

# The myth of the flat start-up: Reconsidering the organizational structure of start-ups

Saerom (Ronnie) Lee 

The Wharton School, University of Pennsylvania, Philadelphia, Pennsylvania, USA

## Correspondence

Saerom (Ronnie) Lee, The Wharton School, University of Pennsylvania, Philadelphia, PA 19104, USA.  
Email: saeroms@wharton.upenn.edu

## Funding information

Kwanjeong Educational Foundation

**Research summary:** There has been an ongoing debate over whether start-ups should be “flat” with minimal hierarchical layers. To reconcile this debate, this article distinguishes between creative and commercial success (i.e., novelty vs. profitability), and examines how these outcomes are variously influenced by a start-up’s hierarchy. This study suggests that while a flatter hierarchy can improve ideation and creative success, it can result in haphazard execution and commercial failure by overwhelming managers with the burden of direction and causing subordinates to drift into power struggles and aimless idea explorations. I find empirical support for this trade-off using a large sample of game development start-ups. These findings offer one resolution to the debate by sorting out the conditions under which hierarchy can be conducive or detrimental to start-ups.

**Managerial summary:** Academics, management gurus, and popular media outlets have argued that “authoritarian,” tall hierarchies are outmoded and will be supplanted by “egalitarian,” flat structures. In recent years, this argument has been largely substantiated by a few “successful” flat start-ups, such as Valve, Zappos, Github, Medium, and Buffer. As these nascent firms constantly garner much attention for their egalitarian ideal—which itself is a signal of their rarity—the myth that start-ups should be flat (often referred to as “flat organization,” “holacracy,” or “boss-less firm”) has

become widespread among entrepreneurs. My study cautions against this myth, suggesting that adding a few hierarchical levels of managers can substantially help start-ups achieve commercial success and survive in their hostile environments, albeit at a potentially marginal cost of creativity.

#### KEY WORDS

entrepreneurship, hierarchy, organizational design, organizational structure, start-up

## 1 | INTRODUCTION

The dirtiest word in Silicon Valley is bureaucracy. ... [L]ayers of bureaucracy...are seen as the enemy of speed and efficiency. (Miller, 2014)

Entrepreneurs tend to regard the organizational structure of hierarchy as “a bureaucratic threat to their entrepreneurial souls” (Gulati & DeSantola, 2016) and are thus reluctant to impose hierarchical layers of managers (Davila, Foster, & Jia, 2010; Puranam, 2014). Some have gone on to successfully launch a “flat” start-up with minimal layers (e.g., Valve, Github, and Zappos), thus gaining much popular acclaim and scholarly interest (Burton et al., 2017; Lee & Edmondson, 2017, p. 38; Puranam & Håkonsson, 2015). This distaste for hierarchy among entrepreneurs is perhaps not surprising if we consider two assertions on new entrepreneurial ventures in the classic literature on organization design. One such assertion is that such small and undifferentiated firms do not require layers of managers for coordination (Blau, 1970; Burton & Obel, 2004, pp. 389–391; Mintzberg, 1979, pp. 310–311). The other is that as these small nascent firms typically operate in dynamic and hostile industries (Santos & Eisenhardt, 2009), they should be flat to streamline their decision-making process and flexibly adapt to their environments (Burns & Stalker, 1961; Covin & Slevin, 1989; Garicano & Posner, 2005).

However, recent qualitative studies have started to raise questions about this popular narrative that start-ups should be flat. These studies point out that, albeit smaller and less differentiated than their mature counterparts, start-ups face coordination problems because these nascent firms not only lack formalized routines (i.e., standard operating procedures; Sine, Mitsuhashi, & Kirsch, 2006; Stinchcombe, 1965, pp. 148–150) and informal coordination devices (e.g., shared culture or norms; McEvily, Soda, & Tortoriello, 2014; Meier, Stephenson, & Perkowski, 2019), but also have insufficient resources to designate integrating committees (Baker & Nelson, 2005). As start-ups have no other coordination devices that can substitute for hierarchy, Sutton and Rao (2014, p. 107) suggest that “even [these] small organizations can’t function without hierarchies.”

To date, this theoretical debate over whether start-ups should be flat has not been empirically addressed, partly due to the challenges in gathering a large sample dataset of start-ups and measuring their hierarchy (Burton, Colombo, Rossi-Lamastra, & Wasserman, 2019; Keum & See, 2017). Thus, as Burton et al. (2019, p. 2) remark, “the empirical evidence on the

antecedents and consequences of organizational design choices in entrepreneurial ventures is both limited and mixed.” Hence, much remains to be learned about how the existing theories on organizational design apply to these nascent firms (Dushnitsky & Matusik, 2019; van de Ven, Ganco, & Hinings, 2013).

To address this gap and reconcile the debate on whether start-ups should be flat, I bring to the forefront two conditions that start-ups face but that the popular narrative has largely neglected. The first condition is that these nascent firms lack alternative coordination devices (Sine et al., 2006; Stinchcombe, 1965, pp. 148–150). The second is that although start-ups strive for both creative success and commercial success (i.e., novelty vs. profitability; Kaul, 2013; Shane & Venkataraman, 2000), they often realize one at the expense of the other (Kaplan & Vakili, 2015; Uzzi & Spiro, 2005, p. 468). This is because these two outcomes are shaped by distinct processes, thus requiring different structural configurations (Csaszar, 2013; Csaszar & Eggers, 2013; Keum & See, 2017). Building on these two conditions, my study suggests that a flatter hierarchy, on the one hand, can improve ideation and creative success by stimulating cross-fertilization of disparate ideas and reducing premature filtering of ideas by managers. However, this benefit can come at the cost of execution and commercial success. I theorize that, without other systematic ways to coordinate, a flatter hierarchy can overwhelm their managers with the burden of direction and conflict resolution, cause subordinates to drift into power struggles and aimless exploration of ideas, and result in major employee turnovers. Put differently, adding hierarchical levels can compensate for the lack of alternative coordination devices and thus be beneficial for start-ups in achieving commercial success and ultimately in their survival, albeit at the expense of creative success.

This study finds empirical support for this performance trade-off by analyzing start-ups in the video game industry. As one of the fastest-growing venues for entrepreneurial activities (Bies, 2017), this industry provides a unique context that allows me to overcome the aforementioned empirical challenges in data collection and measurement. That is, for a large number of start-ups, detailed data is available on their employees (e.g., names, job titles), critics' product review ratings, and sales figures. I use this rich dataset to compute a well-established measure of hierarchy (i.e., the number of hierarchical levels derived from text analysis of job titles) and two distinct measures of performance (i.e., creative and commercial success).

My study contributes to the literature on entrepreneurship and organizational design by offering one resolution to the debate over whether start-ups should have flat hierarchies. In doing so, this study first qualifies the conventional wisdom that start-ups should be (and are thus) flat by providing rare empirical evidence of the heterogeneity and trend in the hierarchy among start-ups. This paper then shows that adding a hierarchical level may entail a trade-off between creative and commercial success—that is, it may increase commercial success at the expense of creative success. This trade-off suggests that hierarchy is not universally good or bad for start-ups, but rather its efficacy depends on the type of performance they pursue. Finally, in support of the theorized mechanism, the results demonstrate that the benefit of adding a hierarchical level decreases when start-ups have informal coordination devices (i.e., social capital that employees develop through prior collaborations). Overall, by sorting out the conditions under which hierarchy can be conducive or detrimental to start-ups, my study offers a contingent perspective that helps reconcile the debate on the organizational structure of start-ups.

## 2 | THEORY AND HYPOTHESES

This section begins by drawing upon the literature on entrepreneurship to define a start-up and to explain why these nascent firms face coordination problems in achieving creative and commercial success. I then build on the literature on organizational design to outline how coordination can, in principle, be achieved through hierarchy and to discuss how this vertical division of tasks may entail a trade-off between creative and commercial success in the context of start-ups. This theoretical discussion is supplemented by qualitative observations of start-ups, as the goal here is not only to consider whether a start-up's performance increases or decreases with its hierarchy, but also—and perhaps more importantly—to articulate the mechanisms specific to start-ups.

### 2.1 | Start-ups and their two performance measures

In line with prior studies on entrepreneurship (e.g., Carland, Hoy, Boulton, & Carland, 1984; Decker, Haltiwanger, Jarmin, & Miranda, 2014; Wasserman, 2012), this study focuses on “high-growth start-ups”—that is, newly established firms (of any size) that aim to introduce innovations and achieve organizational growth.<sup>1</sup> To survive and grow in their hostile environments, these nascent firms need to not only flexibly generate novel ideas to differentiate their products, but also efficiently execute those ideas to make profits (Kaul, 2013; Shane & Venkataraman, 2000). Simply put, start-ups actively strive for both creative success and commercial success.<sup>2</sup>

However, in many cases, they only realize one at the expense of the other (Kaplan & Vakili, 2015; Uzzi & Spiro, 2005, p. 468). This trade-off between these performance outcomes is perhaps best reflected in the following comment by an entrepreneur in the video game industry:

Ideally, every effort brings both artistic [i.e., creative] and commercial success. Sometimes, you only get one or the other. (Ruggiero, 2017; comment in brackets added)

This trade-off occurs partly because achieving both creative and commercial success presents a daunting challenge for start-ups. To realize creative success, a start-up must, by definition, set its product apart from existing ones by generating a variety of novel ideas (Fleming et al., 2007). However, a purely novel product incorporating all these ideas may not be commercially successful because it is so eccentric or incoherent that consumers find it difficult to understand, appreciate, and adopt (Hargadon & Douglas, 2001). To gain acceptance from consumers and attain commercial success, the start-up thus needs to prevent aimless exploration of novel

<sup>1</sup>This article distinguishes these nascent, growth-oriented firms from small businesses (of any age) operating with a minimum number of employees (in colloquial terms, “mom-and-pop stores”). This is because these two represent different types of entrepreneurship: whereas the former is geared toward scaling up its business over time, the latter typically has no desire to grow its business (Carland et al., 1984; Wasserman, 2012, p. 6). As the latter remains considerably small throughout its lifetime and thus does not necessitate hierarchical structure (Blau, 1970; Mintzberg, 1979), this study focuses on the former.

<sup>2</sup>These two performance outcomes apply not just to creative industries, but broadly to various other contexts (see, e.g., Bangle, 2001; Eggers & Kaul, 2018; Fleming, Mingo, & Chen, 2007; Kaplan & Vakili, 2015).

ideas (Bangle, 2001; Gulati & DeSantola, 2016) and “cut features *diplomatically* when it is in the best interest” (Grossman, 2003, p. 109; emphasis added) without generating severe conflicts among employees who may be emotionally attached to their ideas. Should the start-up *tactlessly* screen out ideas, it may lose its distinctiveness as well as its employees, and ultimately fail.

## 2.2 | Coordination problems in achieving the two performance measures

Therein lies a central problem for start-ups in achieving creative and commercial success: coordination (Burton & Obel, 2004, p. 84; Puranam, 2018, p. 67). That is, because start-ups struggle to mobilize new resources (Baker & Nelson, 2005), they must integrate their employees' creative efforts and resolve conflicts among their employees to make the most out of their limited resources.

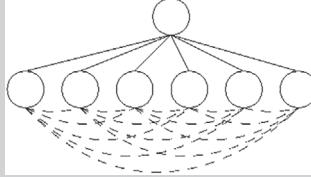
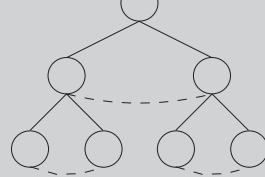
Although start-ups are not as large and horizontally differentiated as to their mature counterparts, they find it particularly challenging to integrate efforts and resolve conflicts because these nascent firms are yet to develop formal coordination devices (e.g., standard operating procedures; Sine et al., 2006; Stinchcombe, 1965, pp. 148–150). Moreover, start-ups typically lack informal devices, such as a common language, shared culture, and norms (McEvily et al., 2014; Meier et al., 2019), because they mostly consist of inexperienced employees who are new to the industry and, except for the co-founders, have rarely worked together before (Grossman, 2003, p. 307; Wasserman, 2012, pp. 229–232). Though it may seem feasible for these relatively smaller firms to coordinate through informal communication (Burns & Stalker, 1961, pp. 121–122; Simon, 1947/1997, pp. 213–215), such informal coordination becomes cumbersome and breaks down even in a small group with less than 10 members (Hackman, 2002, pp. 116–122; Steiner, 1972, p. 96). Given this lack of other systematic ways to coordinate employees, hierarchy can play an important role for start-ups.

