informs
http://pubsonline.informs.org/journal/stsc

STRATEGY SCIENCE
Articles in Advance, pp. 1–24
ISSN 2333-2050 (print), ISSN 2333-2077 (online)

Birds of a Feather … Enforce Social Norms? Interactions among Culture, Norms, and Strategy

Hongyi Li,[a] Eric Van den Steen[b]

[a] University of New South Wales Business School, Kensington, New South Wales 2015, Australia; [b] Harvard Business School, Boston, Massachusetts 02163

Contact: hongyi@hongyi.li, https://orcid.org/0000-0003-3826-6210 (HL); evandensteen@hbs.edu, https://orcid.org/0000-0002-0114-460X (EV)

Received: October 1, 2019
Revised: June 10, 2020; February 16, 2021
Accepted: February 18, 2021
Published Online in Articles in Advance: May 14, 2021

https://doi.org/10.1287/stsc.2021.0129

Abstract. Does culture eat strategy for breakfast? This paper investigates the interactions among corporate culture, norms, and strategy, in order to better understand this issue and related questions. It first shows, through microfoundations, how the forces that drive toward "culture as shared beliefs and preferences" also cause the emergence of social norms (when these beliefs and preferences are unobservable), with people even endogenously enforcing norms that go *against* their own beliefs and preferences. The latter effect may cause a disconnect between the organization's norms and its underlying beliefs and preferences. The paper then makes predictions on the kinds of shared beliefs, preferences, and norms that are likely to emerge, and when they are most likely to emerge. It shows, for example, that social norms are more likely in attractive organizations, for behaviors that have modest personal consequences, and on dimensions along which employees depend on others' choices to a moderate degree. Building on these insights and on the definition of "strategy as core guidance," we finally discuss how corporate culture and strategy interact. We show, in particular, how culture is not just an input into strategy but also a direct substitute for strategy and often even a competitor. The results throw light on ways to deal with the "breakfast" issue.

Supplemental Material: The e-companion is available at https://doi.org/10.1287/stsc.2021.0129.

Keywords: culture • strategy • norms • social norms • shared beliefs and preferences

1. Introduction

This paper studies the close interaction between corporate culture and social norms, and what that implies for strategy—for example, how to think about the claim that "culture eats strategy for breakfast," which paints culture as an almost existential threat to strategy.

To motivate this research, it is useful to first clarify some of the concepts. Most importantly, corporate culture has been defined in myriad—and often somewhat vague—ways. But a common theme throughout these definitions is that corporate culture consists, at its core, of a group's shared beliefs, assumptions, values, or preferences that then drive that group's behaviors (Deal and Kennedy 1982, Schein 1985, Kotter and Heskett 1992). In this paper, we therefore take as our starting point the very simple definition of culture as "shared beliefs and preferences" (Van den Steen 2010a, b), which has been shown to be quite powerful (as we will discuss). An important wrinkle, however—and a core element of our theory—is the fact that such beliefs and preferences cannot be directly observed; they must instead be inferred from behavior. (It is this fact that will give people reason to behave in ways that they would otherwise not.) To avoid potential ambiguity, we must therefore specify this definition of corporate culture somewhat further to "shared *perceived* beliefs and preferences." The concept of social norms, too, has been defined in various ways. Most definitions, however, are a variation on "a socially enforced informal rule of behavior in a group," which we will take as our definition in this paper. We will, finally, define strategy in the "core guidance" sense: as "a small set of core choices to guide the company's managers and employees in their decision making" toward the objective (Van den Steen 2018).

The motivation, then, to study the interaction between corporate culture and social norms is that scholars such as O'Reilly and Chatman (1996) have claimed that social norms make culture into a control mechanism. As we will show, analyzing the interaction between culture and norms in detail not only gives new insight into social norms, into corporate culture itself, and into such claims, but also helps us understand the relationship between corporate culture and strategy, such as that famous claim that culture eats strategy for breakfast. The purpose of this paper is therefore to carefully study this (close) interaction between culture and social norms and its implications by developing microfoundations for culture, norms, and how they interact with each other and with strategy.

At the heart of this paper is then its result that some of the forces that drive toward shared beliefs and preferences also indirectly cause the emergence of social norms (when these beliefs and preferences are unobservable) and may cause people to enforce norms *even if the norm goes against* their own beliefs and preferences. The mechanism behind this result, which has a wide range of implications, builds on a combination of two things: the strong forces toward shared beliefs and preferences (Schein 1985, Chatman 1991, Van den Steen 2010b) and the fact that beliefs and preferences are often unobservable and must therefore be inferred from people's behavior. In particular, this combination generates a pressure for people to act *as if* they hold the "right" beliefs and preferences, in order to fit in. This results in a social norm, with people following an informal "rule" and taking "appropriate" actions that may conflict with their own beliefs and preferences in order to avoid social sanctions. Moreover, when enforcement is observable, people may indeed enforce norms that go *against* their own beliefs and preferences—again in order to act as if they have the right beliefs. Social sanctions thus become part of the norm itself. In this theory, social norm enforcement and sanctions are thus *not* (just) about the actions themselves but often more about what the norm violation reveals about the person's beliefs or preferences. By following and enforcing norms, people show that they are of the right type—or at least not (too much) of the wrong type. We develop this argument in qualitative terms but also illustrate it with a very simple model of a setting where social norms emerge and where people will end up enforcing norms that they do not agree with.[1]

After explaining how social norms emerge in this context, we explore some important implications of this mechanism. A first important implication is that norms make it look "as if" more people share the beliefs/preferences than actually do, that is, norms act as some kind of a multiplier on shared beliefs and preferences. This multiplier effect not only affects the costs and benefits of shared beliefs/preferences, but it also provides an explanation for how cultures can sometimes change quite suddenly and clarifies how people can move between seemingly conflicting cultures. Second, and of particular importance in this context, norms also provide an important channel for culture to function as a control mechanism—with important implications for strategy, to which we return. (Yet, we will also argue that corporate culture exerts control *directly*, beyond its control through norms.) Third—and in the reverse direction—norms may also induce selection and thus create more shared beliefs and preferences. Finally, we also explore how this mechanism may cause a disconnect to develop between an organization's norms and its underlying (i.e., actual) beliefs and preferences, with important implications for culture and culture change.

Apart from these implications for the nature and functioning of culture, the theory also has implications for the prevalence and nature of norms. We show, for example, that norms are more likely in attractive organizations and in settings where people depend on others' choices to an intermediate degree. This last result suggests that the task environment—in particular how much people depend on others' choices—can be used to affect the emergence or extinction of norms.[2] The theory also has implications for the content of norms. It shows, for example, that norms are more likely for behavior that is easy or cheap to adhere to and for work-related behaviors. These comparative statics not only matter for practice, but also provide potential tests of the theory.

The paper then discusses what the theory implies for culture change, where it suggests some new mechanisms, and especially for the interaction between culture and (business) strategy, in particular for the question whether culture eats strategy for breakfast.[3] With respect to the latter, the emergence of norms has the fundamental implication that culture is not just one of the inputs to strategy or part of its execution, as it is traditionally considered: culture is also a direct *substitute* for strategy—in its core function as a guide for decisions (Van den Steen 2017)—and then often *competes* with strategy for control. This may explain why "culture versus strategy" is so high on the agenda for managers. The paper explores this issue in some depth and discusses what its implications are for managing this interaction. The paper, finally, also concisely discusses its implications for the re-emergent debate in some of the field on how to best define corporate culture (Chatman et al. 2019).

1.1. Two Culture Perspectives

Before getting into the literature review, it is helpful to first clarify one important distinction that may otherwise create confusion. That is the distinction between *corporate* culture and culture from an *anthropological* perspective (e.g., Geertz 1973 or Hofstede 1980). Whereas the idea of corporate culture emerged from ideas in anthropology—and that origin leaves an important imprint—it really helps to see the differences between these two. The most important difference is the fact that anthropological culture arises in a setting of a more or less fixed group and thus forms mainly through learning and experience. Corporate culture, on the contrary, emerges in a setting with strong selection and self-selection forces, like hiring and firing. (Getting "fired" from your society is much less likely than getting fired from a company.) Both formal and informal evidence suggests that these selection and self-selection forces tend to overwhelm learning or evolutionary forces in organizations. Moreover, whereas preferences may not change much through learning, they can easily be selected or screened. Finally, the hiring and self-selection in companies often reflects the fact that the organization was created with an explicit objective, in contrast to "life" or society at large. So corporate culture will form faster, will form mainly through selection and screening, will have a more intentional side to it, and will be built more around shared *preferences* than anthropological culture, which is slow, unintentional, and almost exclusively about beliefs and cognition. This thus explains, for example, the increased stress in corporate culture definitions on "preferences" and decreased stress on "learning as the origin," relative to anthropology. (It may also explain some lingering inconsistencies in early (seminal) work on corporate culture, such as Schein (1985), which applied the anthropological notions to a corporate context.) An important implication is that one should be wary about directly extending anthropological insights to a corporate context and should understand the limits of conducting corporate culture research on a fixed group (as that arguably excludes the primary forces toward homogenization in a corporate strategy context). It also suggests that corporate culture will likely tend to have a more "regulative" bent—to use Dimaggio's (1994) distinction—relative to anthropological culture's more "constitutive" bent. Finally, corporate culture and anthropological culture work on different planes. They are not competing concepts but distinct—though related—phenomena with their own rules and properties. The focus in this paper is squarely on corporate culture.

1.2. Literature

This paper is connected to a diverse set of literatures. The literature closest to this paper is obviously the work in organizational economics on "corporate culture as shared beliefs and preferences," such as Van den Steen (2010a, b) and Li (2016) and its antecedents, such as Crémer (1993) and Lazear (1995)—and ultimately, obviously, Schein (1985) and Kotter and Heskett (1992). Because this work will be discussed in more detail in Section 2, we will discuss here some more distant connections. In particular, it is important to observe that there are in fact two quite different views on corporate culture in the economics literature. Apart from the view of "culture as shared beliefs and preferences" that this paper uses, there is also an approach—starting with Kreps (1990) and including work such as Carrillo and Gromb (1999), Weber and Camerer (2003), and Gibbons and Henderson (2012)—that essentially defines corporate culture as an equilibrium, that is, a self-enforcing convention: in situations where aligning actions with those of others matters and it may not be obvious to the participants what action to align on—such as walking on the right-hand side versus on the left-hand side—culture is about creating a common expectation that a given equilibrium will be played, which then makes it optimal for everyone to fulfill this expectation.[4] This relates closely to the idea of "norms as conventions," which is quite different from "social norms": norms as conventions exist only if there are multiple equilibria and are self-enforcing (as driving on the wrong side of the road is a bad idea), whereas social norms are independent of the multiplicity of equilibria and rely on sanctions (possibly internalized). Whereas "culture as equilibrium" is a very interesting perspective, it is clearly different from the perspective in this paper.[5]

Within the broader economics literature, the closest paper is Bernheim (1994), who presents a nice model in which people follow a "norm" to show that their preferences are not too "deviant." But the current paper differs from Bernheim in important ways. First, both the reason for caring about others' preferences and the ultimate "social sanction" are exogenous in Bernheim's paper but endogenized here. Second, and as a consequence, Bernheim has no equivalent to the explicit links with corporate culture as shared beliefs and preferences, to people enforcing norms that go against their preferences, or to the links with strategy.[6] The same differences apply to the work by Bénabou and Tirole (2006, 2011) around the idea that actions can be signals for a person's beliefs or preferences and how that relates to norms.

Some of this paper's results on norms may seem related to the literature on games with a "repeated prisoner's dilemma" and similar games. (See, e.g., Kuhn 2019 for a broad survey.) Whereas some results could indeed be derived in such games, the underlying mechanisms would be quite different and thus also the further implications and comparative statics. In our context, for example—unlike in that literature—knowing the actual beliefs and preferences of early participants gives insight in the likely evolution and eventual outcome. Similarly, the early sorting and selection that plays an important role in our setting is absent from such games. (We will return to this in Subsection 3.3.) Our work also builds on the closely related results of Kreps et al. (1982) on how reputational concerns can support cooperation-like equilibria in such games. But the context and research questions are obviously very different.

