groups: A cognitive perspective. *Strategic Management Journal*, *14*(2), 103–123.

Reger, R. K., & Palmer, T. B. (1996). Managerial categorization of competitors: Using old maps to navigate new environments. *Organization Science*, *7*(1), 22–39.

Renzino, A. (2018). *Tennessee short term rental unit act*. Retrieved from https://nastra.org/news/tennessee-short-term-rental-unit-act

Sedikides, C. (1993). Assessment, enhancement, and verification determinants of the selfevaluation process. *Journal of Personality and Social Psychology*, *65*(2), 317–338.

Sedikides, C., & Strube, M. J. (1997). Self-evaluation: To thine own self be good, to thine own self be sure, to thine own self be true, and to thine own self be better. In P. Z. Mark (Ed.), *Advances in experimental social psychology* (Vol. 29, pp. 209–269). San Diego, CA: Academic Press.

Srinivasan, R., Haunschild, P., & Grewal, R. (2007). Vicarious learning in new product introductions in the early years of a converging market. *Management Science*, *53*(1), 16–28.

Tetlock, P. E., & Manstead, A. S. (1985). Impression management versus intrapsychic explanations in social psychology: A useful dichotomy? *Psychological Review*, *92*(1), 59–77.

Thomas, H., & Pollock, T. (1999). From I-O economics' S-C-P paradigm through strategic groups to competence-based competition: Reections on the puzzle of competitive strategy. *British Journal of Management*, *10*(2), 127–140.

Thomas, H., & Venkatraman, N. (1988). Research on strategic groups: Progress and prognosis [1]. *Journal of Management Studies*, *25*(6), 537–555.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty—Heuristics and biases. *Science*, *185*(4157), 1124–1131.

Van Knippenberg, D., Dahlander, L., Haas, M. R., & George, G. (2015). Information, attention, and decision making. In *Information, attention, and decision making* (Vol. 58). Manor, NY: Academy of Management Briarcliff.

Yu, F. F. (2008). Analyst coverage and earnings management. *Journal of Financial Economics*, *88*(2), 245–271.

Zervas, G., Proserpio, D., & Byers, J. W. (2017). The rise of the sharing economy: Estimating the impact of airbnb on the hotel industry. *Journal of Marketing Research*, *54*(5), 687–705.

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: Luger, J. (2023). Who depends on why: Toward an endogenous, purpose-driven mechanism in organizations' reference selection. *Strategic Management Journal*, *44*(8), 2035–2059. https://doi.org/10.1002/smj.3486