## 2.3 | Hierarchy as a coordination device

In principle, hierarchy achieves coordination by formally dividing decision-making tasks into smaller components and vertically delegating those components among managers at each hierarchical level (Puranam, 2018, pp. 106–126; Simon, 1947/1997, p. 7).<sup>3</sup> This vertical division of tasks narrows each manager's span of control and restricts her subordinates' cross-relationships (Blau & Scott, 1962/2003, p. 139; Burton & Obel, 2004, pp. 169–170). Here, the span of control refers to the number of subordinates that directly report to each manager through the vertical chain of command (Urwick, 1956), whereas a cross-relationship refers to a lateral channel for

<sup>3</sup>In line with the entrepreneurs' interpretation, the term “hierarchy” in this study specifically refers to a formalized dimension of organizational structure represented by the number of hierarchical levels (Burton & Obel, 2004, pp. 75–77). This study distinguishes hierarchy from “informal hierarchy” (i.e., an unofficial power/status ranking; Anderson & Brown, 2010, p. 3). This is because the former formally delineates authorities that employees officially accept when joining the firm (March & Simon, 1958/1993, p. 110). In contrast, the latter emerges over time through social interactions among employees, thus accompanying power struggle, politicking, and conflicts (Kilduff, Willer, & Anderson, 2016). To make this distinction clear, this study refers to the latter as “informal structure” (McEvily et al., 2014, p. 300).

**TABLE 1** An illustration of how adding a hierarchical level decreases each manager's span of control and her subordinates' cross-relationships

		Hierarchy
	(a) Flatter	(b) Taller
		
Number of employees	7	7
Number of hierarchical levels	2	3
Average span of control	6	2
Average cross relationships	15	1

*Note:* In the figures, a circle depicts an employee. A solid line between the circles represents a vertical chain of command, whereas a dashed line represents a cross-relationship.

informal communication among the subordinates within the same span of control (Graicunas, 1937, pp. 192–193).

To illustrate, consider a hypothetical start-up with seven employees (Table 1). If it adopts two hierarchical levels (the left figure), its manager has a span of control of six and her subordinates share 15 cross-relationships. In turn, adding a hierarchical level (the right figure) decreases the average span of control to two and the average number of cross-relationships is limited to one.<sup>4</sup> More formally, assuming the unity of command (Fayol, 1949/2013, pp. 24–25) and a homogeneous span of control throughout the firm, the average span of control can be derived by solving for  $S$  in the formula  $N = \frac{S^L - 1}{S - 1}$ , where  $N$  is the number of employees and  $L$  is the number of hierarchical levels (Csaszar, 2021).<sup>5</sup> In turn, the number of cross-relationships can be calculated by  $\frac{S(S-1)}{2}$  (Graicunas, 1937, pp. 192–193). Albeit approximations, these equations will suffice to establish that given the number of employees, adding a hierarchical level decreases each manager's span of control and her subordinates' cross-relationships. This decrease reduces each manager's burden of direction and conflict resolution, thus facilitating coordination (Puranam, 2018, p. 113).

In what follows, I hypothesize specific ways in which adding a hierarchical level, by restricting each manager's span of control and her subordinates' cross-relationships, may influence creative and commercial success in the context of start-ups.

<sup>4</sup>In principle, two same-sized firms with the same number of hierarchical levels can have a different average span of control. This is possible when the “unity of command” (Gulick & Urwick, 1937, p. 9) is violated, and thus a subordinate employee reports to multiple upper-level employees. Although this violation may occur in large firms with a matrix structure (Burton & Obel, 2004, pp. 60–66), this type of complex structure is rare among start-ups.

<sup>5</sup>Blau and Scott (2003, pp. 168–169) offer a coarser approximation  $S = \sqrt[L-1]{N}$ , which also indicates that adding a hierarchical level decreases the average span of control.

### 2.3.1 | Hierarchy and creative success

A start-up's creative success may benefit from a flatter hierarchy. By specifying fewer hierarchical levels, a flatter hierarchy broadens the cross-relationships among employees and provides them more autonomy (Ghiselli & Siegel, 1972). Because these employees also lack formalized routines to abide by, they can more freely exchange and self-organize around ideas (Lee & Edmondson, 2017, p. 37; Puranam & Håkonsson, 2015). As Saxenian (1996, p. 76) illustrates in her classic book on start-ups in Silicon Valley and Boston's Route 128:

The elimination of direct hierarchical lines of authority and the creation of autonomous groups stimulated informal communication and generated an immense reservoir of new technological ideas.

As this quotation suggests, the autonomy of self-organizing can stimulate the cross-fertilization of disparate ideas (Fleming et al., 2007), thereby promoting the start-up's ideation and creative success.

Adding a hierarchical level confines these cross-relationships (Puranam, 2018, p. 131) and requires ideas to be processed through the vertical chain of command (Fayol, 1949/2013, pp. 34–36). One drawback of this vertical information processing is that as ideas are imperfectly communicated and selectively filtered by more hierarchical levels, novel ones are more likely to be distorted and fall through the cracks (Carzo & Yanouzas, 1969; Csaszar, 2012; Reitzig & Sorenson, 2013). This process of premature filtering is illustrated by a game developer:

Ideas often get shut down *prematurely* just because some people with the power to veto an idea simply don't understand it. (Schreier, 2015; emphasis added)

This premature filtering can prevent its employees from generating and sharing novel ideas, as these employees self-censor novel ideas that might be deemed too risky and share "safer" ones that seem to conform with the ideas of their supervisors (Keum & See, 2017; Reitzig & Maciejovsky, 2015; Tost, Gino, & Larrick, 2013). Thus, by reducing the number of novel ideas generated, shared, and selected, adding a hierarchical level can impede the start-up's ideation and creative success.

In sum, the discussion above leads to the following hypothesis:

**Hypothesis (H1).** *The number of hierarchical levels in a start-up is negatively associated with its subsequent creative success.*

### 2.3.2 | Hierarchy and commercial success

Unlike its creative success, a start-up's commercial success can be hampered by a flatter hierarchy. That is, by dramatically increasing each manager's span of control, a flatter hierarchy imposes a heavy cognitive burden on the manager to be responsible for more subordinates (Graicunas, 1937). As these managers are overwhelmed by the burden of direction and conflict resolution (Burton & Obel, 2004, p. 169; Puranam, 2018, p. 113), they tend to disengage from their supervisory roles (Blau & Scott, 1962/2003, p. 238), causing their subordinates to lose direction and drift into aimless exploration of ideas (Gulati & DeSantola, 2016; Sine et al., 2006,

p. 123; Urwick, 1956, p. 43). This is portrayed by a former employee of Valve, a game company renowned for its flat hierarchy:

Without managers to keep them in line, ... many jockeyed for projects that weren't suited to their skills. (Maier, 2013)

Because these subordinates lack not only their superiors' supervision but also formalized routines to follow (Stinchcombe, 1965, pp. 148–150), they may alternatively gravitate toward an informal structure (i.e., unofficial rankings based on status and respect) for directions (Freeman, 1972; Gruenfeld & Tiedens, 2010, p. 1262). As one entrepreneur reflects:

We began to realize that by building a company with a flat org. structure, ... we had centralized all the decision-making, and we were relying on a *secret implicit structure* to make progress. (Savage, 2015; emphasis added)

Because this informal structure is not clearly defined or officially agreed upon, subordinates may disagree with its directions and struggle with politicking and dysfunctional conflicts until this structure stabilizes (Kilduff et al., 2016). Even if it stabilizes, it is unlikely to be functional because the winner of such power contests tends to be one based on attractiveness or dominance, rather than competence (Tarakci, Greer, & Groenen, 2016). As the former Valve employee recalls:

There are popular kids that have acquired power in the company, then there are the troublemakers who actually want to make a difference. ... productivity suffered and communication broke down as well. (Maier, 2013)

These power struggles and conflicts can prevent subordinates from focusing on substantive issues (Greer & van Kleef, 2010) and result in the “lowest-common-denominator compromises” (Wasserman, 2012, p. 133). When these conflicts spin out of control, they can even lead to major employee turnovers. For instance, CloudFlare, a start-up once proclaimed to be flat with no hierarchical levels, lost one-fifth of its employees, who cited “the lack of a clear mid-level reporting structure” (Gulati & DeSantola, 2016). By letting employees drift into aimless exploration and grapple with dysfunctional conflicts, a flatter hierarchy can result in haphazard execution and commercial failure.

Adding a hierarchical level can forestall such aimless exploration and dysfunctional conflicts by decreasing each manager's span of control (Carzo & Yanouzas, 1969, p. 189). With a smaller span of control, each manager can concentrate on providing clear guidance and can exercise her formal authority in prioritizing ideas and making legitimate decisions (Blau, 1970; Urwick, 1956). This clarity in guidance and decision-making authorities eliminates uncertainty in the relationships among subordinates (e.g., who should be doing what, as well as how and when they should do it) and results in more predictable behaviors (Mintzberg, 1979, p. 83; Sutton & Rao, 2014, p. 107). If those interactions become intractable and evolve into dysfunctional conflicts, the managers can leverage their authority to take control of the situation, resolve those conflicts, and keep things on track (Lawrence & Lorsch, 1967/1986, pp. 146–151). As one entrepreneur succinctly puts it:

[Without hierarchy, start-ups] lose structure altogether and fall foul of many of the trappings of letting creativity prosper without control. (Freeman, 2013)

By curbing aimless exploration and dysfunctional conflicts, adding a hierarchical level prevents chaotic execution and improves commercial success.<sup>6</sup>

Taken together, the above arguments suggest the following hypothesis:

**Hypothesis (H2).** *The number of hierarchical levels in a start-up is positively associated with its subsequent commercial success.*

### 3 | METHODS

#### 3.1 | Data collection

In general, examining the relationships between a start-up's hierarchy and its creative/commercial success is empirically challenging for two main reasons. First, because start-ups are, by definition, newly established and thus predominantly private, it is difficult to gather a large sample dataset of these nascent firms and their performance. Second, even when such a dataset is obtainable, their hierarchy is hard to measure because their organizational charts are not publicly available. Because of these challenges in data collection and measurement, empirical research on the organizational structure of start-ups has remained sparse (Burton et al., 2019; Keum & See, 2017).

This study overcomes these challenges by leveraging a unique large-sample dataset of start-ups in the video game industry, which is one of the fastest-growing venues for entrepreneurial activities (Bies, 2017). This dataset was collected from three major sources widely cited by industry practitioners: MobyGames, GameRankings, and VGChartz. MobyGames keeps a comprehensive database of more than 190,000 games worldwide and, for each game, specifies the game development studio (henceforth, studio) and its employees' full name and job title.<sup>7</sup> I used this data to identify whether a studio is newly established and to measure its hierarchical levels. MobyGames also provides specific information on game attributes (e.g., genre, theme, business model) and financing method (e.g., whether the studio financed independently or whether it was funded by publishers), which is used to account for the task- and financing-level characteristics.

<sup>6</sup>A taller hierarchy may not necessarily hinder execution and commercial success by slowing the decision-making process, because the greater time required for decisions to pass through more hierarchical levels can be offset by the time needed to resolve conflicts and reach a consensus in a flatter one (Carzo & Yanouzas, 1969; Eisenhardt, 1989).

<sup>7</sup>MobyGames database has been used as a source of data in several studies on the video game industry (e.g., Katila, Piezunka, Reineke, & Eisenhardt, 2021; Lee & Csaszar, 2020; Mollick, 2012). Before measuring the variables, I cleaned this dataset according to the following steps. First, among the 20 functional domains categorized by MobyGames (i.e., "Administration," "Art/Graphics," "Audio," "Business," "Companies," "Creative Services," "Customer/Technical Support," "Design," "Localization," "Marketing," "Production," "Programming/Engineering," "Public Relations," "Quality Assurance," "Support," "Technology," "Thanks," "Video/Cinematics," "Writers," and "Others"), the category "Thanks" was excluded because it mainly consists of friends and family, rather than employees. Next, multiple job titles of an individual for a given game were merged into one—for instance, the two job titles (i.e., "CEO" and "Executive Producers") of Mikael Hed for the game *Angry Birds* were combined into one (i.e., "CEO / Executive Producers"). Finally, the names of all 753,407 individuals were manually checked to filter out fake employees. Name disambiguation was not required because this database provides a unique identifier (i.e., "developerId"; e.g., for Mikael Hed, "444,500").

GameRankings aggregates over 240,000 critics' review ratings from both offline and online sources, including those before its launch in 1999 (e.g., GamePro's review for Nintendo's *Super Mario 64*, which was posted a few months after this game was released in 1996), and computes their average (in terms of percentages), which is used as a measure for creative success. In turn, VGChartz documents the global sales figures for games that sold more than 10,000 units, which are used to measure commercial success. These figures from GameRankings and VGChartz were connected to the MobyGames data using a fuzzy string-matching algorithm on the game title (Python's FuzzyWuzzy package with a similarity threshold of 95 to account for typos and minor differences) and then by manually checking false or ambiguous matches.

Although these websites provide a unique dataset to investigate the hierarchy of start-ups and its performance implications, they have potential limitations. First, MobyGames specifies information not on the organizational chart but on the employees' job titles, the use of which may vary across studios (Bethke, 2003, p. 51), thus potentially leading to an imperfect measure of hierarchy. Second, GameRankings selects its sources and converts all types of ratings into percentages, and this aggregation process can be subject to potential biases (Schreier, 2012). Finally, VGChartz claims that for every game in every region, it extrapolates sales figures provided by retailers to estimate global sales figures, the reliability of which has been questioned (Carless, 2008). I later discuss how I mitigate these concerns using various robustness checks.