A very different literature that this paper also connects to is obviously the literature in organizational behavior on culture and norms. The very definition of culture as "shared beliefs and preferences" was originally intended to capture Schein's (1985) and Kotter and Heskett's (1992) definitions in more formal and concrete terms. Moreover, the early work in organizational economics on shared beliefs and preferences formalized, then eventually went beyond, some of Schein's and others' ideas. Schein (1985), however, does not include norms in his definition of culture and barely touches it in his seminal work. The current paper grew in part out of an effort to try to reconcile O'Reilly and Chatman's (1996) important ideas about "culture as a control mechanism" with culture as shared beliefs and preferences, especially in the context of strategy—where this connection is critical. Whereas O'Reilly and Chatman (1996) take values and social norms essentially as given, see norms as the direct translation of values into behavior, and see "the distinction between them [as] one of emphasis," this paper instead studies in detail the connection between beliefs/preferences on the one hand and norms on the other—in particular, how one may result from a drive for the other. We believe that this paper is the first to study culture from this particular angle. Interestingly, Chatman and O'Reilly's (2016) three dimensions of culture—strength, consensus, and content—emerge very naturally and very clearly from the theory and model in this paper, suggesting that this link may be relevant to their theory.

There is also a related literature that focuses on norms by themselves. This literature is so extensive that Legros and Cislaghi (2020) give an overview of literature reviews, such as Cialdini and Trost (1998), Young (2015), or Bicchieri et al. (2018). Within the wider sociology and psychology literature on social norms, the work that comes closest to this paper is the "social identity" theory of social norms such as Turner et al. (1987). This work, however, does not relate to this paper's main theory; instead, it relates only to a variation on the theory—related to "identity"—that we will discuss later. Willer et al. (2009) informally conjecture that people may enforce norms they do not agree with in order to signal their type and demonstrate this experimentally. But their work is focused solely on the actual norms with no connection to corporate culture and takes these norms and the social sanctions again as exogenous, whereas here this result is derived endogenously as part of a broader theory of culture—on how a drive for one gives indirectly rise to the other—with particular implications for culture change and for strategy. At the same time, their work definitely shows that people indeed behave consistent with these predictions. Other perspectives on norms from the sociology and psychology literature focus on efficiency (Coleman 1990) or on socialization (Parsons 1951) and are therefore far removed from this work. There is also a literature in philosophy that establishes logical conditions for norms to exist, building on game-theory-like ideas. Bicchieri (2006), for example, argues that "conditional preferences"—"I prefer to choose X *if* other people choose X and *if* these others expect that I will choose X"—can lead to norms. In the current paper, people turn out to behave as if their preferences resemble conditional preferences, but these behaviors derive here from more primitive assumptions (and there are also some differences from conditional preferences). Importantly, the direction of causality goes here from norms to these "as if" preferences, rather than the other way around.

Finally, this paper is obviously also connected to the strategy literature on the connection between culture and (business) strategy.[7] The academic literature on this topic is surprisingly limited.[8] Most of this literature, exemplified by Barney (1986), has approached culture as an input to strategy, like a resource or an asset, and then focused on the question how culture can create a competitive advantage. We will discuss that perspective further in Section 5 and explain how this paper suggests a very different perspective (without rejecting the view of culture as a resource or asset). Culture also arises in more practitioner-oriented work (e.g., Groysberg et al. 2018, Levine 2019) and textbook chapters on business strategy implementation (e.g., Grant 2016, Saloner et al. 2001), but these discussions do not go into detail on why exactly culture is such a challenge for strategy, and instead offer fairly general advice to consider culture when developing strategy or to make sure that the strategy and culture fit. The purpose of this paper, on the other hand, is to help develop a richer perspective based on a more detailed understanding of the mechanisms underlying the interaction between culture and strategy.

Overall, the contribution of this paper is to provide a detailed and formal perspective on the close connection between shared beliefs/preferences and social norms and how that furthers our understanding of the interaction between culture and (business) strategy and related questions.

As this paper sits squarely at the intersection of a number of different fields, we will start in Section 2 with a brief overview of earlier work (in economics) on "culture as shared (perceived) beliefs and preferences," including some pointers on how this definition relates to other definitions of culture and some discussion of the distinction between preferences and values. Section 3 presents the basic mechanism by which norms emerge and discusses some of its comparative statics. Sections 4 and 5 discuss the implications of the emergence of norms for, respectively, organizations (including for corporate culture) and for strategy. Section 6 presents the formal model and its analytical results, and Section 7 discusses the implications for the definition of corporate culture. Section 8 concludes.

2. Culture as Shared (Perceived) Beliefs and Preferences

This paper started originally from the research question how culture as shared beliefs and preferences

(Van den Steen 2010a) relates to norms, in particular when extended to settings where beliefs and preferences are not publicly observed. This section discusses this definition of corporate culture and summarizes some of the relevant earlier results, whereas the next section will then link culture to the emergence of social norms.

The (antecedent) definition of culture as "shared beliefs and preferences" was originally developed to capture the approach of Schein (1985) and Kotter and Heskett (1992) in formal-economic terms—linking it to the vast literature that these approaches inspired even though it also productively differs from them.[9] This definition turned out to be not only simple and transparent but also very powerful in terms of its implications and predictions, as we will immediately discuss.[10] The original theory, however, was (for tractability) developed in a complete-information context where beliefs and preferences are observable. Extending it—for this paper—to the more realistic setting where beliefs and preferences are *un*observable required a clarification whether it refers to *actual* versus perceived beliefs/preferences. Because culture in the (academic and practitioner) literature turns out to refer to the perceived beliefs and preferences, this paper formally defines culture as "shared (perceived) beliefs and preferences." Because this definition inherits the attractive properties from its antecedent—as the two coincide under complete information—it is helpful to discuss these properties in some more detail first.[11] The rest of this section is thus about results from the complete-information setting that are relevant in this context.

One important advantage of this simple definition of culture—as shared beliefs and preferences—is that it makes it easy to analyze and understand its nature, properties, and implications. Van den Steen (2010a) used a formal (complete-information) model to show how culture as shared beliefs and preferences leads to more delegation, less monitoring, higher satisfaction, higher motivation, faster coordination, less influence activities, and more communication, but also to less experimentation and less information collection. The reason why culture—so defined—has such a wide range of performance effects is because shared beliefs and preferences resolve agency problems (at the source), given that agency problems find their origin in conflicting beliefs, preferences, and objectives. That paper also discussed in detail the close connection between beliefs and preferences in this context, that is, how they both "direct" decision making in similar ways. Van den Steen (2010b) then studied the *origins* of such culture and formally showed that shared beliefs and preferences have many of the key properties attributed to culture by Schein (1985): they develop through screening, self-sorting, and joint learning; are particularly influenced by the CEO's beliefs and preferences; and persist despite turnover. The paper also made some new predictions, for example, that such culture will be stronger among important employees, in older and more successful firms, and when employees make important decisions (as in professional firms). And the paper reinforced the idea—helpful for understanding the effects of culture more broadly—that it is important to distinguish two dimensions of culture: the *content* of the beliefs/preferences and the *strength* of culture, that is, the degree to which these beliefs/preferences are shared (Kotter and Heskett 1992). Li (2016) showed formally how fast growth may lead to fragmentation of such culture, that the fragmentation may persist even after growth ends, and that it is particularly likely with growth driven by acquisitions.[12] Epstein and Lindner-Pomerantz (2017) investigate under what conditions outlier-cultures can survive, whereas Prasad and Tomaino (2020) show how the strength of culture interacts with the availability of resources to determine economic outcomes. On the empirical side, Campbell (2012) and Abernethy et al. (2015) provide very interesting evidence on the effects of the employee selection aspect of such culture and how that determines outcomes.

Overall, defining culture as "shared beliefs and preferences" turned out to be quite powerful: despite being very simple, it generated lots of implications—some previously associated with culture and some new. But it is important to note that, despite being derived from Schein (1985) and Kotter and Heskett (1992), the definition also differs from theirs. Whereas these differences arose out of the need to translate their views to the very concrete and transparent elements of an economic model (and were originally thought of as a downside), they have in fact proved to be more of a benefit than a disadvantage. Discussing some of these differences therefore gives useful insight into the concept of corporate culture.

One important such difference between this definition and many definitions in the culture literature, including Schein's (1985) and Kotter and Heskett's (1992), is that we define culture here in terms of *preferences* instead of *values*, which differ in ways that go to the heart of culture.[13] Preferences are just likes, dislikes, and attitudes, with varying emotions. Values, on the other hand—used in the plural—are preferences that have an additional moral dimension and a general sense of desirability. The Collins Dictionary, for example, defines values (plural) as "moral principles and beliefs," whereas Parks and Guay (2009)—cited by Chatman and O'Reilly (2016)—define it as "conceptions of the desirable."

So is culture about morally desirable values or may it also encompass more plain and more prosaic beliefs and preferences? The current paper explicitly and purposefully uses "preferences" for two reasons. First, only considering values excludes some mundane but performance-relevant shared beliefs and preferences, such as a tolerance for chaos versus a desire for structure or a preference for teams versus working alone. Second, and closely related, restricting culture to *desirable* values risks distancing it from some of the most difficult management problems—caused by negative cultures like those characterized by complacency, favoritism, or much worse.[14] That being said, we will show why, even though culture is not restricted to values, it will (in equilibrium) tend to consist more of values than of vices or general preferences.

Of course, broadening culture to shared preferences may raise some questions. Can we have a culture of "liking blue," for example? But this question points less to a problem with the definition than to an understudied and important research question: along which dimensions are shared beliefs, shared preferences, or norms most likely? As we will see, our theory makes (falsifiable) predictions on these questions. It suggests, for example, that culture is more likely along dimensions that end up affecting others—thus more likely along "liking structure" than along liking blue. So this theory suggests that such questions should be decided by theory and empirics rather than restricted by the definition. A final concern could be that "shared beliefs and preferences" is too broad and hurts the predictive power of the theory. But this literature has, on the contrary, already produced a very rich set of predictions.

Another difference is that proposed definitions such as Schein's (1985) sometimes include ingredients like the *origin* of the beliefs/values/preferences. Schein (1985), for example, includes in his famous proposed definition that the shared "assumptions" must have been "learned [by the group] as it solved problems" and "taught to new members." But this would imply that shared beliefs that became shared purely by screening or self-selection—which turn out to be more powerful mechanisms—would *not* constitute a culture, which contradicts other results of Schein's (1985) work.[15] So it is not clear whether Schein (1985) was intended as necessary and sufficient conditions, as we would need for a formal definition, or more as a description of a common and easy-to-understand case.[16]

With the definition of culture as shared (perceived) beliefs and preferences clarified, the next section will discuss how this leads to the emergence of social norms. Before diving in, however, let us clarify our use of the terms "norms" and "social norms" more explicitly. We will use the term *norm* to refer to a "set of expected or standard behaviors" and *social norm* to refer to the enactment of such a norm in a group through social enforcement. In particular, we will say that a social norm exists in a group when (1) a large part of the group adheres to a particular norm and (2) at least some people adhere to that norm because they fear social sanctions if they do not. It is also helpful in the ensuing discussion to keep in mind that norms and social norms are about (observable) behaviors, whereas beliefs and preferences are about mental (and thus in principle unobservable) underlying processes.[17]

3. From Shared Beliefs to Social Norms

The core idea of this paper is that the forces that drive toward the emergence of culture—as shared (perceived) beliefs and preferences—also give rise to social norms (when beliefs/preferences are unobservable) with people enforcing norms even if they go against their personal beliefs and preferences. The key animating force in this mechanism is a "*pressure for perceived similarity*," that is, a pressure to *look* similar to others in terms of beliefs and preferences—of which we will immediately discuss the sources. As similarity in beliefs and preferences is not directly observable, it must be inferred from the person's actions and behavior, thus creating a pressure for people to *act* similar to others. This generates a social norm (by our definition): an informal rule of behavior, enforced through social sanctions in whatever form the pressure takes. These sanctions can be some form of punishment or can simply entail withholding benefits that are provided to those with the right (perceived) beliefs and preferences.