### 3.2 | Defining a start-up for empirical analyses

To be consistent with the theoretical conceptualization, I empirically define a start-up as a studio at the point when it releases its *first* game. These studios are, on average, <2 years old because it typically takes 2 years to develop a triple-A game—namely, a game with massive development and marketing budgets funded by major publishers (e.g., Infinity Ward's game *Call of Duty*, which raised \$500 million from its publisher Activision; Schreier, 2017, p. 200).

Defining start-ups as such results in a pooled cross-sectional dataset that mitigates potential selection bias, as more than 90% of studios fail after their first product. Furthermore, it mitigates potential measurement errors. This is because newly established studios can be accurately represented by their first game's credit (i.e., the list of employees and their job titles), as these start-ups generally focus all employees on their first game and lack the resources to simultaneously develop multiple games. However, after the first game, studios typically invest the profit from their previous game to run multiple projects and hire administrative staff. As these concurrent projects and staff are not systematically reported at the firm level in a given year, the credits after the first game can incompletely represent the studios, thus biasing the measurement of variables (especially their hierarchy and size).

To avoid these empirical concerns, I acquired a sample that matches the above empirical definition in the following way. First, I measured the variables as described in the following section. I then limited the observations to the first game that each company developed. Among these 6,510 observations, 276 games with more than one studio were removed to exclude potential spin-outs, joint ventures, and project teams of incumbents. This removal yielded a sample of 6,234 start-ups newly established during the period from 1971 to 2015.

### 3.3 | Measurement

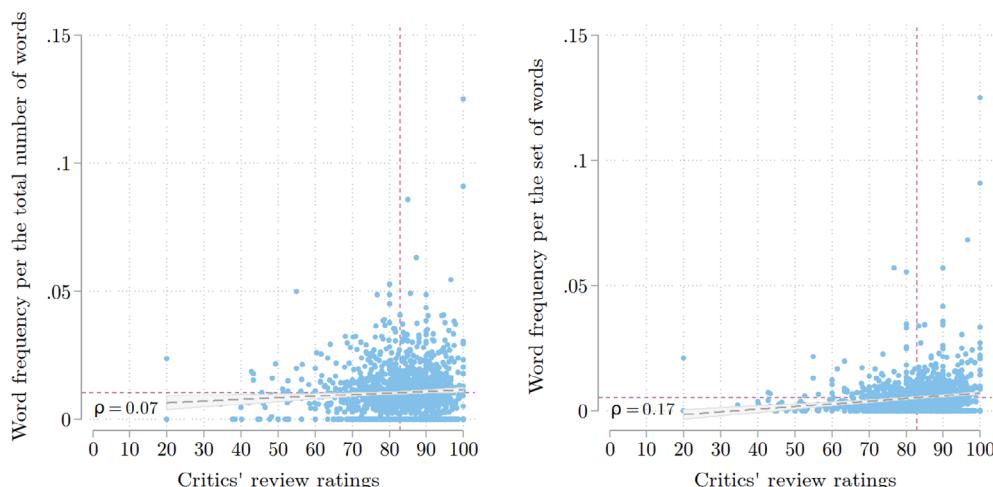
#### 3.3.1 | Dependent variables

As noted earlier, creative success is defined as the novelty of the product, whereas commercial success is defined as its profitability. Following prior work on creative industries (e.g., Uzzi & Spiro, 2005), creative success was measured by the average critics' review rating (on a scale from 0 to 100). This rating, to a large extent, reflects its novelty, as indicated by the fact that (1) terms such as "novel," "creative," "unique," "groundbreaking," "new," and "original" appear disproportionately in the reviews with above-60 ratings and (2) their word frequency is positively correlated with the rating (Figure 1). Thus, this rating is the industry standard used by practitioners as "an important benchmark" (Schreier, 2012)—so important that publishers "stick review score bonuses into contracts, offering extra payouts to [studios] whose games hit a certain threshold on video game aggregation sites like Metacritic or GameRankings" (Schreier, 2017, p. 216).

In turn, commercial success was measured by the number of units sold globally (in millions), which is comparable to the typical measure of commercial success (i.e., attendance) in studies of the movie industry (e.g., Zuckerman & Kim, 2003). These sales figures were used because information on profit, revenue, or return on investment is unavailable. As these figures reflect cumulative demand rather than profitability, I account for development costs by adding various task-level characteristics (e.g., the use of 3D technology).

#### 3.3.2 | Independent variable

In line with prior studies (e.g., Burton & Obel, 2004, p. 75; Lee & Csaszar, 2020; Puranam, 2018, pp. 106–108) and entrepreneurs' interpretation, I measured hierarchy in terms of the number of



**FIGURE 1** The scatter plot of creative success and the word frequency of terms relevant to novelty (i.e., "novel," "creative," "unique," "groundbreaking," "new," "original," "innovative," "rare," "inventive," "innovatory," "innovational," "unconventional," and "unorthodox") in critics' reviews. The left figure shows the word frequency per the total number of words, whereas the right shows that per the set of words (to remove the duplicated words). The line represents the linear prediction with its 95% confidence interval, whereas the dashed line indicates the mean of each variable. (a) Per the total number of words and (b) per the set of words

hierarchical levels. This number was computed by first categorizing each employee's job title into one of 12 levels (i.e., "Owner," "President," "VP," "CEO," "C-Suite," "Head," "Director," "Manager," "Producer," "Lead," "Supervisor," and "Other") and then counting the number of unique levels.

These 12 levels were derived from qualitative data, verified through interviews with practitioners, and measured using text analysis through the following five steps. First, from a rich array of qualitative data (e.g., Bethke, 2003; Grossman, 2003; Kent, 2010; Mencher, 2002; Schreier, 2017; Spaulding, 2009), I extracted keywords that indicate a hierarchical level and categorized those terms into 12 levels. Second, from the same data, I confirmed that these levels follow a vertical chain of command. For example, one founder explains:

Each lead ... will typically *report to* a director, who in turn *reports to* a studio head or president. (Spaulding, 2009, p. 2; emphasis added)

Third, through interviews with practitioners located in the US, South Korea, and Japan, I verified that these 12 categories typically represent distinct hierarchical levels. Fourth, to measure hierarchy for the empirical analyses, I build upon Lee and Csaszar's (2020) method to categorize each employee's job title into one of the 12 levels. This is done so by applying to Rules 1–12 in Table 2 in ascending order until a match is found—that is, if a job title includes terms relevant to that level. The list of relevant terms (see the third column in Table 2) includes abbreviations (e.g., "snrvp" for senior vice president) and typographical errors (e.g., "cheif" is a typo of "chief") that frequently appear in the MobyGames database. Finally, after categorizing all the job titles into these levels, I counted the number of hierarchical levels with at least one employee.

### 3.3.3 | Controls

For the employees' characteristics that may influence performance, I computed their amount and breadth of experience (Eggers, 2012; Shane, 2000), social capital developed through prior collaboration (Meier et al., 2019), and gender diversity (Hoogendoorn, Oosterbeek, & van Praag, 2013). Here, the amount of experience is measured by the average number of prior games that employees worked on, and its breadth by the average diversity of functional domains which employees worked in (for an illustration, Appendix A). In turn, social capital is computed by the average pair-wise social distance among employees in the prior five-year collaboration network (for an illustration, Appendix B). Finally, gender diversity is calculated by the proportion of female employees. As the information on gender is not specified in the MobyGames database, I used Python's gender-guesser package to predict each employee's gender with the full name.

### 3.3.4 | Fixed-effects

To account for unobserved variation that may be driving the results, I included the following dummy variables. At the task level, I added dummies for game characteristics that imply task complexity and development costs (e.g., genre, theme, and released platforms). At the

TABLE 2 The rules to categorize each employee's job title into a hierarchical level

Rule	Hierarchical level	If the job title includes any of these terms	Examples
1	Owner	owner, founder, chairman, creator, created, or made	“Created by,” “Created and Developed by,” “Chairman,” “Made by”
2	President	president or presidente (but not vice)	“President,” “President and CEO,” “President & CEO,” “President, North America”
3	VP	vp, evp, avp, svp, snrvp, vice president, or vice presidente	“Vice President,” “Vice President of Marketing,” “VP of Marketing,” “Senior Vice President”
4	CEO	ceo or any combination of {chief or cheif} and {executive, exec, executive, or executiver}	“CEO,” “Chief Executive Officer”
5	C-Suite	cco, cdo, cfo, cho, cio, clo, cmo, coo, cpo, cso, cto, or both chief and officer	“COO,” “CFO,” “Chief Creative Officer,” “Chief Operating Officer”
6	Head	head	“Head of Production,” “Studio Head,” “Head of Marketing,” “Head of Development”
7	Director	director, directo, diercto, dir, or dierctor	“Art Director,” “Director,” “Technical Director,” “Creative Director”
8	Manager	manager, mgr, or gm	“Project Manager,” “Product Manager,” “QA Manager,” “Production Manager”
9	Producer	producer	“Producer,” “Executive Producer”
10	Lead	lead or leader	“Lead Programmer,” “Lead Artist,” “Lead Tester,” “Lead Designer”
11	Supervisor	supervisor	“Supervisor,” “QA Supervisor,” “Music Supervisor,” “Test Supervisor”
12	Other	(includes none of the above)	“Testers,” “Programmers,” “Artists”

financing level, I introduced a binary dummy for whether the studio published the game independently without the help of publishers (i.e., indie), a binary dummy for whether it was funded by a major publisher (i.e., triple-A), and binary dummies for each of the top 40 publishers. At the employee level, I included 10 size-quantiles to compare among start-ups of a similar size in terms of the number of employees.<sup>8</sup> Finally, I used year dummies to control for common macroeconomic changes that influence all games. The inclusion of these fixed-effects, therefore, absorbs any variation attributable to invariant characteristics of the task, financing method, firm size, and macro-environment, and thus allows me to focus on the variables of interest.

<sup>8</sup>The 10 size-quantiles were used (1) to mitigate multicollinearity concerns because  $\log(\text{Size})$  is highly correlated ( $\rho = .59$ ) with *Hierarchy* (Section 4.1) and (2) to not impose any arbitrary assumptions in defining size categories given the lack of agreement in the literature on how to categorize firms by size (Burton & Obel, 2004, p. 171; Carland et al., 1984, p. 354). Likewise, there is no industry-specific size categorization by the number of employees, and industry practitioners typically categorize firms by revenue, rather than the number of employees (e.g., see NAICS Industry Code 511210 in the U.S. Small Business Administration's *Table of Small Business Size Standards*).

### 3.4 | Model specification

An ideal experimental design is to randomly assign hierarchy and examine how it affects creative and commercial success. Unfortunately, as such randomized designs were not attainable,<sup>9</sup> I apply the fixed-effects model to mitigate potential selection concerns with an extensive set of controls and fixed-effects. Hence, the empirical model is specified as follows:

$$\text{Performance}_i = \alpha + \beta \text{Hierarchy}_i + \Gamma \text{CONTROL}_i + \text{FIXED}_i + \varepsilon_i$$

where for start-up  $i$ ,  $\text{Performance}_i$  stands for one of the two performance outcomes (creative and commercial success),  $\text{Hierarchy}_i$  for the number of hierarchical levels,  $\text{CONTROL}_i$  for the vector of controls, and  $\text{FIXED}_i$  for the vector of fixed-effects (note that because most founders establish only one studio and each studio has only one observation, both founder and studio fixed-effects are not included in  $\text{FIXED}_i$  and that only one index,  $i$ , was used for simplicity).<sup>10</sup> The coefficients  $\beta$  and  $\Gamma$  are the estimated parameters, and  $\alpha$  and  $\varepsilon_i$  are the intercept and the random error term, respectively. In this model, the standard errors are clustered by genre.

## 4 | RESULTS

### 4.1 | Descriptive statistics and correlation matrix

Table 3 presents the descriptive statistics and the correlation matrix. First, the descriptive statistics show a substantial variation in the variables of interest. Among 6,234 start-ups, roughly a quarter (1,725 observations) were rated by critics, while <10% (494 observations) sold more than 10,000 units. Among these start-ups with over 10,000 unit sales, commercial success is highly skewed, as observed by Andersson et al. (2009, pp. 310–311).<sup>11</sup> This left-censored, right-skewed distribution of commercial success implies that most game development start-ups fail (Grossman, 2003, pp. ix–x), which is comparable to the fact that most start-ups, in general, fail. Given the left-censoring in the measures of creative and commercial success, the

<sup>9</sup>Among the candidates for a valid instrument, the most plausible one was the sudden leak in 2012 of Valve's *Handbook for New Employees*, which discusses why this company adopted a flat hierarchy. Another was the first article released in 2007 on "Holacracy," which lays out Ternary Software's principles of a flat hierarchy. Because both these documents would have influenced how start-ups perceive, shape, and legitimize a flat hierarchy—as they have done so for companies such as Zappos and Medium—but would be uncorrelated with their creative/commercial success, I tested as an instrument the dummies for post-2012 and post-2007. Unfortunately, both dummies and their variations incorporating information on founding locations turned out to be a "weak" instrument.