But there is more: this mechanism also drives people to enforce norms, even norms that go against their own beliefs and preferences. The reason is that people who really hold the "right" beliefs and preferences will enforce the norm by applying the sanctions. But since these sanctions are often observable, engaging in these sanctions becomes part of the observable behavior of someone with these right beliefs and preferences, to which the pressure for perceived similarity also applies. This creates a pressure to sanction deviators. And *norm enforcement itself becomes part of the norm.*[18]

So let us now return to the pressure for perceived similarity. Where does this come from? We will discuss here in detail one important mechanism that is particularly relevant for organizations, then (in Subsection 3.1) mention briefly two substantial variations on the main mechanism. The main mechanism results from two forces that interact. The first force comes from the fact that in organizations, people often depend on other people's decisions and thus want these other people to make the right decisions. In the face of differences in beliefs, preferences, and objectives, making the right decisions often means making the decisions that the focal actor would have made. The focal actor therefore wants these other people to have the same beliefs and preferences. For example, if someone believes that technical excellence is critical to a company's success, that person will not only care deeply about technical excellence but will also prefer to work with others who believe that technical excellence is key (as they will help to achieve success). Such person will therefore hire others with similar beliefs (or preferences), promote them, and give them resources. An academic example is that someone who cares a lot about research typically wants colleagues who also care about research so that these colleagues put appropriate weight on research when hiring or allocating funds. We will use the term "functional homophily" to refer to this preference for similar others (i.e., "homophily") because it affects you (i.e., "functional"). Functional homophily is a very common and strong force in organizations. Moreover, it goes beyond direct personal interactions and applies to any other person in the organization who makes decisions that affect you. Van den Steen (2010b) already showed, in a setting where beliefs and preferences were assumed to be observable, that this leads directly to shared beliefs and preferences and thus to culture. But as we discussed earlier, beliefs and preferences are often not visible, so that you cannot just select a person who is similar but instead need to select a person who *appears similar*. Therefore, people who *conform* are more likely to be hired, to be promoted, and to get resources to work with, resulting in a pressure for perceived similarity (to those who are in control).

The second force, which amplifies the first, is the fact that when people have *shared* beliefs and preferences, their attempts to influence beliefs and preferences all go in the same direction and reinforce each other, so that homogeneity begets homogeneity. Consider, for example, the effect of functional homophily in an organization where the first two members need to select a third. If these two existing members are opposites (along the relevant dimensions) then their efforts to hire similar others will go in opposite directions and cancel each other out. If, on the other hand, the two existing members are similar along the relevant dimensions, then they will agree on hiring a similar third, so that homogeneity begets more homogeneity. Translated to a setting where beliefs and preferences are not observable, homogeneity leads to more unanimous and stronger norm enforcement and thus to stronger norms. As such, homogeneity in the sense of shared beliefs and preferences *amplifies* the forces that lead to the selection of similar people and to the enforcement of norms. This second force also makes things like "critical mass" important and leads to predictions about the dynamics of culture based on the current degree of sharedness.

3.1. Variations on the Main Mechanism

There are two variations on the main mechanism above that also lead to a pressure for perceived similarity. All these mechanisms have the second force—how homogeneity amplifies—in common. They all differ in the first force, which was functional homophily in the first mechanism. A first alternative force to functional homophily is obviously regular homophily, the well-known and well-documented effect that people simply *like* other people who are similar to themselves and build relations with them (Huston and Levinger 1978, McPherson et al. 2001). In the current context, the key implication of homophily is that if one actor wants to be liked by another, it helps if that actor is perceived to be similar to them, that is, it results in a pressure for perceived similarity. And although there is little you can do about that on visible characteristics, the situation is very different for beliefs and preferences, as before.

A second alternative force is driven by people's need for identity, in the sense of a person's beliefs about themselves (Erikson 1968, Tajfel and Turner 1979, Fiske and Taylor 1991, Akerlof and Kranton 2000). Such identity is derived in part from group membership. I may believe, for example, that I'm a tough person because I belong to a gang that has a reputation for being tough. But the need for identity makes the members of the group care about the group's reputation and therefore about the other group members' actions. If someone in the supposedly-tough group cries, then that may damage the group's image as being tough and thus hurt the members' identity. Members are therefore pressured, through sanctions and other means, to conform to the group's desired image, thus creating a pressure for perceived similarity to the group's desired image. But as each group member conforms, the result is effectively a pressure for perceived similarity among them. Note, however, that a group's desired image may differ from its true nature. Although this causes a misalignment between the group's norms and its underlying beliefs and preferences, it still results in norms. We will discuss another—more important—source of misalignment.

Notice that these different mechanisms can have very different comparative statics, which makes it possible to test and discriminate among them. Identity, for example, suggests stronger norms when group membership is more salient. Functional homophily is stronger when people are affected more by others' decisions.

3.2. A Preference for Differences?

There are obviously also cases where people may prefer to work with others who are *different* from themselves. People may simply prefer diversity, for example, or see a need for trying different things. This observation does not contradict the theory. Instead, it produces some useful insights into this theory of culture and norms.

First, if there are dimensions along which people prefer to work with others who are different then *these dimensions* will simply not become part of the culture or of the norms. In as far as it is predictable along what dimensions that will be, this can give us some predictions about the dimensions along which we are more (or less) likely to see culture and norms. For example, we might hypothesize that people will generally like agreement on work practices (order versus structure, team versus individual, open fights versus consensus, …) but that certain people will like diversity in backgrounds or in experience. Second, a preference for diversity along some dimension often causes a preference for similarity along a related dimension. For example, if I like diversity of backgrounds, then I typically want to work with people who have a similar like for diversity of backgrounds. This would result in a culture with a shared preference for diversity of backgrounds. If I think experimentation is important, then I may want to work with someone who has different beliefs on the most likely solution but similar beliefs on the importance of experimentation. Third, some preferences for differences are not "stable" in a group. If I'm lazy, I may prefer to work with someone who is not lazy, so that that other person does all the work. But that other person will likely *not* prefer to work with me because I'm lazy; instead they also want to work with someone who is not lazy. Overall, a preference for differences does not contradict the theory, but can instead be helpful to make predictions about the types of homogeneity and norms that are more likely.

3.3. Some Comparative Statics

Whereas Sections 4 and 5 will discuss the implications of the main result—on the close connection between shared beliefs/preferences and social norms—respectively for organizations and for strategy. There are, however, some insights from this theory for norms themselves that are worth discussing. (Note that, whereas the formal results of Section 6 clearly support these insights, the illustrative nature of the model means that they do not prove them in full generality.[19] The results thus come in part from qualitative reasoning.)

The following insights are illustrated explicitly by the results in Section 6:

- Norms are more likely in attractive organizations (from which exclusion is thus costly) or when strong sanctions can be credibly applied.
- Norms are more likely when the normative behavior is less costly to the participant. We would thus expect to see a lot of relatively "cheap" norms (such as dress codes, behavior codes, word usage, etc.) and fewer 'expensive' norms (such as houses to buy or type of partner to have).
- Norms are more likely in settings where participants' outcomes depend to a moderate degree on others' decisions and behavior.

This last point deserves some more explanation. Our original intuition when starting this project was that norms only emerge when there is sufficient interdependence: without dependence on others' behavior, there is no reason to enforce norms. But the formal analysis of Section 6 showed, at first to our surprise, that there can also be too much interdependence to sustain norms! The reason is that, when interdependence is very strong, the organization is not attractive to people who disagree with its shared beliefs or preferences. Unafraid of exclusion, such participants see no need to adhere to the norm and will thus violate it, causing the norm to fail. But because norm violators will be subsequently excluded, this norm failure will ultimately result in an organization with fully shared (underlying) beliefs and preferences.[20] This comparative static (on interdependence) is of particular interest as it suggests an interesting path to potentially affect the emergence of norms: forcing or loosening cooperation may affect the emergence (or extinction) of social norms. The result also implies that norms are more likely in settings where people interact a reasonable amount (so they are affected more by others' behavior). This potentially explains why we see norms in organized settings more than in unstructured groups.

The last point also implies some predictions regarding the type of norms and regarding the type of shared beliefs/preferences that are likely to emerge. In particular, norms and shared beliefs/preferences are more likely around behaviors that affect other people in the group. Norms in a company are therefore more likely to be work-related. For example, norms and shared beliefs/preferences are more likely about the optimal level of open conflict or of structure than about liking the color blue. (An exception may be a design studio, where color preferences *do* affect others. Similarly, academic groups can have social norms about which theories or assumptions are more right or "acceptable.") This is also illustrated by two famous early accounts of social norms, the Bank Wiring Room in the Hawthorne Studies (Roethlisberger and Dickson 1939) and the Goldbricking paper (Roy 1952), which are both about settings where employees' pay was affected by others' output.

There is also an important higher-level comparative static, at least from a conceptual perspective. Social norms will only emerge—according to our mechanism—in settings where beliefs and preferences are private information. In particular, in a complete-information setting, participants cannot hide their different beliefs or preferences by conforming to some norm. That removes any incentives to conform and the norm will simply not emerge. Culture *does*, however, emerge in such a complete-information setting, as in Van den Steen (2010b). (With "perceived" and "actual" coinciding, culture is now again simply the shared beliefs and preferences.) Corporate culture can thus exist without norms but norms always go together with culture.

By informally analyzing variations on a model in the style of Section 6, we also conjecture the following further insights:[21]

- Norms are more likely for behavior that is socially desirable or at least acceptable, that is, norms are more likely for values than for vices or than for regular preferences: if the behavior is socially desirable, then following the norm has other benefits beyond avoiding social sanctions and also makes enforcing the norm more rewarding.[22,23]
- Norms are more likely in settings where there is a common future and where that future matters more: if two people are going to interact a lot and that interaction matters a lot, then they care more about each other's behavior and thus also about each other's beliefs and preferences; moreover, with a common future, social sanctions also have more bite.
- Vulnerable members of the organization—those who are easy to sanction or who bear strong consequences from the sanctions—are more likely to adhere to norms because it is very important for them to avoid the sanctions. Invulnerable members are, by the same logic, less likely to adhere.
- Members whose beliefs and preferences are common knowledge or beyond doubt are less likely to adhere to norms (that go against their preferences or beliefs), as they do not need to prove their type.
- A norm requires a critical mass of participants who are perceived to hold the norm-supporting beliefs and preferences: without such critical mass, the social sanctions are not strong enough to make the norm worth obeying.[24]
- Early members of an organization have more influence on norms than later members: norms form almost automatically around early members' beliefs and preferences, whereas later members can only affect the norm by overthrowing it, which is difficult.

It may at first be a bit surprising that we argued earlier that culture should be defined in terms of preferences rather than values, but now show that norms are nonetheless more likely for values than for vices or regular preferences. There are, however, big differences between including a property in a definition versus showing that it tends to hold in equilibrium. First, including the property in the definition simply excludes the opposite property from ever holding and that is problematic in this case: some cultures have shared vices and such vicious cultures are important to understand because of their destructive potential. Second, deriving a property is more powerful than simply imposing it as it becomes a *prediction* instead of an assumption, which makes the theory not only more powerful but also more testable.

One central characteristic that we have not fully addressed up to this point is the role of the strength of beliefs and preferences.[25] Part of the issue is that one of its key effects is partially outside the reach of this paper: strongly held beliefs and preferences will lead to more sorting and thus to more homogeneity in underlying shared beliefs and preferences. This is not only because people are less willing to follow norms that go against strongly held beliefs and preferences (per the earlier result that norms are more likely for "cheap" behavior) but also because people tolerate less deviation by others, thus either pushing them out of the organization or leaving themselves. In terms of actual norms—for fixed beliefs and preferences—people with strongly held beliefs and preferences are less willing to follow norms (because such behavior is not cheap) but also, as a somewhat countervailing force absent from the formal analysis in this paper, more willing to enforce such norms. Overall, strongly held underlying beliefs/preferences will lead to more homogeneity in (actual) beliefs/preferences but has mixed effects on norms for those who strongly hold these beliefs/preferences.

The results of Srivastava et al. (2018) are interesting in the context of all these predictions—if we take a somewhat broad view of their interpretation. Their paper studies how enculturation determines success by looking at how people's email styles come to resemble that of their colleagues. If we interpret adhering to an email style as adhering to a norm, then we can see how the results compare with the theory. Fast adoption of norms—which may measure how close norms are to personal beliefs and preferences—predicts high retention, which would thus be consistent. Not adopting norms predicts getting "eliminated" from the organization. And a decrease of adherence to norms after having adopted norms previously predicts voluntary exit, which seems to capture the effect that some feel that adherence to norms is too costly and leave or that their desire to be in the organization has decreased for other reasons and now they do not feel a need to adhere to norms any more. Finally, adopting an email style is a relatively cheap behavior, along which we thus may expect norms to develop.