<sup>10</sup>Here, it is important to clarify two empirical assumptions for the second hypothesis. First, this hypothesis theoretically assumes that commercial success increases with the number of hierarchical levels, but at a decreasing rate. Yet, it is empirically tested as a linear relationship because boundedly rational entrepreneurs are presumably aware of this diminishing return and thus hardly adopt extreme hierarchical structures, for instance, where each employee occupies a hierarchical level. Second, this hypothesis is theoretically derived conditioned on creative success. However, the empirical model excludes creative success in examining the relationship between hierarchy and commercial success because creative success is an intermediate outcome of hierarchy and thus can bias the regression estimates for commercial success (Angrists & Pischke, 2008, pp. 64–68). Note that the results are consistent when controlling for creative success (Table S9).

<sup>11</sup>The right skewness of commercial success is accounted for by applying the natural logarithm. This transformation also applies to the amount of experience, to which one is added to avoid computing the logarithm of zero.

TABLE 3 Descriptive statistics and correlation matrix

	Descriptive statistics						Correlation matrix											
	Obs.	Mean	SD	Min	Max		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Dependent																		
(1) Creative (review rating)	1,725	66.24	14.96	10	100	1.00												
(2) Commercial (million units) <sup>a</sup>	494	0.97	2.53	0.01	30.26	0.30	1.00											
Independent																		
(3) Hierarchy	6,234	2.28	2.03	1	12	-0.03	0.08	1.00										
Controls																		
(4) Experience <sup>a</sup>	6,234	2.45	4.07	0	64	0.05	-0.04	0.47	1.00									
(5) Breadth	6,234	0.09	0.12	0	0.84	0.03	-0.09	0.25	0.69	1.00								
(6) SocialCapital	6,234	0.12	0.21	0	1	0.04	0.03	0.27	0.61	0.54	1.00							
(7) GenderDiversity	6,234	0.12	0.15	0	1	-0.02	-0.10	0.29	0.11	0.03	0.09	1.00						
Fixed-effects																		
(8) 3D	6,234	0.09	0	1	-0.04	-0.11	0.22	0.13	0.10	0.06	0.04	1.00						
(9) LicensedTitle	6,234	0.10	0	1	-0.12	0.14	0.21	0.16	0.10	0.13	0.05	0.03	1.00					
(10) BusinessModel	6,234	0.14	0	1	0.08	0.04	-0.07	-0.05	-0.04	-0.08	-0.03	-0.06	-0.07	1.00				
(11) Indie	6,234	0.31	0	1	0.07	0.01	-0.31	-0.31	-0.22	-0.26	-0.13	-0.12	-0.18	0.21	1.00			
(12) Triple-A	6,234	0.32	0	1	-0.05	0.04	0.49	0.40	0.24	0.29	0.18	0.13	0.18	-0.15	-0.44	1.00		
(13) Size <sup>a</sup>	6,234	26.11	1	2,450	0.13	0.23	0.59	0.27	0.12	0.15	0.17	0.16	0.11	-0.04	-0.16	0.28		

<sup>a</sup>In the correlation matrix, the natural logarithm was applied to commercial success, experience, and size to account for their right skewness.

main results may be interpreted for start-ups that reach a certain level of success (for the descriptive statistics and correlation matrix for the observations used in the regression analyses, see Table S1).

The descriptive statistics also show that a typical start-up consists of fewer than 30 employees (mean *Size* of 26.11) who, on average, have worked on approximately two projects before joining the start-up (mean *Experience* of 2.45) and are highly specialized in a few functional domains (mean *Breadth* of 0.09).<sup>12</sup> These statistics are consistent with various observations by industry practitioners (e.g., Schreier, 2017, p. 64; Spaulding, 2009, p. 6), thus implying that my data is a representative sample that closely matches the population of game development start-ups.

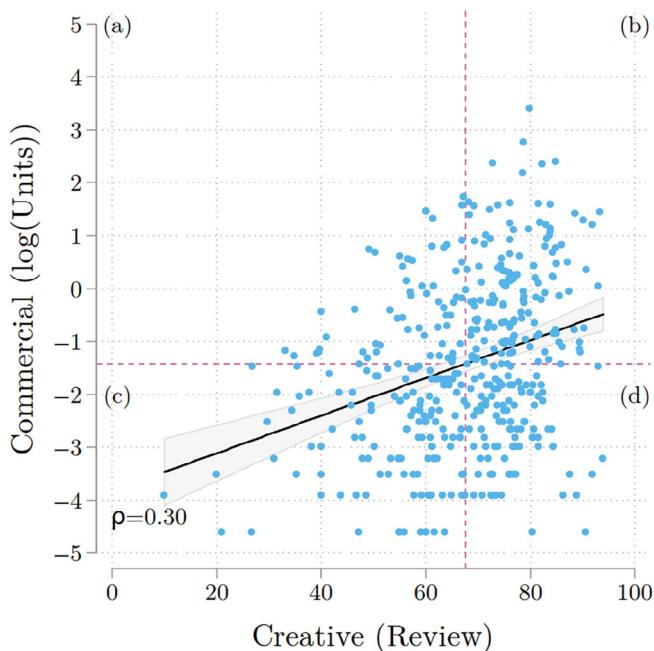
Next, I turn to the correlation matrix. This matrix indicates that the independent variable *Hierarchy* is highly correlated ( $\rho = .59$ ) with the control log(Size), which is thus discretized into 10 size-quantiles and included as fixed-effects in the main analyses. Another highly correlated control ( $\rho = .47$ ) is log(Experience), which is an important source of knowledge that influences start-ups' performance (Eggers, 2012; Shane, 2000). To investigate whether these variables raise multicollinearity concerns, I conducted the following three diagnostic tests. First, I computed the variance inflation factors and the condition indexes, each of which shows a maximum value (1.76 and 5.38, respectively) well below its customary threshold (10 and 30, respectively; Belsley, Kuh, & Welsch, 1980, p. 112; Kutner, Nachtsheim, & Neter, 2004). Second, I excluded log(Experience) and obtained robust results (Table S8). Finally, because these two diagnostics may misleadingly dismiss multicollinearity concerns (Kalnins, 2018), I ran separate regressions (a) with only *Hierarchy*, (b) with only log(Experience), and (c) with both variables, all of which exhibit consistent signs and relatively stable magnitudes for the two variables (Table S8). Together, three diagnostics imply that the main results are not severely biased by multicollinearity.

The correlation matrix also shows that creative and commercial success are positively correlated. This correlation is expected given that creative success partially explains commercial success: novel ideas tend to be more commercially successful. However, what is surprising is the significant amount of divergence between these two measures. This divergence is illustrated in Figure 2, which plots creative success (x-axis) against commercial success (y-axis). This figure shows that more than 40% of start-ups achieved one or the other: that is, received an above-average review rating but sold below-average global sales units, or vice versa (and are thus in quadrants (a) and (d)).

## 4.2 | Exploratory data analysis

To explain this divergence between creative and commercial success, hierarchy should substantively vary across start-ups. According to the descriptive statistics in Table 3, a typical start-up has roughly two hierarchical levels (mean *Hierarchy* of 2.28), which aligns with the common belief that start-ups are homogeneously flat. But, this mean can be misrepresentative because whether start-ups are flat will vary by their size (Blau, 1970; Burton & Obel, 2004, p. 168). If I

<sup>12</sup>To understand why the mean *Breadth* of 0.09 corresponds to a high level of functional specialization, consider an individual with 20 units of experience. If 19 of these units fall in the same functional domain, she would have  $Breadth = .095 (= 1 - (19^2 + 1^2)/20^2)$ , which is larger than the mean value.



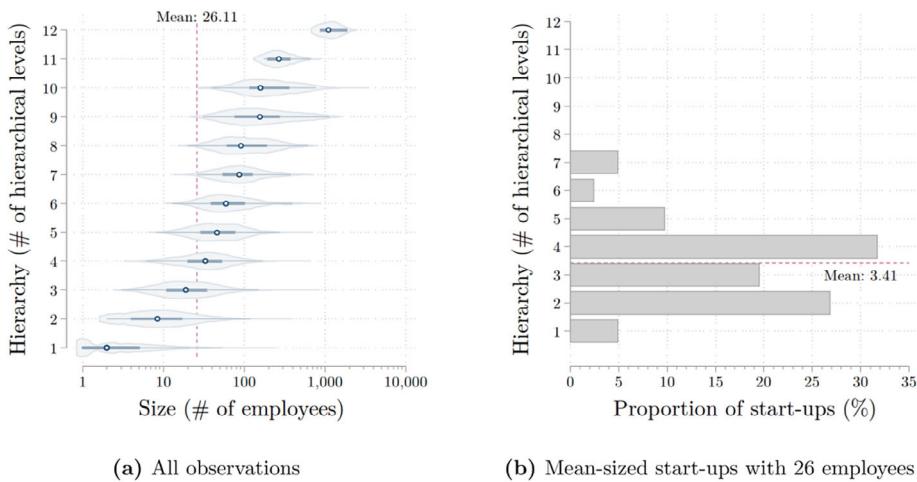
**FIGURE 2** The scatter plot of creative and commercial success. The line represents the linear prediction with its 95% confidence interval, whereas the dashed line indicates the mean of each performance measure

thus plot the number of hierarchical levels by the number of employees (Figure 3a), the two variables show a strong positive relationship.<sup>13</sup>

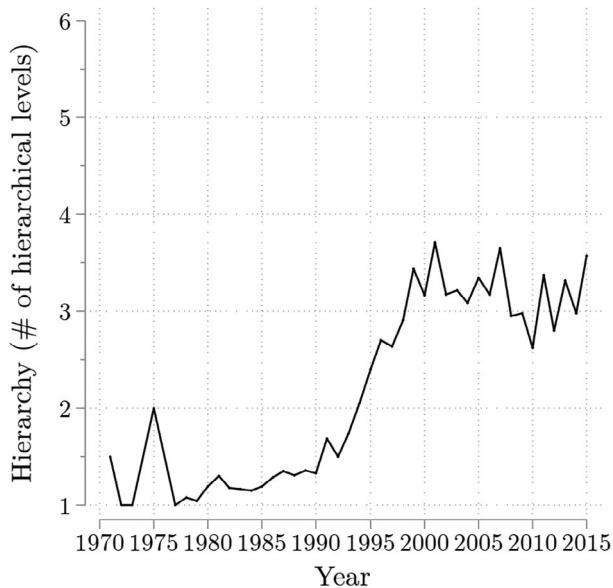
Here, if we focus just on the same-sized start-ups (i.e., observations vertically above a given point on the  $x$ -axis in Figure 3a), we can see that they vary substantively in terms of their hierarchy. For instance, among the mean-sized firms with 26 employees, the number of hierarchical levels ranges from one to seven (Figure 3b). This variation is congruent with Baron and Hannan's (2002, p. 9) observation that "[t]hough some observers might think that most start-ups look pretty much the same, or that the appropriate organizational design... for a high-tech venture is 'obvious,' the data suggest otherwise." Adding to their observation, this finding qualifies the conventional wisdom that start-ups are homogeneously flat.

In fact, start-ups at entry have become taller over the past four decades (Figure 4). That is, the average number of hierarchical levels when start-ups released their first game has increased from 1.3 in the 1980s to 3.2 in the 2010s. This finding is surprising because it counters not only the widespread distaste for hierarchy among practitioners but also the conventional wisdom in the literature on organizational design. This literature has posited that firms will flatten to streamline the decision-making process and flexibly adapt to today's rapid technological and economic changes (Burns & Stalker, 1961; Lawrence & Lorsch, 1967/1986). It has further asserted that firms will "flatten" (or "delayer") because more advanced information technology

<sup>13</sup>The observations with a large number of employees and hierarchical levels (in the upper right-hand corner of Figure 3a) represent the game development start-ups whose first product was a triple-A game (e.g., Infinity Ward and its first game *Call of Duty* mentioned in Section 3.2). To mitigate the concern that such extreme observations are driving the main findings, I ran a robustness test excluding those with more than 250 employees, and found consistent results (Table S5).



**FIGURE 3** The variation in start-ups' hierarchy. The left figure includes all 6,234 observations (i.e., start-ups of any size), whereas the right is for 43 mean-sized start-ups with 26 employees. (a) All observations and (b) mean-sized start-ups with 26 employees



**FIGURE 4** The trend towards a taller hierarchy. The line represents the average number of hierarchical levels in start-ups that released their first game in a certain year

(IT) will decrease the role of the middle managers in coordinating information (Rajan & Wulf, 2006; Zhou, 2013).

One plausible explanation for this “tallening” of start-ups is the advancements in IT (Edwards, 2005; Grossman, 2003, p. x; Spaulding, 2009, pp. 4–7). During the 1990s, major IT advancements (e.g., CD/DVDs, the Internet, and multi-core CPUs) were widely introduced and commercialized. These advancements have enabled studios to create more media-rich,

complex games (e.g., Massively Multiplayer Online Role-Playing Games), which require decomposition into smaller task components (e.g., programming, graphics, network) and a team of specialized individuals dedicated to a particular component. To integrate the efforts of these specialists, start-ups have introduced more hierarchical layers to compensate for their lack of alternative coordination devices. As one practitioner highlights:

Now, ... [studios] employ programming teams divided into sub-teams of coders dedicated to four or five specialties... requiring more individuals to take leadership positions on a given team. (Spaulding, 2009, pp. 6–7)

Thus, in contrast to the conventional wisdom about flattening of mature firms, this observation suggests that by increasing task complexity, more advanced IT can, in fact, “tallen” start-ups.<sup>14</sup>

Given this large-scale evidence of variation and trend in the hierarchy of start-ups, I discuss below the regression results of how this variation translates into a difference in their performances in terms of creative and commercial success (for the results, see Table 4).

### 4.3 | Regression results

Part A of Table 4 reports the main regression results using the fixed-effects model with standard errors clustered by genre. For ease of reading, only the baseline model (excluding all fixed-effects) and the full model (including all fixed-effects) are presented. The results of these models are consistent when the fixed-effects are sequentially added (Tables S3 and S4). Note that, as the measures of creative and commercial success are left-censored, these results may be interpreted for start-ups that reach a certain level of success.

I begin by examining the relationship between a start-up's hierarchy and its creative success. In Hypothesis 1, I proposed that a start-up's hierarchy is negatively associated with its creative success, because adding a hierarchical level can hamper cross-fertilization and prematurely filter out novel ideas. Supportive evidence for this hypothesis is provided in Models 1 and 2. Model 1 shows that when including just the controls, the coefficient of hierarchy is negative ( $p=.002$ ). This remains consistent in Model 2 ( $p<.001$ ) even when accounting for the vast list of fixed-effects. In Model 2, the coefficient size of hierarchy indicates that one additional hierarchical level is correlated with a 1.14-point decrease in the average review rating, which is a minor change representing <1.7% of the mean of 66.24 (or 0.08 standard deviations). Thus, these results offer empirical support for Hypothesis 1.

I move on to establish the relationship between hierarchy and commercial success. Hypothesis 2 states that a start-up's hierarchy is positively correlated with its commercial success, because adding a hierarchical level can prevent aimless exploration and dysfunctional conflicts among employees. Support for this hypothesis is displayed in Models 3 and 4. Model 3 (the baseline model) shows that commercial success is positively associated with hierarchy ( $p=.017$ ). In Model 4, the positive coefficient of hierarchy remains stable in terms of size ( $p=.004$ ), despite adding the long list of fixed-effects. The coefficient in Model 4 implies that

<sup>14</sup>I offer further empirical support for this conjecture regarding “tallening” by using Massively Multiplayer Online Role-Playing Games as a moderator (see Appendix S2).

TABLE 4 Regression results

	Part A		Part B		Part C			
	Main analyses		Testing the mechanism		Assessing the scope condition			
	(1) Creative	(2) Commercial	(2) Commercial	(log(Units))	(1) Creative	(2) Commercial		
	(Review)	(log(Units))	(log(Units))	(Review)	(Review)	(log(Units))		
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
Independent								
Hierarchy	-0.45 (0.13)	-1.14 (0.28)	0.09 (0.04)	0.14 (0.04)	0.22 (0.09)	0.22 (0.07)		
Moderator								
Hierarchy × Breadth					-0.47 (0.36)	-0.29 (0.16)		
Hierarchy × SocialCapital								
Hierarchy × Size								
Micro (1–9 employees)							-2.33 (1.35)	0.15 (0.46)
Small (10–49)							-1.07 (0.21)	0.17 (0.06)
Medium-sized (50–249)							-1.12 (0.34)	0.14 (0.06)
Large ( $\geq 250$ )							0.71 (0.41)	-0.06 (0.11)
Controls								
Log(experience)	2.23 (1.27)	1.26 (1.64)	-0.21 (0.15)	-0.10 (0.23)	-0.05 (0.25)	-0.06 (0.25)	1.11 (1.69)	-0.13 (0.19)
Breadth	-6.43 (6.13)	-0.17 (6.66)	-3.22 (0.99)	0.45 (2.31)	2.44 (2.18)	0.65 (2.37)	0.77 (7.54)	1.12 (1.91)
SocialCapital	0.61 (2.42)	4.34 (4.14)	1.56 (0.44)	0.11 (0.85)	0.12 (0.93)	1.33 (1.14)	3.66 (3.74)	0.06 (0.64)
GenderDiversity	-1.29 (2.20)	-6.24 (3.45)	-1.86 (0.71)	-1.10 (1.18)	-1.08 (1.19)	-1.17 (1.17)	-5.71 (3.21)	-1.31 (1.13)
Fixed-effects								
Task-level								
3D		-2.40 (1.14)		-0.59 (0.23)	-0.63 (0.23)	-0.67 (0.24)	-2.38 (1.17)	-0.49 (0.23)
LicensedTitle		-4.27 (1.44)		0.49 (0.23)	0.49 (0.22)	0.46 (0.21)	-4.24 (1.38)	0.43 (0.25)

TABLE 4 (Continued)

	Part A		Part B		Part C	
	Main analyses		Testing the mechanism		Assessing the scope condition	
	(1) Creative	(2) Commercial	(2) Commercial	(1) Creative	(2) Commercial	
(Review)	(log(Units))	(log(Units))	(log(Units))	(Review)	(log(Units))	
Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
BusinessModel	2.37 (1.95)	-0.30 (0.58)	-0.25 (0.62)	-0.25 (0.59)	2.41 (1.96)	-0.08 (0.39)
Genres	Y	Y	Y	Y	Y	Y
Themes	Y	Y	Y	Y	Y	Y
Platforms	Y	Y	Y	Y	Y	Y
Financing-level						
Indie	0.13 (0.87) -2.48 (1.04)	0.68 (0.64) -0.22 (0.39)	0.66 (0.68) -0.30 (0.42)	0.61 (0.71) -0.26 (0.39)	0.35 (0.86) -2.40 (1.09)	1.12 (0.86) 0.08 (0.42)
Triple-A						
Publishers (top 40)	Y	Y	Y	Y	Y	Y
Employee-level						
Size	10Q	10Q	10Q	10Q	NSF/EC	NSF/EC
Macro-level						
Year	Y	Y	Y	Y	Y	Y
No. observations	1,725	1,477	494	370	370	1,477
R-squared	0.01	0.26	0.05	0.63	0.63	0.27
						0.63

Note: Part A reports the main analyses on the relationships between hierarchy and creative/commercial success. Parts B and C display the supplementary analyses using moderators to test the theorized mechanism and the scope condition of firm size. Standard errors clustered by genre in parentheses. The number of observations decreased since singleton observations were dropped. For size fixed-effects, "10Q" stands for 10 size-quantiles, whereas "NSF/EC" for NSF and EC's size categorization.

one additional hierarchical level is correlated with a 14% increase in the number of units sold globally. In sum, these results lend supportive evidence for Hypothesis 2.

Together, Models 1 to 4 allude to an interesting observation that, for start-ups, adding a hierarchical level may entail a trade-off between creative and commercial success. What is surprising—especially considering the empirical context of the video game industry—is that one additional level is correlated with only a small decrease in creative success: that is, a 0.08 standard deviation decrease in average review rating. In contrast, it is associated with a large increase in commercial success: that is, a 14% increase in global sales units. For an average start-up in my sample, this increase translates to ~14,000 units, or \$0.84 million in revenue (assuming an average price of \$60; Yan & Gilbert, 2018). This observation hints that, even in creative industries, adding a hierarchical level could be substantially beneficial for start-ups in achieving commercial success at the very marginal expense of creative success.

#### 4.4 | Testing the theorized mechanism

Part B of Table 4 presents the results using moderators to test the theorized mechanism. For Hypothesis 2, I argued that adding a hierarchical level can lead to better commercial success by preempting aimless exploration and dysfunctional conflicts among employees. Because this mechanism is difficult to observe across a large sample, I test it using two moderators—namely, the employees' breadth of experience (henceforth, breadth) and social capital. The intuition (for details, see Appendix C) is that the more breadth the employees have, the more cognitively flexible they are, and thus the less likely to engage in misunderstandings and dysfunctional conflicts (Greer, de Jong, Schouten, & Dannals, 2018, p. 594). In turn, the more social capital that employees have developed through prior collaborations, the more mutual understanding and trust they share, and thus the less likely to engage in dysfunctional conflicts (Meier et al., 2019). Because both moderators would decrease the extent to which dysfunctional conflicts arise, I expect that hierarchy will have the largest positive relationship with commercial success when employees lack breadth or social capital (i.e., when dysfunctional conflicts are the most severe), and that this positive relationship will decrease as these moderators increase.

In Models 5 and 6, the coefficients of *Hierarchy* are positive ( $p=.018$  and  $p=.006$ , respectively). Although the interaction terms (i.e., *Hierarchy*  $\times$  *Breadth* and *Hierarchy*  $\times$  *SocialCapital*) are both negative as expected, only the coefficient for the interaction between hierarchy and social capital is significant ( $p=.095$ ). Because social capital has a value between 0 and 1, the coefficient of *Hierarchy* can be interpreted as the relationship between hierarchy and commercial success given *SocialCapital* = 0. Hence, these results show that the positive relationship between hierarchy and commercial success is the largest when employees lack social capital and that this positive relationship decreases as these moderators increase. Hence, Model 6 supports my theorized mechanism.

#### 4.5 | Assessing the scope condition of firm size

As I defined start-ups as nascent firms of any size (theoretically in Section 2.1 and empirically in Section 3.2), a potential question would be whether the observed relationships between hierarchy and creative/commercial success vary across different sizes—and if not, are purely driven by the extremely small and large start-ups. To investigate whether this is the case, I extend my

analyses by interacting hierarchy with size. Because the 10 size-quantiles used in the main analyses are specific to my data and are thus difficult to generalize and because there is no industry standard for categorizing studios by the number of employees, I instead adopt a size categorization defined by the National Science Foundation's (NSF) *Business R&D and Innovation Survey—Microbusiness* and the European Commission's (EC) *Recommendation of May 6, 2003 Concerning the Definition of Micro, Small and Medium-Sized Enterprises*. This categorization labels a firm as a “micro-business” if it has fewer than 10 employees, a “small enterprise” if between 10 and 49 employees, a “medium-sized enterprise” if between 50 and 249 employees, and a “large enterprise” if at least 250 employees.

Part C of Table 4 displays the results for the marginal effects of hierarchy on creative/commercial success by these four size categories. These results show that its marginal effects are significant for small and medium-sized start-ups, but not for micro-businesses and large enterprises. Along with the robustness checks excluding these two categories (Table S5), these results suggest that extremely small and large firms are not driving the main findings. Moreover, they imply that hierarchy may become consequential when start-ups grow beyond micro-businesses to small/medium-sized enterprises, thereby providing empirical support to the burgeoning stream of research on scaling (DeSantola & Gulati, 2017; Lee & Kim, 2021; Sutton & Rao, 2014).

These results also offer two interesting observations. First, as firm size increases, the marginal effect of hierarchy on creative success seems to increase, whereas that on commercial success seems to decrease. Second, the trade-off between creative and commercial success seems to flip for large start-ups—that is, adding a hierarchical level may be conducive to creative success but be detrimental to commercial success. These observations seemingly contradict the conventional wisdom that, as firms grow, they impose more hierarchical levels to gain efficiency in execution, at the cost of flexibility in ideation (Blau, 1970; Burton & Obel, 2004, pp. 168–171). As these observations cannot be further tested with the present data, more research is needed to determine whether (and if so, why) large start-ups may face the opposite trade-off between creative and commercial success.

## 4.6 | Robustness checks

Although the video game websites offer a rare dataset to observe start-ups' hierarchy and performance outcomes, they carry some limitations in measuring the variables of interest. Also, despite delving into the empirical context, exogenous variation in hierarchy was unobtainable. As these data limitations raise potential empirical concerns that deserve careful considerations, I ran an extensive array of stress tests (for details, see Appendix S3). These tests include (a) sequentially adding controls and fixed-effects, (b) subsampling, (c) using alternative clustered standard errors, (d) using alternative model specifications, (e) using alternative measurements, (f) addressing multicollinearity concerns, (g) controlling for creative success when regressing commercial success on hierarchy, (h) controlling for span of control, and (i) addressing potential selection based on omitted variables, such as start-ups' founding locations (Saxenian, 1996), their founder's prior employment (Baron, Hannan, & Burton, 1999), their employees' ethnic backgrounds (Gruenfeld & Tiedens, 2010, p. 1263), their star developers (Baron, Burton, & Hannan, 1996), their horizontal division of tasks (Sine et al., 2006), their game's platforms (Rietveld & Eggers, 2018), and their publishers (Katila et al., 2021). These test results (reported in Tables S3–S12) are robust.