3.4. Misalignment between Norms and Underlying Beliefs/Preferences

Occasionally, the mechanisms discussed can lead to a misalignment between the norms of an organization and its underlying beliefs and preferences. In fact, we will immediately explain why it is (in principle) possible that the organization enforces norms that go directly against *everyone's* shared beliefs and preferences—and even when everyone is aware of that misalignment. Although such an extreme case is unlikely in practice and therefore more for illustrative purposes, the real possibility of more moderate misalignment has important implications for the broader theory of culture and for practice. Some of these implications will come up in the next sections of the paper. This section focuses on the mechanism behind the misalignment.

So how can such misalignment happen? The fundamental issue is that participants follow norms because of how they *think* others will react to their actions. At the same time, participants may over time lose track of who the people are with beliefs or preferences that really support the norm, for example because they lost track of who actually started the norms versus who just went along or because they joined after the norm was already established. If there is turnover and participants cannot reliably distinguish who has beliefs/preferences that really support the norm, then at some point all the people with beliefs/preferences that gave rise to the norm may have left, but the norm is still enforced because people are not fully aware of that turnover; and in the face of uncertainty, people may then find it safer to follow and enforce the norms.[26] In fact, even if a person is sure that all the original "instigators" have left, that person may keep following and enforcing a norm that goes against her own beliefs and preferences because she is not sure whether *others* know that fact and because she is afraid that these others may sanction her if she fails to follow or enforce the norm.[27,28]

3.5. A Note on Assumptions

It is probably useful to discuss a few assumptions implicit in the theory that are likely taken for granted by some readers but that are either nonobvious or raise questions for others. This is especially true as this work sits at the intersection of multiple literatures and in an area where these assumptions may really matter.

One such assumption is that people's preferences do not change. Even though it is well-understood that preferences might be malleable to some degree, we do not include that possibility in the analysis or in our more qualitative discussions, for a number of reasons. First, assuming fixed preferences is a conservative assumption, in the sense that it would become easier to explain the results if we dropped it (and let everyone's beliefs and preferences just change to, say, the majority's). Second, explaining all norms by preference change does not ring true to us and conflicts with our own experience. So we still need an explanation of norms for situations where preferences do not change, and thus we need a theory like this one. Finally, we believe that the theory would hold up even if reasonable degrees and types of malleability are included, but including these assumptions goes beyond the research intended here.

A second assumption is that people are rational, including that their beliefs are (correctly) updated as new information becomes available or when new information can be inferred. We understand the concerns about unbounded rationality, but we see rationality simply as an approximating assumption that captures "logical" behavior and that works sufficiently well in most settings. Moreover, its analytical power and its ability to develop theories "out of sample" makes it a powerful tool. We believe that the test is ultimately in the pudding: does the theory that emerged from assuming rationality pass the "reasonability" test in terms of what it demands from rationality? We strongly believe that this is the case here.

4. Implications of Norms Driven by Shared Beliefs and Preferences

With the relationship between corporate culture (as shared (perceived) beliefs and preferences) and social norms established and with basic comparative statics for norms themselves derived, we now turn to the *implications* of such social norms and culture. This section looks at general implications for organizations, whereas Section 5 will focus on the implications for strategy in particular.

4.1. Norms as Multipliers

A first important effect of such social norms is that the underlying shared beliefs/preferences may seem more wide-spread than they actually are. In particular, because some people who do not share the beliefs/preferences act as if they do, it will be *as if* they really share these beliefs/preferences. In other words, norms work as a *multiplier* on shared (perceived) beliefs and preferences.

This matters in two broad ways. First, it matters because it helps our understanding of how organizational culture really works. It explains, for example, how people can move between organizations with very different, even conflicting, cultures and nevertheless "fit in" everywhere: if the person has relatively weak beliefs/preferences along the relevant dimensions, they can function in each organization by adhering to the prevailing norms. This effect also matters directly for culture change as it implies that there will often be a (considerable) group of members of the organization who follow the norms but are very open to a different culture. Recognizing a conflict between social norms and underlying (actual) beliefs and preferences may be key to finding leverage to change organizational culture and norms.[29] Finally, "norms as multipliers" also has empirical implications when testing the effects of culture. As already pointed out, this can be an interesting angle to interpret work such as Goldberg et al. (2016) or Srivastava et al. (2018) who derive measures of culture from people's similar style in emails. In as far as the styles are reflective of norms, the ability to observe norm adoption provides some interesting angles.

A second broad way in which "norms as multipliers" matters is because it affects how the cost and benefits of homogeneity, as identified by Van den Steen (2010a), will play out. This norms as multipliers effect has two opposing effects on the costs and benefits of homogeneity. On the one hand, "norms as multipliers" can potentially extend these costs and benefits to employees who act as if they hold the shared beliefs and preferences by following the norm. For some of the effect, however, this will only be the case if the norms are enforced independent of the agency relation under consideration. For example, a principal will feel comfortable delegating to an agent who may not share his or her beliefs *only if* the actions are observed by others who will enforce the norms. On the other hand, there is also a reverse effect: "norms as multipliers" can weaken the benefits and costs of homogeneity if the principal is unsure whether the agent will obey the norm. For example, when faced with some probability of "wrong" actions and, due to the norm, unable to tell the difference between those who really share the beliefs and preferences versus those who just go along with the norms, the principal may decide to forgo delegation or reduced monitoring altogether.[30]

4.2. Norms as a Control Mechanism

A second and closely related effect of social norms is to provide an important channel for culture to function as a control mechanism, as proposed by O'Reilly and Chatman (1996). Given that control has long been viewed as one of the core functions of management and organization (e.g., Fayol 1916), the fact that culture can exert control (through norms) is obviously significant in and of itself. And whereas there exist a lot of formal control mechanisms—from direct monitoring and (subjective) evaluations to incentives—culture and norms as control mechanisms have distinct advantages, for example because their reach is very different and because they may come across as less openly controlling. (Obviously, culture is no panacea as it is itself difficult to control and may thus drive behavior in counter-productive ways.)

Norms are obviously not the only mechanism through which corporate culture as shared beliefs and preferences exerts control over an organization's actions. Shaping shared (actual) beliefs and preferences serves as a *direct* form of control—after all, people with the "right" beliefs and preferences will take the right actions, even in the absence of norms. There is also an interesting relation between these two forms of control: attempts at culture change may fail to really change (actual) beliefs and preferences, yet, by influencing the emergent norms, achieve the ultimate goal of exerting control over actions.

"Norms as a control mechanism" may seem to be just a different way to look at "norms as a multiplier on shared beliefs and preferences" in that norms make people behave as if they hold the shared beliefs and preferences. But the two perspectives ultimately focus on very different aspects of the same broad phenomenon, which reflect the two fundamental dimensions of culture identified by Kotter and Heskett (1992): the degree to which beliefs and preferences are shared versus the content of these beliefs and preferences. Norms as multipliers is about the degree to which beliefs and preferences are (or seem) shared—and is in principle independent of the content of these beliefs and preferences. "Norms as control" is all about the content of the actions and of the beliefs and preferences that they reflect, and is in principle independent of whether these are shared.

This effect also clarifies the following seemingly paradoxical feature of culture as a control mechanism. Norms as control have the most bite for people who do not share the underlying beliefs and preferences; so if norms were purely about people following their beliefs and preferences, then norms would stop effecting control. "Norms as control" is thus most relevant in settings with *partially* shared beliefs and preferences.

4.3. Norms as Signals

The theory also implies that norms are often not necessarily about the behavior itself, but about what that behavior reveals about the person, and thus about other behaviors that the person may engage in. Norms then act as signals for underlying beliefs or preferences. For example, a norm of wearing ties is presumably not driven by a deeply held belief that "wearing ties is important" but, for example, by a shared belief that formal dress reflects respect for customers and that respect for customers is important. In the model of the paper, norm enforcement is indeed largely about what norm violations reveal about the person's preferences.

This "norms as signals" observation has important implications. First, inferring underlying beliefs and preferences from norms may be difficult, as famously observed by Schein (1985, p. 20–21). People may wear ties because of the underlying belief that it reflects respect or because of an underlying preference for "elegance." But, second, understanding the true driver is fundamental when it comes to changing norms. If some norm is about signaling a particular belief, then the norm can sometimes be changed by providing an acceptable alternative to signal that belief or by changing the actual belief itself. But both of these are possible only if you understand what that belief or preference really is. Third, it also raises the question whether some types of norms should be considered as just an expression of culture, rather than a real part of culture. For example, is culture about wearing ties or about showing that you care about customers? We return to that question.

This "norms as signals" also connects to a very different role of norms: norms give new participants a window into the underlying shared beliefs and preferences of the organization. With respect to the distinction between descriptive versus injunctive norms (Deutsch and Gerard 1955), note that in this theory descriptive and injunctive norms essentially coincide: people learn how to behave from observing what others do, but what they implicitly learn is the behavior that satisfies the injunctive norms; their reason for following descriptive norms is precisely because the descriptive norms are also injunctive norms.

4.4. Norms and Self-Selection: From Norms Back to Shared Beliefs and Preferences

Up to this point, we have looked at how shared (perceived) beliefs/preferences lead to social norms. But there is also an important effect in the reverse direction, that is, from social norms to shared beliefs/preferences. In particular, norms force people to behave as if they hold the underlying shared beliefs and preferences. Doing so is fine for people who do share these beliefs and preferences but is often uncomfortable for people with different beliefs and preferences. This is especially true for those who hold such conflicting beliefs and preferences strongly. Such people may leave the organization and they are more likely to do so if their own beliefs and preferences conflict more or if they hold these conflicting beliefs and preferences more strongly. It follows that exactly those people who may weaken the culture (and violate the social norm) most are also most likely to leave, thus reinforcing the culture in a virtuous cycle. This is thus another way in which norms end up strengthening culture, but now in a dynamic sense: the norms, generated by culture as shared beliefs/preferences, lead to self-sorting that further spreads that same culture.

This mechanism can also be used in a very different way to change culture, as we discuss now.

4.5. Culture Change

This paper's perspective on the relation between corporate culture (as shared perceived beliefs/preferences) and social norms has important implications for ways to change an organization's culture and its norms. For this discussion, it is important to observe up front that it is very difficult to change an organization's underlying (actual) beliefs/preferences and that such a change may require drastic measures. Note that there are two mechanisms through which such change can happen (Schein 1985, Chatman 1991, Van den Steen 2010b). By far the most effective mechanism is turnover: replace people who hold the "wrong" beliefs and preferences with people who hold the right beliefs and preferences. Considerable turnover, both forced and voluntary, is indeed typical for settings where organizations have changed their culture. The second mechanism is to try to modify the actual beliefs and preferences that people hold. Although there are some exceptional circumstances where this may happen relatively quickly,[31] this is almost always a slow and difficult process, if it happens at all. By and large, then, changing beliefs and preferences seems to happen mainly through turnover, both forced and voluntary, which is not only painful but also very costly, including the costs of lost experience and specific human capital, and human cost.

Building on these observations, the paper first of all implies a fundamental cautionary tale about trying to change norms (i.e., behaviors) without tackling the underlying (actual) shared beliefs and preferences. In particular, doing so may be futile, as the organization will likely revert to the old norms once the pressure is removed. (Personal observations suggest that such "reversion" is quite common.) As will be discussed, temporarily imposing new norms *may* lead to a real change in underlying beliefs and preferences, and then have a long-term impact, but only under certain conditions. So one should always be wary about the sustainability of apparent short-term changes in norms.

But the paper does also suggest a number of concrete leverage points that can be used to *facilitate culture change* (in the sense of a long-term change in shared (perceived) beliefs and preferences) or a permanent change in norms. The first and most important lever is the fact that there typically is a, potentially large, group of "conformers": people who have relatively weak beliefs or preferences and who just go along with the prevailing norms. In other words, their behaviors simply conform to the norm to mask any potential divergences in their beliefs or preferences from the organization's culture. Such conformers can often switch quite easily from one culture regime to another, simply by adapting their behavior and thus their *perceived* beliefs or preferences. This also explains why organizations can really change their culture without having 100% turnover. (Note that in the model of Section 6, employees who adapt to a changing culture will not necessarily be perceived as having the wrong beliefs and preferences, but simply as having weak beliefs and preferences.) To assess the likelihood of success with a desired culture change, it is therefore helpful to have some assessment of the share of people who hold the current shared beliefs and preferences strongly versus those who just conform. Note that this also predicts that employees hired after the culture change are more likely to fit with the new culture than those who remained, consistent with the empirical findings of Campbell (2012).

A second important leverage point for culture change—closely related to the first—is the potential disconnect, discussed in Subsection 3.4, between *perceived* beliefs and preferences and *actual* beliefs and preferences. For culture, it is the perceived beliefs and preferences that matter. If there is a perception that most employees, especially powerful employees, share a specific set of beliefs and preferences, then that will sustain norms and enforcement of norms that are consistent with these perceived beliefs and preferences. But in case of a misalignment between perceived and actual beliefs, revealing the actual beliefs and preferences, making them common knowledge, may drastically alter the forces that sustain norms and may lead to rapid culture change. When culture change is needed, it is thus again helpful to try to discover the actual versus perceived beliefs and preferences, as that may provide a powerful lever for culture change.

A third leverage point is the observation that norms are ultimately the (superficial) expressions of deeper-held beliefs and preferences. This has two implications. First, one set of beliefs and preferences can manifest itself in very different, even conflicting, norms. For example, really caring about patients can be expressed in a norm to prescribe patients the medications they ask for or in a norm to just do what is best for the patient, independent of how the patient feels about that.[32] It follows that norm changes may be possible without necessarily changing the underlying shared (actual) beliefs and preferences. Second, and relatedly, norms differ in how strongly they align with the underlying beliefs and preferences, making some norms easier to change than others. In both cases, developing a clear understanding of the shared beliefs and preferences that drive the norms is very helpful to assess the potential to change some of these norms.

A fourth lever is the use of "norms as selection" as a tool to facilitate culture change, as suggested in Subsection 4.4. In particular, temporarily forcing different norms may lead to self-sorting toward people who hold beliefs and preferences consistent with these enforced norms. But this works only under certain conditions. The issue is that unless there is a change in the underlying beliefs/preferences that drive the norms, the organization will revert to its old norms as soon as the forcing ends. If participants realize that, they have limited reason to sort: they can just live through the temporary discomfort and be back to their old ways. So a *vicious/virtuous cycle* emerges: the prospect of no change in underlying beliefs and preferences removes the incentives to self-sort and thus ensures that there will be no change. Three things are therefore important to make this work. The first is a credible prospect of a substantial period with the new norms, enforced or otherwise, for example by an organization change that creates some commitment to the new norms.[33,34] The second is to encourage turnover that furthers the change in underlying beliefs and preferences; for example, by facilitating employees' voluntary departure, by explicitly using forced turnover as a response to norm violations, and by making the new norms very visible in recruiting. The third is to be very cognizant of the vicious or virtuous cycle by creating critical mass and being proactive about managing perceptions, for example by making any progress very visible.

As a final lever, this paper suggests that a change in the task environment that creates or breaks interdependence along certain task dimensions or along certain social dimensions may be helpful to change culture and norms. In particular, such changes in the task environment may alter the forces that sustain the norms and that lead to shared beliefs and preferences. If, for example, employees are forced to cooperate more and work more in teams, they will start to care more about the beliefs and preferences of their teammates, setting in motion a process toward more shared beliefs and preferences and toward social norms, and the other way around.

Overall, whereas wholesale culture change is typically a very difficult and slow process, this paper points out that (1) there are less-common cases (that can be identified) where wholesale culture change is a lot easier, and that (2) sometimes changes in part of the culture (e.g., part of the norms) may be much easier and may achieve the intended purpose.

4.6. Effects of Shared Beliefs and Preferences Themselves (Beyond Norms)

All these effects of norms raise the important question whether *all* effects of shared beliefs and preferences may in fact go through norms.[35] The answer, however, is unequivocally that shared beliefs and preferences matter independent of norms, as a lot of the effects of shared beliefs and preferences do not depend on norms. This is obvious from Van den Steen (2010a): that paper does not have any social norms (as it does not have private information about beliefs or preferences), yet shared beliefs and preferences have a wide range of effects.

5. Culture, Norms, and Strategy: Does Culture Eat Strategy for Breakfast?

For strategy, the most important challenge related to culture is the claim that culture eats strategy for breakfast, that is, the claim that in a battle between strategy and culture, culture (somehow) wins.[36] This section discusses this issue, what it says about the ways that culture and strategy interact, and how to potentially deal with it.

Traditionally, the strategy literature has viewed culture as an input—in the sense of a resource or an asset—that can help or hurt the strategy (e.g., Barney 1986). A frugal culture, for example, supports a low-cost strategy but may hurt a high-end strategy. Hence, a company needs a culture that fits its strategy, just like it needs a brand that fits its strategy. In this traditional view, culture is one of many inputs to strategy and following a strategy with the wrong culture or with the wrong brand will lead to failure. In that sense, the wrong culture may "eat" your strategy, just like the wrong brand or the wrong know-how. And the solution, in this view, is to change and find the right culture or brand.

This perspective misses, however, a critical aspect of the relationship between culture and strategy. Executives do not typically ask "Which is more important, strategy or brand?", but they *do* ask "Which is most important, strategy or culture?" Somehow the relationship between culture and strategy thus seems more existential than that between brand and strategy. The theory in this paper does actually explain the above observation. In particular, it shows that culture is not just one of many inputs into strategy or a component of strategy execution, though it is that, too; culture is additionally a direct *substitute* for strategy and often competes with it for control. In particular, the "core guidance" definition of strategy as "a small set of core choices to guide the company's managers and employees in their decision making" toward the objective (Van den Steen 2018) makes clear that the role of strategy is to guide the company's decisions. Strategy is thus essentially a control mechanism to achieve coordination and direction. But culture is a control mechanism too, through norms and through the direct influence of shared beliefs and preferences. This makes strategy and culture (as shared beliefs and preferences) indeed *substitutes and potential competitors*. To return now to the motivating observation: because brand is seen as part of strategy and is not some competing control mechanism, juxtaposing "brand versus strategy" makes little sense;[37] culture, on the other hand, is seen as a potential alternative or competitor to strategy, and then the juxtaposition makes perfect sense. To say it in a different way, culture may pose an existential challenge to strategy (whereas brand does not).

Strategy unfortunately has a weak hand when it needs to compete with culture. The issue is that strategy needs credibility (in the form of reliability) to be effective (Van den Steen 2017): employees take strategy as guidance only if they believe that others will follow it. This often relies on a virtuous cycle where people follow the strategy (only) because they believe that others will follow it, which is fragile. Social norms, on the other hand, derive their credibility from stable beliefs and preferences. A conflict between culture and strategy therefore often turns into a vicious cycle where culture wins. As a consequence, culture can interfere with the effectiveness of strategy on an almost existential level and devour strategy for breakfast. The traditional response to this issue has been to make strategy adapt to culture. But this response creates important problems of its own. In particular, as a control mechanism, strategy does have distinct and important advantages over culture. Strategy, when done well, is more under the direct control of management and much more malleable than culture; it can reach where culture cannot; and it can be made much more clear and unambiguous. Simply adapting strategy to culture forgoes many of strategy's advantages and extends culture's inertia into strategic inertia.

Some of these issues can be illustrated with the experience of the Ford Motor Company. The phrase "culture eats strategy for breakfast" was in fact popularized by Mark Fields, a former Ford CEO. Ford had indeed had a widely publicized experience with this "strategy for breakfast" issue at a time when Fields was climbing the ranks. In particular, when Bill Ford (great-grandson of Henry Ford) became Chairman and CEO of the Ford Motor Company in 2001, Ford had a strong culture of being a "real car company"—exemplified by its "built Ford tough" tagline. The culture had everyone in the company focused on massive gas-guzzling SUVs. Bill Ford decided to redirect Ford's strategy toward more environmentally friendly cars, going against that "real car culture." Despite having his name on the building and his family's voting shares behind him, however, Bill Ford's intended strategy never got beyond concept vehicles and long-term plans as he failed to get buy-in from almost any part of the company. As Ford himself later reflected, "I couldn't find anyone else in the company at a remotely high level who shared where I thought this company had to go." And with shared beliefs and preferences in direct opposition to the intended strategy, everyone understood that they could safely disregard that strategic push.[38] When Mark Fields became CEO, he really pushed the idea that culture eats strategy for breakfast, sticking closely to the traditional culture (while making more symbolic moves on the environmental front). But simply adapting strategy to culture leads to inertia. Indeed, Fields was ultimately let go as CEO *exactly because* he failed to produce a sufficiently radical strategy, especially with respect to electric vehicles.

The theory of this paper does suggest a number of complementary approaches to this issue. The first approach is based on the observation that the core of the problem is commitment: culture creates commitments that may conflict with strategy. This raises the question what other commitments could be deployed to counter this effect of culture. Whereas leadership—by the board, the CEO, or the management team—will obviously play an important role in this, that is not the only possible path. There are many moves and actions that can create a commitment. Think, for example, of Cortez' burning of his ships. It is in this respect interesting to note that Fields' removal sent a clear signal to the Ford organization that the board was committed to a more radical path, making it easier for his successor to make more credible commitments in this direction.

The second approach starts from the observation that the solution should not be thought of as a black-or-white choice between "strategy adapting to culture" versus "culture adapting to strategy." In contrast to how it is often thought of, culture is typically not an aligned monolith but instead a complex patchwork of different (at times conflicting) beliefs and preferences and norms that follow from them. Changing one aspect of culture, when well targeted, may be all that is needed and will be much more feasible than a wholesale change of culture. The optimal approach is therefore typically a mix of adapting to some norms while either countering or modifying others.

Third, the earlier discussions of culture change also provided pointers to potential ways to modify culture and norms: leveraging the existence of conformers, leveraging potential disconnects between perceived and actual beliefs and preferences, or channeling beliefs and preferences into norms that are more compatible with the strategy. For example, caring about students may be expressed in a norm to make classes into easy entertainment or in a norm to make classes difficult and challenging. Two very different strategies can thus be consistent with the underlying beliefs and preferences, as long as these can be channeled into the right norm. This is indeed one of the approaches that Ford took after Fields' departure: by building its EV strategy around the Mustang—one of its most revered brands—it is trying to make EVs part of that real car company culture. GM seems to be following a similar approach by bringing back the Hummer as an EV.

Finally, strategy will slowly shape culture, as it will lead to self-sorting toward people who align with the strategy as discussed in Subsection 4.4. So a strategy change, when actually implemented, should become easier over time as the culture adjusts.

Overall, the theory makes a strong case for the importance of developing a detailed understanding of culture and norms and their "structure." Small conflicts, contradictions, and inconsistencies may be helpful leverage points for modifying the relevant dimensions of culture so as to make it carry the right strategy. Moreover, it makes clear that strategy change, culture change, and leadership should be approached in an integrated manner. Although bad strategy may get eaten by culture, great strategy knows how to be carried by culture (and make culture do the hard work).

6. A Formal Model

This section presents a very simple model that shows the forces at the heart of this paper and illustrates some of its core insights. The model captures an organization where people may disagree on the optimal course of action (because of different preferences or beliefs); each person is affected by the choices of all others in the group; and each person can (over multiple periods) try to force others out. Because of the decision externalities, each person prefers to work with others who share her beliefs or preferences and will try to remove those with different views or preferences. Beliefs and preferences, however, are not observable and must be inferred from behaviors. The analysis then shows that in such a setting—after an early shake-out that establishes shared preferences or beliefs—social norms indeed develop, with people even enforcing norms against their own preferences (to be seen as being of the right type). The model also derives comparative statics in line with the earlier-discussed results.

Whereas this formal analysis provides solid support for the paper, the model is too specific—due to its limited focus—to provide a complete formalization of the theory (which thus requires further research). Instead, the purpose of the model and analysis is threefold. First, the model and analysis illustrate the mechanism and make very transparent exactly how the logic works. Second, the model and analysis work as a basic check on the more qualitative argument. Third, and closely related to the second, the model also pushes the boundaries of the original argument. In this particular case, the analysis actually revealed an important omission in the original qualitative arguments that then led to new insights, as we will discuss. (The online appendix has a more formal and precise statement of the model and result and contains the proofs.)