Although all the above analyses validate my findings, the concern of selection based on unobserved variables may yet remain. To further mitigate this concern, I follow prior empirical

**TABLE 5** Results of Oster's (2019) test

			Oster's test		
Specification		$R_{full}$	(1) $R_{max} = R_{full}$	(2) $R_{max} = 1.3R_{full}$	(3) $R_{max} = 1$
Creative	Model 2	0.26	-164.42	-12.27	-2.51
Commercial	Model 4	0.63	-73.89	-2.81	-2.20

Note: For these results, I use the R-squared of Models 2 and 4 in Table 4 (with the largest R-squared) and three values of  $R_{max}$  (i.e.,  $R_{full}$ ,  $1.3R_{full}$ , and 1).

studies (e.g., Mian & Sufi, 2014; Starr, Frake, & Agarwal, 2019) and conduct a series of diagnostic tests using a formal approach proposed by Altonji, Elder, and Taber (2005) and developed by Oster (2019). This method computes a parameter  $\delta$ , the size of which reflects the amount of variation that the unobservable should explain (relative to the observables) to reduce the coefficient of hierarchy to zero and thus nullify the results. Specifically,  $|\delta|=1$  implies that the unobservable would need an explanatory power as strong as all the observables to invalidate the results, whereas  $|\delta|>1$  suggests that the unobservable would need a stronger power than do all the observables.

To estimate  $\delta$ , this test requires (a) a set of advanced controls that are unrelated to the set of proportionally related unobservables and (b) a value of  $R_{max}$  which represents the R-squared from a hypothetical regression that includes both observed and unobserved variables. For the advanced controls, I used all of the fixed-effects. For  $R_{max}$ , I employed the following three values. Following Oster's (2019) experimental evidence and recommendation, I applied  $R_{full}$  and  $1.3R_{full}$ , where  $R_{full}$  represents the R-squared of the most saturated model (i.e., Models 2 and 4 in Table 4). To be the most stringent, I also used the maximum possible value of R-squared (i.e., one).

Table 5 reports the results of these diagnostic tests. The results for creative success (the first row) imply that to reduce the coefficient of hierarchy to zero, the amount of variation explained by an unobservable needs to be 2.51–164.42 times as large as the amount of variation explained by the observables included. In turn, for commercial success (the second row), this amount is between 2.20 and 73.89. Because these values are considerably larger than one—even when assuming the maximum possible value of R-squared—and because the full models include an extensive array of controls and fixed-effects, these results suggest that selection bias from unobservables is unlikely to overturn the main findings.

## 5 | DISCUSSION

How start-ups should configure their hierarchy is a key area of inquiry for both organization scholars and practitioners. To address this question, my work provides a large sample study using a novel dataset of game development start-ups, and supplements these empirical analyses with qualitative observations. I discuss the theoretical and managerial implications of my findings below.

### 5.1 | Theoretical implications

My study contributes to the literature on entrepreneurship and organizational design on five fronts. First, this study offers large-scale evidence of the heterogeneity in the hierarchy of start-

ups, albeit within a single industry. Thus far, it has been widely taken for granted that these nascent firms should be and are thus flat (DeSantola & Gulati, 2017, p. 657). This conventional wisdom has remained untested partly because empirical research on the organizational structure of start-ups has been stymied by data limitations (Burton et al., 2019; Keum & See, 2017). Along with recent observations (e.g., Baron & Hannan, 2002), my study qualifies this common belief by showing that same-sized, small nascent firms (on average, <2 years old) can significantly vary in their hierarchical levels. By shedding light on this important yet understudied variation, this study complements the recent work (e.g., Baron et al., 1996, 1999; Colombo & Delmastro, 2008) and calls for more research on the initial conditions that lead to the emergence of organizational structure at the very early stage of the firm life-cycle.

Second, this article presents a boundary condition for the concept of “flattening” by showing a trend toward a taller hierarchy at entry among game development start-ups. My qualitative analyses suggest that this trend has occurred partly because more advanced IT has increased the task complexity in game development, requiring more task decomposition and specialization. To integrate these specialized task components, start-ups have added more hierarchical levels to compensate for their lack of alternative coordination devices. This “tallening” of start-ups stands in stark contrast to the flattening of mature firms (Rajan & Wulf, 2006; Zhou, 2013), which have other systematic ways to coordinate (e.g., formalized routines, informal structure, or organizational culture; Lawrence & Lorsch, 1967/1986, pp. 12–13; Puranam, 2018, pp. 128–144). Hence, these contrasting trends imply that whether firms flatten may depend on whether they have other coordination devices that can substitute for hierarchy.

Third, this study highlights the underappreciated importance of organizational structure for start-ups. As employees are the main source of competitive advantage for start-ups (Eggers, 2012; Shane, 2000), it would seem intuitive that an effective organizational structure could play a crucial role in their performance—arguably even more so than in that of mature firms (Burton et al., 2019). But prior research on start-ups, to date, has paid little attention to their organizational structure (Burton et al., 2019; DeSantola & Gulati, 2017; Wasserman, 2012, p. 3), mostly focusing instead on their resource mobilization (Clough, Fang, Vissa, & Wu, 2019). On the other hand, extant theories on organizational structure have primarily been examined in the context of large, mature organizations (Colombo & Grilli, 2013; Sine et al., 2006). This tendency to neglect the organizational structure of start-ups is mainly due to the presumption that these firms are simply too small and undifferentiated to necessitate any form of organizational structure (Blau, 1970; Mintzberg, 1979), which could hamper flexibility in adapting to their hostile environment (Burns & Stalker, 1961; Lawrence & Lorsch, 1967/1986). However, this dominant scholarly view has largely underestimated their coordination challenges by overlooking the fact that these nascent firms lack alternative coordination devices (e.g., standard operating procedures; Sine et al., 2006; Stinchcombe, 1965, pp. 148–150). Given this deficiency, this study finds that a start-up’s hierarchy has significant relationships with its creative and commercial success. My results also demonstrate that these relationships become considerable when start-ups grow beyond micro-businesses (1–9 employees) to small/mid-sized (10–249 employees) enterprises, thus providing empirical support to research on scaling (DeSantola & Gulati, 2017; Lee & Kim, 2021; Sutton & Rao, 2014). These findings imply that even for these small nascent firms in such a hostile environment as the video game industry, their hierarchy is a consequential factor that warrants further investigation. Hence, this study answers recent calls for more empirical research on how the existing theories on organizational design apply to new ventures (Burton et al., 2019; Dushnitsky & Matusik, 2019; van de Ven et al., 2013).

Fourth, my work speaks more broadly to the longstanding debate over whether the hierarchy is conducive to firm performance. Although the pervasiveness of hierarchy—even in start-ups, as shown in this study—would seem to indicate the answer is yes, prior studies have offered opposing arguments (for reviews, see Anderson & Brown, 2010; Lee & Edmondson, 2017; Puranam, 2018). To reconcile this debate, I distinguished between two measures of a firm's performance (i.e., creative and commercial success; Kaplan & Vakili, 2015; Kaul, 2013; Uzzi & Spiro, 2005) and showed that these two measures can be variously associated with its hierarchy. These results suggest that hierarchy is not universally good or bad, but rather its efficacy depends on the type of performance firms pursue. Put differently, firms should organize differently depending on their goal—that is, “flatten” their hierarchy for creative success, whereas “tallén” for commercial success. Extending the burgeoning stream of research that revisits hierarchy and its behavioral mechanisms (e.g., Csaszar, 2013; Keum & See, 2017; Lee & Csaszar, 2020; Lee, Lee, Lee, & Braha, 2020; Reitzig & Maciejovsky, 2015), my study thus helps resolve the contrasting assessments of hierarchy by sorting out the condition under which hierarchy can be conducive or detrimental to firms.

Finally, this study contributes to research on the interplay between formal and informal structures. Although this interplay received much attention in the foundational studies on organization theory (e.g., Blau & Scott, 1962/2003; March & Simon, 1958/1993; Simon, 1947/1997), it has since been largely neglected as the two subsequent streams of research on formal and informal structures have remained disconnected (McEvily et al., 2014). As bridging the gap between these two streams can afford a better understanding of organizational functioning and performance (Puranam, 2018, pp. 128–144), recent studies have started to reinvigorate the discussion on the interplay between formal and informal structures (e.g., Clement & Puranam, 2018; Slade Shantz, Kistruck, Pacheco, & Webb, 2020). Adding to this growing scholarly effort, my mechanism tests offer rare empirical evidence that hierarchy (i.e., one dimension of formal structure) can compensate for informal structure (e.g., social capital) or the lack thereof. These results suggest that, as these structures are deeply intertwined, it is important to theoretically and empirically conceive of formal and informal structures in combination, rather than in isolation, to account for their true performance implications.

## 5.2 | Managerial implications

[Start-ups'] managers tend to worry about avoiding bureaucracy but are blind to the danger of chaos. (Davila et al., 2010)

My work offers managerial implications for how start-ups should organize to capture entrepreneurial opportunities. For decades, academics, management gurus, and popular media outlets have argued that “authoritarian,” tall hierarchies are outmoded and will be supplanted by “egalitarian,” flat structures (Leavitt, 2005; Sutton & Rao, 2014, p. 107; Urwick, 1956). In recent years, this argument has been largely substantiated by a few “successful” flat start-ups, such as Valve, Zappos, Github, Medium, and Buffer (Burton et al., 2017; Lee & Edmondson, 2017, p. 38; Puranam & Håkonsson, 2015; Puranam, 2018, p. 137). As these firms constantly garner much attention for their egalitarian ideal—which itself is a signal of their rarity—the myth that start-ups should be flat (otherwise referred to as “flat organization,” “holacracy,” or “boss-less firm”) has become widespread among entrepreneurs (Davila et al., 2010; Gulati &

DeSantola, 2016). As one founder succinctly put it: “Flat [is] startup-y and awesome. Structure [is] BigCorp-y and boring” (Savage, 2015).

However, most of these experiments have failed (Leavitt, 2005, p. 29; Puranam, 2018, p. 137), as these start-ups drifted into aimless exploration and dysfunctional conflicts, eventually abandoning their flat hierarchies (Ferro, 2016; Maier, 2013; Miller, 2014). As the very entrepreneur who once thought “Flat [is] start-up-y and awesome” explains: “We had hoped that being flat would let us move faster and be more creative, but ... we ended up with an unspoken hierarchy that actually slowed down our ability to execute” (Savage, 2015). In line with this anecdote, my study cautions against the myth of the flat start-up, suggesting that adding a few hierarchical levels of managers can substantially help start-ups achieve commercial success and survival in their hostile environments, albeit at a potentially marginal cost of creative success.

### 5.3 | Limitations and future work

Although my study makes progress in understanding the organizational structure of start-ups, it has some limitations which could be addressed by future work. First, the empirical analyses were carried out within a single industry. One plausible scope condition of my findings is that the task is creative and complex. If the task does not call for creativity, hierarchy may not entail a trade-off, but may still be positively correlated with commercial success. In turn, if the task is so simple as to require only a few employees or as to not necessitate any coordination among employees, hierarchy needs not to be imposed. Future research could test the generalizability in different industries. Second, despite the comprehensive list of controls and the extensive array of robustness checks, there might be some form of unobserved heterogeneity. While my qualitative observations give confidence to my theoretical predictions, future work could revisit these predictions using an experimental design that randomly assigns hierarchy. Third, as critics review published games, their review ratings could reflect the game’s overall quality in terms of both its novelty and its execution (Rietveld & Eggers, 2018, p. 312). Future studies could revisit my findings using a more direct measure of creative success, for instance, by asking the critics to review only the game’s description in terms of its novelty but not allowing them to play the game and observe its execution. Fourth, this study concentrated on one dimension of organizational structure, namely hierarchy (i.e., the vertical division of tasks). Future studies could expand my work by exploring other dimensions—for instance, the horizontal division of tasks, for which I surprisingly find in the robustness checks negative associations with both creative and commercial success. Fifth, each start-up’s hierarchy was cross-sectionally measured when it released its first product (on average, <2 years from its founding). Future work could examine how its hierarchy evolves from its founding and changes after its first product. Sixth, this article probed into the consequences of start-ups’ hierarchy. Drawing upon my descriptive findings regarding its variation and trend, future research could investigate the antecedents of hierarchy and its tailoring. Finally, given the motivation, this study focused on start-ups. Although the empirical findings could hold for mature firms, there are potential reasons why the mechanisms will attenuate or be absent. In particular, unlike start-ups, mature firms typically have (a) adequate resources to organize integrating committees (Lawrence & Lorsch, 1967/1986, pp. 12–13), (b) formalized routines (Stinchcombe, 1965, p. 148), and (c) well-established informal structure or culture that can substitute for hierarchy (McEvily et al., 2014; Meier et al., 2019). Future studies could examine the performance implications of hierarchy in the context of mature firms.

## 6 | CONCLUSION

This article began by describing the myth of the flat start-up and then sought to revisit this untested myth. This was done by unveiling intriguing patterns in the hierarchy of start-ups and by investigating how it may entail a trade-off between creative and commercial success. By identifying its relationship with this fundamental dilemma, my work calls for more attention to the organizational structure of start-ups and makes headway in understanding their survival and growth.

## ACKNOWLEDGEMENTS

I am particularly indebted to Felipe Csaszar, Gautam Ahuja, Gerald Davis, Justin Frake, and Scott Page, who provided constructive feedback throughout this project. I also thank Seth Carnahan, Sungyong Chang, Judith Dyer, Christine Feak, Cheng Gao, Lindred Greer, Derek Harmon, Reuben Hurst, Hyo Kang, Samina Karim, Riitta Katila, Daniel Keum, Suresh Kotha, Megan Lawrence, Jeho Lee, Phanish Puranam, Chris Rider, Jordan Siegel, Jim Westphal, Pai-Ling Yin, and Maggie Zhou. In addition, I acknowledge the helpful suggestions from the seminar participants at Michigan Ross, the Wharton School, UIUC, UCL, University of Texas at Austin, SFU, Seoul National University, the Strategy Science Doctoral Workshop (2018), the East Coast Doctoral Conference (2019), the Trans-Atlantic Doctoral Conference (2019), the Consortium on Competitiveness and Cooperation Conference (2019), the INFORMS/Organization Science Dissertation Proposal Competition (2019), and the Strategic Management Society Annual Conference (2019). I am grateful to J.P. Eggers and anonymous reviewers for their valuable comments. Finally, research support from the Kwanjeong Educational Foundation is gratefully acknowledged. All errors in this article are the sole responsibility of the author.

## DATA AVAILABILITY STATEMENT

The data that support the findings of this study were derived from the following resources available in the public domain: MobyGames, GameRankings, and VGChartz.

## ORCID

Saerom (Ronnie) Lee  <https://orcid.org/0000-0001-9831-4910>

## REFERENCES

- Altonji, J. G., Elder, T. E., & Taber, C. (2005). Selection on observed and unobserved variables: Assessing the effectiveness of catholic schools. *Journal of Political Economy*, 113(1), 151–184.
- Anderson, C., & Brown, C. E. (2010). The functions and dysfunctions of hierarchy. *Research in Organizational Behavior*, 30, 55–89.
- Andersson, F., Freedman, M., Haltiwanger, J., Lane, J., & Shaw, K. (2009). Reaching for the stars: Who pays for talent in innovative industries? *Economic Journal*, 119(538), 308–332.
- Angrist, J., & Pischke, J. (2008). *Mostly harmless econometrics: An empiricist's companion*. Princeton, NJ: Princeton University Press.
- Baker, T., & Nelson, R. E. (2005). Creating something from nothing: Resource construction through entrepreneurial bricolage. *Administrative Science Quarterly*, 50(3), 329–366.
- Bangle, C. (2001). How BMW turns art into profit. *Harvard Business Review*, 79, 47–55.
- Baron, J. N., Burton, M. D., & Hannan, M. T. (1996). The road taken: Origins and evolution of employment systems in emerging companies. *Industrial and Corporate Change*, 5(2), 239–275.

- Baron, J. N., & Hannan, M. T. (2002). Organizational blueprints for success in high-tech start-ups: Lessons from the Stanford project on emerging companies. *California Management Review*, 44(3), 8–36.
- Baron, J. N., Hannan, M. T., & Burton, M. D. (1999). Building the iron cage: Determinants of managerial intensity in the early years of organizations. *American Sociological Review*, 64(4), 527–547.
- Beaudoin, M. (2016). Why I quit my dream job at Ubisoft. Retrieved from [https://gamasutra.com/blogs/MaximeBeaudoin/20160125/264180/Why\\_I\\_Quit\\_my\\_Dream\\_Job\\_at\\_Ubisoft](https://gamasutra.com/blogs/MaximeBeaudoin/20160125/264180/Why_I_Quit_my_Dream_Job_at_Ubisoft).
- Belsley, D. A., Kuh, E., & Welsch, R. E. (1980). *Regression diagnostics: Identifying influential data and sources of collinearity*. New York, NY: Wiley & Sons.
- Bethke, E. (2003). *Game development and production*. Plano, TX: Wordware.
- Bies, B. (2017). Gaming entrepreneurs: Do they exist? Retrieved from <https://medium.com/@brianbies/gaming-entrepreneurs-do-they-exist-ba63a46db2e6>.
- Blau, P. M. (1970). A formal theory of differentiation in organizations. *American Sociological Review*, 35(2), 201–218.
- Blau, P. M., & Scott, W. R. (1962/2003). *Formal organizations: A comparative approach*. Redwood City, CA: Stanford University Press.
- Burns, T. E., & Stalker, G. M. (1961). *The Management of Innovation*. London, U.K.: Tavistock.
- Burton, M. D., Colombo, M. G., Rossi-Lamastra, C., & Wasserman, N. (2019). The organizational design of entrepreneurial ventures. *Strategic Entrepreneurship Journal*, 13(3), 243–255.
- Burton, R. M., Håkonsson, D. D., Nickerson, J., Puranam, P., Workiewicz, M., & Zenger, T. (2017). Github: Exploring the space between boss-less and hierarchical forms of organizing. *Journal of Organization Design*, 6(1), 10.
- Burton, R. M., & Obel, B. (2004). *Strategic organizational diagnosis and design: The dynamics of fit*. Boston, MA: Springer.
- Carland, J. W., Hoy, F., Boulton, W. R., & Carland, J. A. C. (1984). Differentiating entrepreneurs from small business owners: A conceptualization. *Academy of Management Review*, 9(2), 354–359.
- Carless, S. (2008). Analysis: What VGChartz does (and doesn't) do for the game biz. Retrieved from [https://gamasutra.com/view/news/109870/Analysis\\_What\\_VGChartz\\_Does\\_And\\_Doesnt\\_Do\\_For\\_The\\_Game\\_Biz](https://gamasutra.com/view/news/109870/Analysis_What_VGChartz_Does_And_Doesnt_Do_For_The_Game_Biz).
- Carzo, R., & Yanouzas, J. N. (1969). Effects of flat and tall organization structure. *Administrative Science Quarterly*, 14(2), 178–191.
- Clement, J., & Puranam, P. (2018). Searching for structure: Formal organization design as a guide to network evolution. *Management Science*, 64(8), 3879–3895.
- Clough, D. R., Fang, T. P., Vissa, B., & Wu, A. (2019). Turning lead into gold: How do entrepreneurs mobilize resources to exploit opportunities? *Academy of Management Annals*, 13(1), 240–271.
- Colombo, M. G., & Delmastro, M. (2008). The determinants of the corporate hierarchy. In *The economics of organizational design: Theoretical insights and empirical evidence* (pp. 108–134). London, UK: Palgrave Macmillan.
- Colombo, M. G., & Grilli, L. (2013). The creation of a middle-management level by entrepreneurial ventures: Testing economic theories of organizational design. *Journal of Economics and Management Strategy*, 22(2), 390–422.
- Covin, J. G., & Slevin, D. P. (1989). Strategic management of small firms in hostile and benign environments. *Strategic Management Journal*, 10(1), 75–87.
- Csaszar, F. A. (2012). Organizational structure as a determinant of performance: Evidence from mutual funds. *Strategic Management Journal*, 33(6), 611–632.
- Csaszar, F. A. (2013). An efficient frontier in organization design: Organizational structure as a determinant of exploration and exploitation. *Organization Science*, 24(4), 1083–1101.
- Csaszar, F. A. (2021). A note on calculating the average span of control. *Journal of Organization Design*, 10, 83–84.
- Csaszar, F. A., & Eggers, J. P. (2013). Organizational decision making: An information aggregation view. *Management Science*, 59(10), 2257–2277.
- Davila, A., Foster, G., & Jia, N. (2010). Building sustainable high-growth startup companies: Management systems as an accelerator. *California Management Review*, 52(3), 79–105.
- Decker, R., Haltiwanger, J., Jarmin, R., & Miranda, J. (2014). The role of entrepreneurship in U.S. job creation and economic dynamism. *Journal of Economic Perspectives*, 28(3), 3–24.

- DeSantola, A., & Gulati, R. (2017). Scaling: Organizing and growth in entrepreneurial ventures. *Academy of Management Annals*, 11(2), 640–668.
- Dougherty, D. (1992). Interpretive barriers to successful product innovation in large firms. *Organization Science*, 3(2), 179–202.
- Dushnitsky, G., & Matusik, S. F. (2019). A fresh look at patterns and assumptions in the field of entrepreneurship: What can we learn? *Strategic Entrepreneurship Journal*, 13(4), 437–447.
- Edwards, R. (2005). Hierarchy of a game development company. Retrieved from <http://ign.com/articles/2005/11/30/hierarchy-of-a-game-development-company>.
- Eggers, J. P. (2012). All experience is not created equal: Learning, adapting, and focusing in product portfolio management. *Strategic Management Journal*, 33(3), 315–335.
- Eggers, J. P., & Kaul, A. (2018). Motivation and ability? A behavioral perspective on the pursuit of radical invention in multi-technology incumbents. *Academy of Management Journal*, 61(1), 67–93.
- Eisenhardt, K. M. (1989). Making fast strategic decisions in high-velocity environments. *Academy of Management Journal*, 32(3), 543–576.
- Fayol, H. (1949/2013). *General and industrial management*. Mansfield Centre, CT: Martino.
- Ferro, S. (2016). What happened when this major company got rid of all its bosses. Retrieved from [https://huffpost.com/entry/why-you-need-a-boss\\_n\\_569fdddbbe4b0fca5ba765409](https://huffpost.com/entry/why-you-need-a-boss_n_569fdddbbe4b0fca5ba765409).
- Fleming, L., Mingo, S., & Chen, D. (2007). Collaborative brokerage, generative creativity, and creative success. *Administrative Science Quarterly*, 52(3), 443–475.
- Freeman, J. (1972). The tyranny of structurelessness. *Berkeley Journal of Sociology*, 17, 151–164.
- Freeman, W. (2013). How to build a stronger development team. Retrieved from <https://develop-online.net/studio-profile/how-to-build-a-stronger-development-team/0117634>.
- Garicano, L., & Posner, R. A. (2005). Intelligence failures: An organizational economics perspective. *Journal of Economic Perspectives*, 19(4), 151–170.
- Garicano, L., & Wu, Y. (2012). Knowledge, communication, and organizational capabilities. *Organization Science*, 23(5), 1382–1397.
- Ghiselli, E. E., & Siegel, J. P. (1972). Leadership and managerial success in tall and flat organization structures. *Personnel Psychology*, 25(4), 617–624.
- Graicunas, V. A. (1937). Relationship in organization. In L. Gulick & L. Urwick (Eds.), *Papers on the science of administration* (pp. 183–187). New York, NY: Institute of Public Administration.
- Greer, L. L., de Jong, B. A., Schouten, M. E., & Dannals, J. E. (2018). Why and when hierarchy impacts team effectiveness: A meta-analytic integration. *Journal of Applied Psychology*, 103(6), 591–613.
- Greer, L. L., & van Kleef, G. A. (2010). Equality versus differentiation: The effects of power dispersion on group interaction. *Journal of Applied Psychology*, 95(6), 1032–1044.
- Grossman, A. (2003). *Postmortems from game developer*. San Francisco, CA: CMP Books.
- Gruenfeld, D. H., & Tiedens, L. Z. (2010). Organizational preferences and their consequences. In S. T. Fiske, D. T. Gilbert, & G. Lindzey (Eds.), *Handbook of social psychology, Chapter 33* (pp. 1252–1287). Hoboken, NJ: Wiley.
- Gulati, R., & DeSantola, A. (2016). Start-ups that last: How to scale your business. *Harvard Business Review*, 94(3), 54–61.
- Gulick, L., & Urwick, L. (1937). *Papers on the science of administration (early sociology of management and organizations)*. New York, NY: Institute of Public Administration.
- Hackman, J. R. (2002). *Leading teams: Setting the stage for great performances*. Boston, MA: Harvard Business School Press.
- Hargadon, A. B., & Douglas, Y. (2001). When innovations meet institutions: Edison and the design of the electric light. *Administrative Science Quarterly*, 46(3), 476–501.
- Hoogendoorn, S., Oosterbeek, H., & van Praag, M. (2013). The impact of gender diversity on the performance of business teams: Evidence from a field experiment. *Management Science*, 59(7), 1514–1528.
- Kalnins, A. (2018). Multicollinearity: How common factors cause type 1 errors in multivariate regression. *Strategic Management Journal*, 39(8), 2362–2385.
- Kaplan, S., & Vakili, K. (2015). The double-edged sword of recombination in breakthrough innovation. *Strategic Management Journal*, 36(10), 1435–1457.