The formal model captures a stylized setting of an organization with N employee positions, where we assume N = 5 for simplicity.[39] The employee in position n needs to choose a public course of action $a_n \in \{L, R\}$ and a private action $b_n \in \{L, R\}$, with a choice between Left (L) or Right (R). The public action is taken up front and visible to all, like setting an official policy, whereas the private action is taken at the end and private information—as in the case of a last-minute switch or some deviation from policy in the field. Employees are randomly selected from an infinite pool of applicants who differ in their preference for L versus R (or, equivalently, who disagree, in the sense of differing priors, whether L or R is the better course of action). In particular, half the pool prefers R and half prefers L. Moreover, both among those who prefer R and among those who prefer L, a proportion $1-\gamma$ are what we call "regular" types, labeled r and l, while a proportion γ are "strong" types, labeled R and L. We assume (for analytical reasons) γ to be vanishingly small.[40] Let $\tau_n \in \{L, R\}$ denote the action preference of the employee in position n. These preferences are private information, but everyone knows the composition of the pool.

The game consists of two parts. (For a formal timeline, see the online appendix.) In the first part, employees are hired and choose their public action, but can also try to eliminate other employees. They will, in fact, get (at most) three opportunities, in three rounds, to eliminate others. Formally, in the first round, after N employees are selected randomly from the pool, they simultaneously choose their public actions from {L, R}. After observing each other's actions, all employees simultaneously and publicly "vote" which of the other employees they would like to "eliminate." Any employee who gets a majority of votes ($> N/2$) against her is immediately eliminated and replaced by a new employee randomly selected from the pool. All replacements happen simultaneously. At the moment an employee is eliminated, all remaining employees incur a cost ϵ for each employee eliminated. (This can be interpreted as the personal cost of turnover, e.g., training, taking over duties, or separation costs, or some empathy with those who are removed.) If no one was replaced, the game proceeds to the second part. But if anyone was replaced, then the game moves to the second round. In that second round, all employees again have the opportunity to choose (or re-choose) their public actions, followed by a second elimination vote (with immediate replacements). ("Old" employees can thus switch their earlier action choice, if they want to.) As before, the game proceeds to the second part if no one was replaced, but (now for the last time) if someone was replaced, there is a third and final round. Employees again have a chance to choose or re-choose their public actions, followed by an elimination vote (with immediate replacements). Now the game proceeds to the second part no matter what. (If any employee was eliminated and replaced in the last round, her public action stands.) Finally, in the second part, employees choose their private actions and their final payoffs are realized. A regular employee in position n gets utility

$$U_n = U + (VJ_{a_n=\tau_n} + vJ_{b_n=\tau_n}) + \sum_{m \neq n}(WJ_{a_m=\tau_n} + wJ_{b_m=\tau_n}),$$

where $J_X = 1$ if condition X is satisfied and equals –1 if it is not. Here, $U > 0$ is some benefit from being part of the organization, $V, v > 0$ are the benefits/costs from the employee's own choices, and $W, w > 0$ are benefits/costs that an employee enjoys/suffers from the choices of others in the organization. An employee's utility when not being part of the organization, that is, the outside option, is normalized to zero (though any costs from eliminating others (ϵ) are sunk). We will finally assume that strong-type employees (types L and R) choose mechanically: they always choose the action of their type $a_n = b_n = \tau_n$ and vote against anyone who they believe prefers the opposite action.[41]

The key point of the analysis now is to show that this setting exhibits a very natural equilibrium, which we call the "social norm equilibrium," in which either an L-norm or an R-norm emerges (in a sense that we will make precise), which is upheld and enforced even by employees who prefer the opposite course of action. This result will also allow us to derive some of the comparative statics we discussed.

To explain what the equilibrium looks like, it is helpful to start with a few observations about the set of possible outcomes. These are unfortunately somewhat technical but are necessary for a precise explanation. Most of these observations are about the beliefs that employees hold about each other. First, because all choices before the end-round are public and everyone starts from the same prior beliefs, everyone will share the same belief about a particular employee's type (with the exception, obviously, of that employee, as she knows her own type). Call this the employee's "perceived type." Second, in the equilibrium that we consider, only a few beliefs are possible as perceived types. Most importantly, because the equilibrium is in pure strategies, if the perceived type of a player puts strictly positive probability on more than one type, say, on R and r, then these probabilities will be proportional to those in the prior, in this case $(1-\gamma, \gamma)$, and we then say that the employee's perceived type is R/r.

(Note that, with this notation, the prior type is $R/r/L/l$ with probabilities $(50-\gamma/2, \gamma/2, 50-\gamma/2, \gamma/2)$.) A third property, which depends on our equilibrium but which seems to hold for any intuitive equilibrium, is that whenever an employee's perceived type puts positive probability on more than one type and one of these is a strong type, then it must also put positive probability on the corresponding regular type. So R/l is not possible and neither is $R/L/l$, but $R/r/l$ is possible. (Due to the setup and the focus on pure strategies, we end up with discrete types and point-mass beliefs. This is done for tractability reasons. In more realistic settings, we would expect types and beliefs to be more continuous.)

We also need to introduce some terminology. If an employee's perceived type includes both r and l, for example $R/r/l$, we say that the employee is *unrevealed*. (Note that in this particular case, the beliefs will put essentially 50/50 weight on l and r, as γ is vanishing, which simplifies the calculations. Note also that in this particular example of $R/r/l$, there is a small R-bias but it is vanishing.) It then follows that, in the equilibrium that we consider, an employee will be either unrevealed (with potentially a small L- or R-bias), or her perceived type is one of the following: $L, l, L/l, R, r, R/r$. Naturally, in the former three cases we will say that the player was L-revealed; in the latter three cases, R-revealed. Note that upon joining the organization, an employee's type is unknown and thus (completely) unrevealed, but the employee's actions can "reveal" her type (partially or fully). We will say that there is an L-majority (and analogously for an R-majority) if the number of L-revealed employees is strictly higher than the number of R-revealed employees.

We can now define norms and social norms in more formal terms and then translate that to the context of our model. Remember that we use the term *norm* to refer to "the set of expected or standard behaviors" themselves and *social* norm to refer to the enactment of such norms in an equilibrium. Consider then first norms. For the purposes of the main paper and the statement of the proposition, an L-norm will consist of the following behaviors: "choose L and vote against anyone who chose R in this round or who violated the norm in a previous round," and analogously for an R-norm.[42] Note that strong-R types will always deviate from an L-norm, because they mechanically choose R, and then be sanctioned (if there is a social norm).

Consider then a social norm. Remember that we say that a social norm exists in a group at a point in time when (in equilibrium) (1) a large part of the group adheres to a particular norm at that point in time, and (2) at least some people adhere to the norm because they fear social sanctions if they do not. In this model, the social norm will always take the form of either an L-norm or of an R-norm. Note that a situation where everyone behaves in a particular way just because they "like" that behavior is not a social norm according to this definition, because no one does so *out of fear of social sanctions*. (For example, eating three meals a day is the norm but that does not make it necessarily a social norm.) We will return to this later.

We now translate this to our model by describing the "social norm equilibrium". (We will discuss the interpretation at the end of this section, after the proposition.) The reason to call it a social norm equilibrium is because of how this will play out: after a first-round shake-out that establishes either an L-majority or an R-majority, the majority-norm will be observed (even by regular minority-type employees, including punishing any deviations toward the action that their type actually prefers) and minority-type employees adhere to the norm out of fear for social sanctions. In fact, with very high probability, no one will violate the norm in the second period and the game will move straight to the second part, that is, the private action round. We will describe this outcome in more detail, after we describe the actual action choices of the employees. The following description of their choices applies both on and off the equilibrium path, except for a few unusual cases (e.g., if a majority-type player is wrongly perceived to be of the minority type) that we deal with in the online appendix.

- Consider any decision point prior to the third (and last) elimination vote. If there is a L-majority, then the employees adhere to the L-norm, and analogous for an R-majority and the R-norm. If there is no majority of either type, then all employees choose a course of action according to their real type and vote to eliminate anyone who was revealed to be of the opposite type. (So an l-type will choose L and vote to eliminate anyone who was R-revealed.)
- In the, potential but unlikely, third elimination vote and for the final private action choices, all employees vote and choose according to their true type. (So an l-type will choose L and vote to eliminate anyone who was R-revealed.) In this case, the norm is not followed anymore because there is no threat of future social sanctions to enforce norms. However, there is only a very small probability that this third round takes place, as it occurs only if a second-round replacement was of the strong-minority type, which happens with probability $\gamma \downarrow 0$.

Beliefs (about employees' types) are determined by Bayes' rule whenever possible. When Bayes' rule does not have any bite—following off-equilibrium outcomes—the equilibrium specifies "skeptical" beliefs: a player who has *ever* violated—or failed to enforce—the norm in any round with a majority is perceived to be the minority type. (We specify the details in the online appendix.) Such skeptical beliefs are intuitive, and essentially prevent employees who have violated a norm from "rehabilitating" themselves simply by later adhering to the norm.

The more detailed outcome of our "social norm equilibrium" is then as follows. *First round – Shake-out:* Once employees have been randomly selected, they all simply choose the action that corresponds to their true type, which thus reveals their type. This always establishes either an (absolute) L-majority or an (absolute) R-majority. The majority-type employees then vote to eliminate all minority-type employees in the elimination stage of that first round. *Subsequent rounds, until third elimination stage – Social Norm:* In subsequent action stages, everyone but strong minority type employees adhere to the majority norm. Note that employees of the regular minority type, by adhering to the norm, remain unrevealed and will not be eliminated. But employees of the *strong* minority type will violate the norm, thus revealing their type. When such a strong minority type violates the norm, the regular minority types will vote with everyone else to eliminate that employee—that is, they enforce a norm that goes against their own preferences—in order to remain unrevealed. Because strong minority types get revealed and eliminated, the unrevealed employees have a slight majority-bias. *Final round – Endgame:* If the game proceeds to the third elimination stage, then everyone votes their true type. In the (final) private action round, everyone takes their preferred action. (This third round serves to close the game. More generally, if there were more rounds in the game, then all but the first and the last rounds would be fully norm-observing rounds. The first round would remain the shake-out round and the last round—pushed now further into the future—would remain the endgame, with the norm observed in the action stage but not in the elimination stage.)

Whereas the main contribution of this formal analysis is to show the emergence of a social norm. It is, however, also helpful to identify culture in the model. In general, an organization has a culture—according to our definition—if there is a critical mass of employees with shared (perceived) beliefs and preferences.[43] But culture is not a black-and-white phenomenon; instead it comes in gradations, from weak to strong cultures. This complicates the delineation of exact definitional boundaries. To keep things simple here, we will say for this model that there is (for sure) a culture if a strict majority of employees share a (perceived) belief or preference that all go in the same direction: either all L or all R. So if a strict majority of employees are perceived to have an L-belief or preference, then the organization has an L-culture, and analogously for an R-culture. Such culture can range from very strong—if all employees are perceived to have, say, an L-belief or preference—to weak—if only a bare majority of employees share (perceived) beliefs or preferences. The most common case in this particular model is actually a medium-strength culture with three or four employees having been revealed to have, say, an L-preference, whereas the remaining one to two employees are of type $L/l/r$, that is, unrevealed but with a slight bias toward L.

The following proposition shows that the social norm equilibrium indeed exists for at least some parameter values, and then further also gives comparative statics for when the equilibrium exists and when it does not. It also shows that an L-norm goes together with an L-culture, as delineated above, and analogously for R. As mentioned before, we focus on a vanishingly small probability of strong types, $\gamma \downarrow 0$. In terms of notation, let Ω be the allowed parameter space $\{(U,V,v,W,w,\epsilon) : U > 0, V > 0, v > 0, W > 0, w > 0, \epsilon > 0\}$ (excluding γ) and let ω be some generic point in Ω.

Proposition 1a. *The "social norm equilibrium" exists (in a nonempty part of the parameter space) and has the following properties:*

1. *On the equilibrium path, there exists (with strictly positive probability) a social norm in at least one round of this equilibrium.*
2. *Whenever this social norm takes the form of an L-norm, the organization will have an L-culture, and analogously for an R-norm and an R-culture.*

Proposition 1b. *This "social norm equilibrium" exists for sufficiently large U, small V, large v, small W, and intermediate w, in the following sense: There exists a subset S of Ω with nonempty interior such that:*

1. *The "social norm equilibrium" exists for any element in the interior of S*
2. *If $\omega = (U,V,v,W,w,\epsilon) \in S$ then a) $w \geq \epsilon$ and b) if $\omega' = (U',V',v',W',w',\epsilon) \in \Omega$ is such that $U' > U$, $V' < V$, $v' > v$, $W' < W$, $\epsilon < w' < w$ then ω' is also in S*
3. *The "social norm equilibrium" does not exist for any $\omega \in \Omega$ outside of S*

So what does this proposition say (beyond what we already discussed above)? Let us start with the comparative statics on when the "social norm equilibrium" exists. The social norm equilibrium exists for high U, that is, for attractive organizations, because that makes it worthwhile to adhere to the norm to avoid being eliminated; for low V, that is, low cost of the personal public actions, because that makes it cheap to signal a type; and for large v, that is, large benefits from personal private actions, because that again makes the organization attractive. Consider, finally, the comparative statics on W and w, that is, how much an employee is impacted by others' actions, which captures the degree of interdependence across employees. These are driven by two forces. On the one hand, you only have reason to enforce a norm or eliminate anyone if their actions affect you. So social norms are only possible when w is large enough ($w > \epsilon$).[44] On the other hand, being part of an organization where everyone follows norms that go against your personal beliefs and preferences is frustrating, especially when their actions really affect you. So very high W and w make an organization unattractive to an employee who disagrees with the norm, so that that employee will not bother to adhere to the norm. (In such case, all minority types will get eliminated, so the organization ultimately has no social norms—as no one takes actions to avoid social sanctions—but perfectly shared beliefs/preferences that make everyone behave (autonomously) in identical ways. We return to this below.) We had originally overlooked this second intuition but the formal analysis brought it out. Overall, social norms are thus more likely with intermediate interdependence. And as interdependence increases, an organization moves from having no shared beliefs or preferences and no social norms, to having partially shared beliefs and preferences with social norms that make everyone behave "as if", to ultimately having completely shared beliefs and preferences but again no social norms.

There are also a few interesting observations about the equilibrium itself. First, people often get eliminated not so much because of their public action (a_n) that triggered the elimination, but because of what that action reveals about their type (and thus about their future private action b_n).[45] Second, the flipside is, of course, that people adhere to a norm to avoid revealing that they have the "wrong" beliefs or preferences. And with elimination votes public, they have to also enforce the norm as if they held the "right" beliefs and preferences. That makes the norm thus serve as an "as if multiplier" on the shared beliefs of the majority type, so that these beliefs appear more widespread (to a naive observer) than they really are. Third, the (off-equilibrium) result that an L- or R-majority generates a social norm but that a stalemate does not, starts to capture the idea that you need a critical mass for a social norm to emerge. (Models with, e.g., more than two types would make this point more clearly.)

It is also helpful to consider two variations on this model, both to clarify what social norms exactly mean here and to relate this to earlier work with complete information, such as Van den Steen (2010a) and Li (2016). In both variations, the outcome superficially resembles our social norm equilibrium but with simple but conceptually important differences. (One could call these outcomes, for example, quasi-norms.) Consider first a complete-information version of this model where employees' types are publicly known from the start. In this case, all employees just choose actions according to their type and vote against employees of the opposite type, independent of what actions these employees take. On the one hand, this is clearly not a social norm (as defined above) as adhering to the norm cannot save you from social sanction and elimination. On the other hand, the outcome has some norm-like characteristics. In particular, through successive eliminations, employees' actual beliefs and preferences would converge over time and thus also their actions. So, some standard of behavior emerges. Moreover, employees of a different type will violate that standard and then be eliminated. However, the key difference is that they would get eliminated either way, independent of the action they take. So, whereas with a social norm, deviation from the norm *causes* elimination, here there is just *correlation*, and thus not a social norm. (In fact, the minority employees deviate from the norm in part because they know they will get eliminated anyways.) It is important to note, however, that we did not exclude a priori that "social norms cannot exist in a world with complete information." This is, instead, a (falsifiable) prediction of the model: if our mechanism is the (only) driver of social norms, then we will not observe social norms in a world with complete information. (In a world with complete information, true social norms could still exist, e.g., through a repeated game equilibrium. But this is an empirical question about the mechanisms causing social norms in actual organizations.) Note that culture *does* emerge in this setting (where perceived and actual beliefs coincide). In fact, a very strong culture emerges, with all employees ultimately sharing (perceived and actual) beliefs or preferences.

The second interesting variation is to consider what happens if there is private information but everyone is of a strong type. Now a player *could* avoid social sanctions by adhering to a norm and posing as the opposite type, yet no one does. As in the case of complete information, the underlying beliefs and preferences and the employees' actions will converge over time. This has even more norm-like characteristics than the complete information case, but is still not a social norm according to our earlier definition because no one adheres to the norm *in order to* avoid the social sanction. We do consider this more of a borderline case, but feel that our definition corresponds best with the common use of the term. In this variation, again, culture *does* emerge.

7. Implications for the Definition and Conceptualization of Corporate Culture

Whereas this paper uses a very specific definition of corporate culture—as shared (perceived) beliefs and preferences—there is a renewed debate unfolding in the management literature on how to exactly define corporate culture. Even though we feel this paper's definition is the right one for a corporate context, it would be a mistake to disregard this debate, especially as the field may well settle on a different definition from ours. It is thus important to ask: if "shared (perceived) beliefs and preferences" does not turn out to be the definition to which the field converges, what can be learned from this paper that is relevant to this discussion on the nature and definition of corporate culture? There are a number of relevant insights.

First, the results of this paper offer some arguments in favor and some arguments against explicitly including norms in definitions of corporate culture. Most obviously, the paper shows that norms may be implied by the shared (perceived) beliefs and preferences. One could therefore argue that it suffices to specify these shared (perceived) beliefs and preferences, that is, that they are a "sufficient statistic" for the norms. In the other direction, however, the paper does not provide any evidence, theoretical or empirical, that *all* relevant norms result from this mechanism.

Second, together with Van den Steen (2010a), the paper does provide a clear argument that shared beliefs and preferences have performance effects beyond norms and should thus be included separately in any definition or conception of corporate culture.

Third, the paper has made clear how the choice between "values" and "preferences" is consequential for the relevance of culture, and made a strong case in favor of preferences.

Finally, this paper makes very clear that it is important to distinguish between the actual beliefs and preferences and the perceived beliefs and preferences. The analysis and results of this paper suggest in fact that, unlike the existing definitions in the literature, culture should be defined as being about the perceived beliefs and preferences (shared among a group of critical size). It turns out that all results of, for example, Van den Steen (2010a), Van den Steen (2010b), or Li (2016) go through under this definition (as there is no private information in these models).

Overall, this paper subscribes to the view that clear definitions—with unambiguous necessary and sufficient conditions—are an important foundation for good research. We therefore hope that this paper at least contributes to that evolving discussion on the definition of corporate culture.

8. Conclusion

This paper defined corporate culture as "shared (perceived) beliefs and preferences"—extending earlier work—and showed that the forces that drive toward such shared (perceived) beliefs and preferences also give rise to social norms, including norm enforcement even by people who disagree with the norm. This result implies a very close connection between corporate culture (as shared (perceived) beliefs and preferences) on the one hand and social norms on the other. The paper then explored the broad implications of this connection for the nature and functioning of culture, for norms, and for their interaction with strategy.

The paper raises a number of interesting further research questions, such as:

- Can we develop and test theory on the dimensions along which culture and norms are more or less likely to evolve—assuming a broad view of culture as shared (perceived) beliefs and preferences and the norms that are connected to it?
- Can, in some settings, formal organization be used to counter the effect of culture in cases where strategy and culture conflict?
- Would experiments that change the task environment be able to encourage the emergence or extinction of social norms?
- Which aspects of corporate culture are (theoretically and/or empirically) most likely to interfere with strategy?

On a higher level, this paper hopes to have made two contributions: (1) encourage the view that microfoundations can really improve our understanding of complex phenomena, and (2) contribute new insights on the nature of corporate culture and on the interaction between culture and strategy.

Acknowledgments

This work benefited from many discussions with, and inspiration from, Robert Akerlof, Ramon Casadesus-Masanell, and John Roberts; from the feedback by participants in the Berkeley Haas Culture Conference and the Strategy Science Conference; and from very helpful feedback from two anonymous referees. The authors owe special gratitude for this paper (and other projects) to Bob Gibbons, the Senior Editor, whose guidance, suggestions, and constructive criticisms (both conceptual and expositional) brought out the best. The authors thank Catherine Barrera for research assistance on an earlier incarnation of this project. Jenny Chatman and Sameer Srivastava gave some instrumental encouragement.

Endnotes

[1] Note that this is intended as a semiformal paper. Although many of the arguments in the paper are illustrated with the model and with formally proved results, the analysis does not provide a fully general proof of all arguments. (See footnote 19.) Such a general formal approach requires a different type of paper and goes beyond the current purpose.

[2] The paper does not claim that all norms are of this type—though it also does not deny that. Its claim is that this is a logical mechanism for norms to emerge, that such norms should obey particular

comparative statics, and that these comparative statics make it possible to test whether this mechanism is at work.

[3] The focus is here on business strategy or organization strategy, not on global or corporate strategy.

[4] Although these definitions of corporate culture are very different, we do conjecture that culture as shared beliefs and preferences can facilitate, and be a determinant of, equilibrium selection and convergence. In fact, Van den Steen (2010a) argued formally in the context of coordination that shared beliefs and preferences facilitate and accelerate equilibrium selection and convergence. Although not made explicit, these shared beliefs and preferences also played an important role in determining *which* equilibrium obtained. Note that this does not contradict the distinction we make here: "equilibrium selection" is one of many *effects* of "culture as shared and preferences," whereas it is at the heart of the *concept itself* for culture as equilibrium. There may be a second connection between the two approaches as building an equilibrium is also about developing some kind of "shared beliefs": shared beliefs about the equilibrium that is being played. This awaits further research.

[5] There is also some other literature that is orthogonal to these two (related) views of corporate culture. First, Guiso et al. (2006) study the effect of *personal* values, as determined by your *national* culture, on economic outcomes. Their definition of culture, rooted in anthropology, does not include any reference to whether these values are "*shared*," which is the key discriminant of *corporate* culture. Second, Guiso et al. (2015) study the effect of how employees *perceive* top management's values–in particular, whether they believe that management keeps its word or that it is honest and ethical. They call these perceptions "culture," independent of whether this is shared among the management team or in the organization. This is obviously a very different use and a very different interpretation of culture and a different focus in terms of research.

[6] Bernheim (1994), like Bénabou and Tirole (2006, 2011), occasionally uses the term "culture" but in an informal sense to refer to the broader social or societal context within which norms develop. Bernheim (1994) discusses, for example, how cultures may differ in how much they value social status, which is the key driver of norms in his model. Used in that sense, culture is a cause of forces toward conformity rather than the effect of such forces.

[7] The global strategy literature also has work on culture as an environmental factor, opportunity, or constraint, such as Ghemawat (2001) or Siegel et al. (2013), but this is quite removed from this paper.

[8] The 128-page strategy reading list for PhD students compiled by a group of faculty in the "Strategy Research Initiative" mentioned the word *culture* just once, as a potential asset (SRI 2009).

[9] Kotter and Heskett (1992) defined culture on the deepest level as "the values that are shared" (p. 4). Schein (1985) defined culture as "shared basic assumptions" (p. 9) but also said that culture consisted of three levels with the middle level being "shared values" (p.15) and the deepest level "shared assumptions." Schein essentially claimed that all shared values are ultimately rooted in shared assumptions. This suggests a much broader interpretation of assumptions—encompassing both beliefs and preferences—than its typical use. This broader interpretation is also suggested by Schein's statement (in his definition) that these assumptions tell you, among other things, "the correct way … to feel" (p. 9) and by the fact that one of his examples of assumptions is that "an individual's rights supersede those of the group" (p. 18), which is more a value than an assumption. Some of the issue here seems to reflect the fact that Schein, as an early (seminal) contributor, was trying to apply the anthropological approach of culture to a corporate culture context, without fully integrating the effect of sorting and selection (even though he explicitly recognized these forces). It took some time before their full implications became clear. Note also that the definition in Van den Steen (2010a) built on earlier work by Crémer (1993) and Lazear (1995), as discussed there.

[10] It is also very effective to teach students and executives about corporate culture.

[11] Although the results that follow were proved only in settings with complete information and thus, strictly speaking, valid only in the context of the earlier definition, we (strongly) conjecture that they extend—with appropriate modifications—to settings with incomplete information and, hence, to the new definition. But showing this formally goes beyond the purpose of this paper.