- Katila, R., Piezunka, H., Reineke, P., & Eisenhardt, K. M. (2021). Big fish vs. big pond? Entrepreneurs, established firms, and antecedents of tie formation. *Academy of Management Journal*. <https://doi.org/10.5465/amj.2018.1197>
- Kaul, A. (2013). Entrepreneurial action, unique assets, and appropriation risk: Firms as a means of appropriating profit from capability creation. *Organization Science*, 24(6), 1765–1781.
- Kent, S. L. (2010). *The ultimate history of video games: From pong to Pokemon—The story behind the craze that touched our lives and changed the world*. New York, NY: Crown.
- Keum, D. D., & See, K. E. (2017). The influence of hierarchy on idea generation and selection in the innovation process. *Organization Science*, 28(4), 653–669.
- Kilduff, G. J., Willer, R., & Anderson, C. (2016). Hierarchy and its discontents: Status disagreement leads to withdrawal of contribution and lower group performance. *Organization Science*, 27(2), 373–390.
- Kutner, M. H., Nachtsheim, C., & Neter, J. (2004). *Applied linear regression models*. New York, NY: McGraw-Hill.
- Lawrence, P. R., & Lorsch, J. W. (1967/1986). *Organization and environment: Managing differentiation and integration*. Boston, MA: Harvard Business School Press.
- Leavitt, H. J. (2005). *Top down: Why hierarchies are here to stay and how to manage them more effectively*. Boston, MA: Harvard Business School Press.
- Lee, E., Lee, J., Lee, J.-H., & Braha, D. (2020). The magic and dark side of hierarchies. Working Paper. Seoul National University.
- Lee, M. Y., & Edmondson, A. C. (2017). Self-managing organizations: Exploring the limits of less-hierarchical organizing. *Research in Organizational Behavior*, 37, 35–58.
- Lee, S., & Csaszar, F. A. (2020). Cognitive and structural antecedents of innovation: A large-sample study. *Strategic Science*, 5(2), 71–97.
- Lee, S., & Kim, J. D. (2021). When do startups scale? Framework, measurement, and large-scale evidence, Working Paper. Wharton School, University of Pennsylvania, Philadelphia, PA.
- Maier, E. (2013). Case study: Valve's flat hierarchy isn't for everyone. Retrieved from <https://inc.com/em-maier/i-drank-the-kool-aid-valve-flat-management.html>.
- March, J. G., & Simon, H. A. (1958/1993). *Organizations*. New York, NY: Wiley.
- McEvily, B., Soda, G., & Tortoriello, M. (2014). More formally: Rediscovering the missing link between formal organization and informal social structure. *Academy of Management Annals*, 8(1), 299–345.
- Meier, S., Stephenson, M., & Perkowsky, P. (2019). Culture of trust and division of labor in nonhierarchical teams. *Strategic Management Journal*, 40(8), 1171–1193.
- Mencher, M. (2002). *Get in the game: Careers in the game industry*. Indianapolis, IN: New Riders.
- Mian, A., & Sufi, A. (2014). What explains the 2007–2009 drop in employment? *Econometrica*, 82(6), 2197–2223.
- Miller, C. C. (2014). Yes, Silicon Valley, sometimes you need more bureaucracy. Retrieved from <https://nytimes.com/2014/05/01/upshot/yes-silicon-valley-there-is-such-a-thing-as-not-enough-bureaucracy.html>.
- Mintzberg, H. (1979). *The structuring of organizations*. Englewood Cliffs, NJ: Prentice-Hall.
- Mollick, E. (2012). People and process, suits and innovators: The role of individuals in firm performance. *Strategic Management Journal*, 33(9), 1001–1015.
- Oster, E. (2019). Unobservable selection and coefficient stability: Theory and evidence. *Journal of Business & Economic Statistics*, 37(2), 187–204.
- Puranam, P. (2014). Managing without authority: Notes on the romance and reality of boss-less firms. Working Paper. INSEAD.
- Puranam, P. (2018). *The microstructure of organizations*. Oxford, UK: Oxford University Press.
- Puranam, P., & Håkonsson, D. (2015). Valve's way. *Journal of Organization Design*, 4(2), 2–4.
- Rajan, R. G., & Wulf, J. (2006). The flattening firm: Evidence from panel data on the changing nature of corporate hierarchies. *Review of Economics and Statistics*, 88(4), 759–773.
- Reitzig, M., & Maciejovsky, B. (2015). Corporate hierarchy and vertical information flow inside the firm—a behavioral view. *Strategic Management Journal*, 36(13), 1979–1999.
- Reitzig, M., & Sorenson, O. (2013). Biases in the selection stage of bottom-up strategy formulation. *Strategic Management Journal*, 34(7), 782–799.
- Rietveld, J., & Eggers, J. P. (2018). Demand heterogeneity in platform markets: Implications for complementors. *Organization Science*, 29(2), 304–322.

- Ruggiero, D. (2017). How game developers define success. Retrieved from <https://youtube.com/watch?v=yVTvw9sKHdA>.
- Santos, F. M., & Eisenhardt, K. M. (2009). Constructing markets and shaping boundaries: Entrepreneurial power in nascent fields. *Academy of Management Journal*, 52(4), 643–671.
- Savage, C. (2015). Ditching flat: How structure helped us move faster. Retrieved from <https://wistia.com/learn/culture/ditching-flat>.
- Saxenian, A. L. (1996). *Regional advantage: Culture and competition in Silicon Valley and route 128*. Cambridge, MA: Harvard University Press.
- Schreier, J. (2012). Why are game developer bonuses based on review scores? Retrieved from <https://kotaku.com/5893595/why-are-game-developer-bonuses-based-on-review-scores>.
- Schreier, J. (2015). One reason Nintendo has such a hard time with new ideas. Retrieved from <https://kotaku.com/one-reason-nintendo-has-such-a-hard-time-with-new-ideas-1680988301>.
- Schreier, J. (2017). *Blood, sweat, and pixels: The triumphant, turbulent stories behind how video games are made*. New York, NY: Harper.
- Shane, S. (2000). Prior knowledge and the discovery of entrepreneurial opportunities. *Organization Science*, 11(4), 448–469.
- Shane, S., & Venkataraman, S. (2000). The promise of entrepreneurship as a field of research. *Academy of Management Review*, 25(1), 217–226.
- Simon, H. A. (1947/1997). *Administrative behavior: A study of decision-making processes in administrative organization*. New York, NY: Free Press.
- Sine, W. D., Mitsuhashi, H., & Kirsch, D. A. (2006). Revisiting Burns and Stalker: Formal structure and new venture performance in emerging economic sectors. *Academy of Management Journal*, 49(1), 121–132.
- Slade Shantz, A. F., Kistruck, G. M., Pacheco, D. F., & Webb, J. W. (2020). How formal and informal hierarchies shape conflict within cooperatives: A field experiment in Ghana. *Academy of Management Journal*, 63(2), 503–529.
- Spaulding, S. (2009). *Team leadership in the game industry*. Boston, MA: Cengage Learning.
- Starr, E., Frake, J., & Agarwal, R. (2019). Mobility constraint externalities. *Organization Science*, 30(5), 961–980.
- Steiner, I. D. (1972). *Group process and productivity*. New York, NY: Academic Press.
- Stinchcombe, A. L. (1965). Social structure and organizations. In J. G. March (Ed.), *Handbook of organizations* (Vol. 7, pp. 142–193). Chicago, IL: Rand McNally.
- Sutton, R. I., & Rao, H. (2014). *Scaling up excellence: Getting to more without settling for less*. New York, NY: Crown Business.
- Tarakci, M., Greer, L. L., & Groenen, P. J. (2016). When does power disparity help or hurt group performance? *Journal of Applied Psychology*, 101(3), 415–429.
- Tost, L. P., Gino, F., & Larrick, R. P. (2013). When power makes others speechless: The negative impact of leader power on team performance. *Academy of Management Journal*, 56(5), 1465–1486.
- Urwick, L. F. (1956). The manager's span of control. *Harvard Business Review*, 34, 39–47.
- Uzzi, B., & Spiro, J. (2005). Collaboration and creativity: The small world problem. *American Journal of Sociology*, 111(2), 447–504.
- van de Ven, A. H., Ganco, M., & Hinings, C. R. B. (2013). Returning to the frontier of contingency theory of organizational and institutional designs. *Academy of Management Annals*, 7(1), 393–440.
- Wasserman, N. (2012). *The Founder's dilemmas: Anticipating and avoiding the pitfalls that can sink a startup*. Princeton, NJ: Princeton University Press.
- Yan, M. & Gilbert, B. (2018). Here's the reason most new console video games cost \$60. Retrieved from <https://businessinsider.com/why-video-games-always-cost-60-dollars-2018-10>.
- Yee, B. (2015). My to-do list: Organizing a producer's work. Retrieved from <https://youtube.com/watch?v=zW8gKpEP-rs>.
- Zhou, Y. M. (2013). Designing for complexity: Using divisions and hierarchy to manage complex tasks. *Organization Science*, 24(2), 339–355.
- Zuckerman, E. W., & Kim, T. (2003). The critical trade-off: Identity assignment and box-office success in the feature film industry. *Industrial and Corporate Change*, 12(1), 27–67.

## SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

**How to cite this article:** Lee, S. (R.) (2021). The myth of the flat start-up: Reconsidering the organizational structure of start-ups. *Strategic Management Journal*, 43(1), 58–92.  
<https://doi.org/10.1002/smj.3333>

## APPENDIX A: MEASURING EXPERIENCE AND BREADTH A

Table A1 illustrates how experience and breadth are measured using a hypothetical start-up with four employees (*A*, *B*, *C*, and *D*). As described in this table, they each have a set of experiences in the following functional domains: “Business,” “Design,” and “Programming.” At the individual level, the amount of experience (i.e., *IndExperience*) was measured by the sum of experience, and the breadth of experience (i.e., *IndBreadth*) by the diversity across all functional domains in terms of  $1 - HHI$ , where *HHI* stands for Herfindahl–Hirschman Index. By averaging, these individual-level characteristics were aggregated to the firm-level (i.e., *Experience* and *Breadth*, respectively).

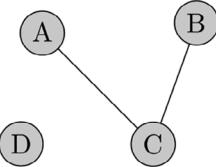
## APPENDIX B: MEASURING SOCIAL CAPITAL B

Table B1 shows an example of measuring social capital. To calculate this variable, I first constructed the collaboration network in a given year (panel (a)), where two individuals have a tie if they collaborated within the previous 5 years (results hold for different time frames; e.g., 3, 7). Then, for each start-up founded in that year, I identified its employees and measured their pairwise shortest path lengths (panel (b)). Because the shortest path length to an isolate (e.g., employee *D*) is infinite, the average shortest path length is infinite for any start-up with at least one “isolate” employee. Since the vast majority of start-ups have at least one isolate who is new to the industry (Grossman, 2003, p. 307; Wasserman, 2012, pp. 229–232), I instead calculated the inverse of pairwise shortest path lengths (panel (c)) and averaged those values (panel

TABLE A1 An illustration of measuring the employees' experience and breadth

Employee	Functional domain			<i>IndExperience</i>	<i>IndBreadth</i>
	Business	Design	Programming		
A	3	2	0	5 = 3 + 2 + 0	$0.48 = 1 - \frac{3^2 + 2^2 + 0^2}{(3+2+0)^2}$
B	1	1	1	3	0.67
C	0	0	1	1	0
D	0	0	0	0	0
			<b>Experience : <math>2.25 = \frac{5+3+1+0}{4}</math></b> <b>Breadth : <math>0.29 = \frac{0.48+0.67+0+0}{4}</math></b>		

**TABLE B1** An illustration of measuring the employees' social capital



(a) Prior collaboration network, where employees A, B, C, and D found a start-up.

	A	B	C	D
A	2	1	$\infty$	
B	2		1	$\infty$
C	1	1		$\infty$
D	$\infty$	$\infty$	$\infty$	

(b) Pairwise shortest path length among the employees.

	A	B	C	D
A	0.5	1	0	
B	0.5		1	0
C	1	1		0
D	0	0	0	

(c) Inverse of the path lengths.

$SocialCapital: 0.42 = \frac{0.5+1+0+0+0}{6}$ 

(d) Average the inverse of the path lengths.

(d)). This average represents how closely the employees are connected through collaboration, and is considered a measure of social capital.

## APPENDIX C: RATIONALES BEHIND THE MODERATORS C

### Breadth of experience

When employees lack breadth by specializing in a functional domain, they tend to narrowly focus on the parts of the information relevant to their domain and neglect the other parts that may be essential to the overall task (Dougherty, 1992, p. 182). This selective processing of information can result in non-overlapping, narrow interpretations which are difficult to integrate into one coherent decision. As employees continue to share these disjoint interpretations and encounter more disagreements and misunderstandings (Garicano & Wu, 2012, p. 6), frustration grows, prompting dysfunctional conflicts (Greer et al., 2018, p. 594). As one entrepreneur illustrates:

With specialization [i.e., less breadth] often comes tunnel-vision. ... People become biased towards their own expertise. It makes decision-making a lot more complicated. More often than not, it's the loudest voice who wins ... even if it doesn't make much sense. (Beaudoin, 2016; comment added)

On the contrary, employees with more breadth can stimulate overlapping, broader interpretations, and flexibly coordinate using various function-specific languages. This cognitive flexibility enables employees to bridge their differences and avoid dysfunctional conflicts.

### Social capital

Dysfunctional conflicts can decrease when employees have strong social relationships developed through prior collaboration. This is because former collaborators develop and share social

capital, such as a common language and tacit routines (Puranam, 2018, pp. 128–144) and a mutual understanding and trust (Stinchcombe, 1965, p. 149). As one founder explains:

[Former collaborators] have history and shared communication. Communication is easy and you know what other people are thinking. You've already agreed on your values. ... [Because of] the trust that you have built up with people, a lot does not have to be said and, in theory, you are comfortable enough to say important things. (Yee, 2015)

Therefore, the more closely connected through collaboration the employees are, the more social capital they share, and thus the less likely that they face dysfunctional conflicts (Meier et al., 2019).