[12] Although the paper formulates the model and results in terms of shared knowledge, it can also be interpreted—and was originally intended—as being about shared beliefs or preferences.

[13] Note that the stated definition in Van den Steen (2010a) is also in terms of "values" but that paper used "values" as a synonym for preferences, simply because it was more in line with the terminology of the management literature.

[14] Defining culture as having some degree of desirability also risks making important performance aspects of culture tautological. What would it mean to show that culture is valuable if culture is restricted to be about desirable values? Moreover, this very discussion suggests that the term "value" has some ambiguity.

[15] This issue seems rooted in the earlier-mentioned distinction between anthropological and corporate culture. As an aside, one could argue that sorting is often preceded by joint learning, but this argument does not negate the conceptual issue that pure sorting/screening would still be excluded by Schein's (1985) definition. Although it is interesting to understand where cultures come from, the perspective of this paper is that that should be a research question instead of part of the definition.

[16] A careful evaluation of "necessary and sufficient" often highlights needs for clarification. For example, a definition that says that culture consists of "shared values and shared norms" must be clear on whether you need *both* values *and* norms or whether one or the other is sufficient; who evaluates whether something is a value or not; etc. Van den Steen (2018) showed, for example, that the most common proposed definitions of "strategy" fail to give "necessary and sufficient" conditions.

[17] We thank Bob Gibbons for this "behaviors versus mental processes" framing, among other things.

[18] Axelrod (1986) refers to such higher-order norms—which specify that those who do not punish norm violations should themselves be punished—as "metanorms," and illustrates the concept with numerous historical examples. Relatedly, in the context of collective-action problems, the need to motivate sanctioning behavior is known as the "second-order free-rider problem" (Heckathorn 1989, Coleman 1990).

[19] To keep the analysis simple and tractable, the model has some less realistic features—like a voting process, a finite number of rounds, and a small number of participants—that limits its generality. Moreover, the proof only shows existence rather than uniqueness of the "norm equilibrium." We do conjecture uniqueness, potentially under some (reasonable) conditions, but that goes beyond the purpose of this paper.

[20] If potential employees could self-select into the organization, such nonconforming employees would avoid joining in the first place, and there would be no norm violations in equilibrium. But there still is no social norm—according to our definition—because no one adheres to the norm *out of fear of social sanctions*.

[21] As discussed earlier, some of the results that we conjecture here could likely also be generated in a model where norms exist as equilibria of infinitely repeated games. For instance, norms of cooperation are easier to sustain in infinitely repeated games when

interactions become more frequent. However, the underlying mechanisms are often quite different. Consider our conjecture about the persistence of early norms. In our model, persistence arises because later members of the organization hide their beliefs and preferences, thus leaving early members with outsize and persistent influence over norms. This is quite different from the mechanisms that generate persistence in relational contracting models such as Chassang (2010), Li, Matouschek, and Powell (2017), or Andrews and Barron (2016). See, also, the discussion in Gibbons and Henderson (2013, section 5).

[22] We thank a referee for pointing out that this does not only apply to vices but also to regular preferences.

[23] Note that the fact that values may be more strongly held than regular preferences also has important implications, as discussed.

[24] This is difficult to see in the model of Section 6 for two reasons. First, there are, for transparency and analytical simplicity, only two types. As a consequence, anyone of unknown type is 50% likely to have your type. This makes it possible to support norms with only a small majority of people who have been revealed. In a setting with a larger number of types, norms will only emerge when there is a substantial group with shared beliefs/preferences. Second, elimination is by simple majority vote in that model. A more onerous elimination process would require a larger critical mass.

[25] We thank a referee for pushing us on this issue and for making the argument that strongly held beliefs/preferences make people more willing to enforce norms.

[26] This could be captured in the model by introducing a vanishing probability that each participant is (unobservably) replaced with a random draw, as in Tirole (1996).

[27] Even worse, in a type of result that may be familiar from the common knowledge literature in game theory, it is possible that norms are enforced even though *everyone* in the organization is aware that the norms conflict with *everyone else*'s beliefs and preferences. The reason is again because of a "beliefs about beliefs" issue: even when everyone knows that fact, each person may be unsure about others' beliefs and, in the face of that uncertainty, may then find it safer to follow and enforce the norm rather than risk getting eliminated. For an introduction to these ideas, see Chwe (2001).

[28] This situation is reminiscent of the fairly well-known "Four Monkeys" scientific experiment reported by Hamel and Prahalad (1996) where monkeys learned from other monkeys not to climb a pole that had bananas at the top because these earlier monkeys had been sprayed when trying to do so. New monkeys did not try to climb the pole even after the spraying was stopped. Unfortunately, there is no evidence, from Hamel and Prahalad (1996) or other sources, that such an experiment—where monkeys get sprayed when they climb a pole with bananas—has really been conducted.

[29] Although obviously not trivial, underlying (actual) beliefs and preferences can be discovered by careful, in-depth, and targeted interviews and observation. One potential approach is discussed by Schein (1985).

[30] Note that this issue is really caused by the existence of a norm because the principal would learn the agent's actual beliefs and preferences in the absence of a norm.

[31] An example would be a setting where current norms or behaviors have disastrous consequences that are not readily observed by those who follow these norms or engage in these behaviors. Pointing out these consequences may have drastic impacts on beliefs and preferences over such behaviors, and then lead to quick changes in norms. But this is a fairly specific setting.

[32] Our interpretation here is that some strongly held belief or preference ("care about patients")—around which underlying shared beliefs or preferences develop—is combined with a much weaker belief about what is best for patients: make them feel better by "prescribing what they want" versus improve their health by "prescribing what they need." When you are almost indifferent between the two beliefs about what is best for patients, then choosing one path, say, "prescribing what they need" —even if you actually believe more in the other—is a cheap way to signal your deeper and more important belief/preference about "caring about patients."

[33] A common theme in experiences with quick norm reversion is that the responsible manager declared far too soon "victory" on the culture change and moved on to other opportunities.

[34] This finding is somewhat reminiscent of Tirole (1996) where a sustained period of good behavior by an agent is necessary to move a firm from a "high" to a "low corruption" steady state. The mechanisms are very different, however. In Tirole (1996), sustained good behavior serves to rehabilitate the reputation of the agent. Here, a sustained period of norm enforcement serves instead to force self-sorting by imposing costs on employees who oppose these norms.

[35] This question matters not only for practice—to really understand the effects of culture—but also for the earlier-mentioned debate in the broader literature on the definition of culture: if not all effects do go through norms, then "shared beliefs and preferences" should be part of a good definition.

[36] The claim is often attributed to Drucker, but there is no evidence that he ever wrote or said it. The phrase is most connected to Mark Fields, former CEO of Ford Motor Company, who popularized it.

[37] Which is more important, a car or a wheel?

[38] Although the resistance to change probably came in many different forms—from dragging feet to open power struggles—a united front of "shared beliefs" or "shared preferences" is a really strong barrier, as no one wants to be seen breaking with the culture.

[39] Because part of the proof (currently) consists of brute force calculations of incentive constraints, it is helpful to keep the number of positions down. We limit for the same reason the number of elimination opportunities to three.

[40] This $\gamma \downarrow 0$ assumption vastly simplifies the determination of beliefs. Similarly, we assume for simplicity a finite number of preference types and strengths; whereas a continuous distribution of preference (or priors) strengths would be more natural, it would complicate the analysis dramatically.

[41] These choices can be endogenized with appropriate utility functions but at a considerable analytical cost.

[42] To deal with some unusual off-equilibrium-path cases, the appendix and proof uses a formal definition of L-norm and R-norm that is slightly more general—though unfortunately somewhat less transparent: an L-norm consists of "choose L and vote to eliminate anyone who is R-revealed." This is a generalization of our more informal statement because anyone who violates the L-norm becomes R-revealed. An example of an off-equilibrium-path case that requires this more general formulation is the situation—impossible on any reasonable equilibrium path—where someone first violates the L-norm in the face of an L-majority but then decides to violate the R-norm in the face of the subsequent R-majority. The generalization resolves some ambiguity on how to interpret this.

[43] We will immediately set the threshold for critical mass to be a strict majority. This criterion, although simple, is arguably too stringent. We choose it for a very practical reason. To see the reason, note that defining culture in terms of a critical mass allows for subcultures. A real study of subcultures, however, goes beyond this paper. By using "strict majority" as the cutoff, we exclude subcultures from the analysis while recognizing their potential importance for a more general theory. We do, indeed, believe that subcultures are an important research question. More work is needed to analyze this in detail and think carefully about the concept of culture in such settings.

[44] The fact that W is not subject to that condition has to do with the specific setup. In most settings, there would be a similar condition on W.

[45] In the third round, for example, norm violators get eliminated (by the majority) even though their public action a_n is frozen, to prevent them from taking their private action.

References

Abernethy M, Dekker H, Schulz A (2015) Are employee selection and incentive contracts complements or substitutes? *J. Accounting Res.* 53(4):633–668.
Akerlof GA, Kranton RE (2000) Economics and identity. *Quart. J. Econom.* 115(3):715–753.
Andrews I, Barron D (2016) The allocation of future business: dynamic relational contracts with multiple agents. *Amer. Econom. Rev.* 106(9):2742–2759.
Axelrod R (1986) An evolutionary approach to norms. *Amer. Political Sci. Rev.* 80(4):1095–1111.
Barney JB (1986) Organizational culture: Can it be a source of competitive advantage? *Acad. Management Rev.* 11(3):656–665.
Bénabou R, Tirole J (2006) Incentives and prosocial behavior. *Amer. Econom. Rev.* 96(5):1652–1678.
Bénabou R, Tirole J (2011) Laws and Norms. CEPR discussion paper no. DP8663.
Bernheim BD (1994) A theory of conformity. *J. Political Econom.* 102(5):841–877.
Bicchieri C (2006) *The Grammar of Society: The Nature and Dynamics of Social Norms.* (Cambridge University Press, Cambridge, MA).
Bicchieri C, Muldoon R, Sontuoso A (2018) Social norms. Zalta EN, ed. *The Stanford Encyclopedia of Philosophy* (Metaphysics Research Laboratory, Stanford University, Stanford, CA), 1–41.
Campbell D (2012) Employee selection as a control system. *J. Accounting Res.* 50(4):931–966.
Carrillo JD, Gromb D (1999) On the strength of corporate cultures. *Eur. Econom. Rev.* 43:1021–1037.
Chassang S (2010) Building routines: Learning, cooperation, and the dynamics of incomplete relational contracts. *Amer. Econom. Rev.* 100(1):448–465.
Chatman JA (1991) Matching people and organizations: Selection and socialization in public accounting firms. *Admin. Sci. Quart.* 36:459–484.
Chatman JA, O'Reilly CA (2016) Paradigm lost: Reinvigorating the study of organizational culture. *Res. Organ. Behav.* 36:199–224.
Chatman JA, Gibbons R, Goldberg A, O'Neill O, Rivera L, Siegel J, Srivastava S, Weber R, Zenger T (2019) Reinvigorating research on organizational culture and its link to strategy. Call for Papers for Special Issue of Strategy Science.
Chwe MS-Y (2001) *Rational Ritual: Culture, Coordination, and Common Knowledge* (Princeton University Press, Princeton, NJ).
Cialdini RB, Trost MR (1998) Social influence: Social norms, conformity and compliance. Gilbert DT, Fiske ST, Lindzey G, eds. *The Handbook of Social Psychology* (McGraw-Hill, New York), 151–192.
Coleman JS (1990) Foundations of Social Theory (Belnap Press, Cambridge, MA).
Crémer J (1993) Corporate culture and shared knowledge. *Indust. Corp. Change* 3(3):351–386.
Deal T, Kennedy AA (1982) *Corporate Cultures: The Rites and Rituals of Corporate Life* (Addison Wesley, Reading, MA).
Deutsch M, Gerard HB (1955) A study of normative and informational social influences upon individual judgment. *J. Abnorm. Soc. Psych.* 51(3):629–636.
Dimaggio P (1994) Culture and economy. Smelser N, Swedberg R, eds. *Handbook of Economic Sociology* (Russell Sage, New York), 27–57.
Epstein GS, Lindner-Pomerantz R (2017) The survival of unique corporate cultures. *Management Decision Econom.* 38(4):622–629.
Erikson E (1968) *Identity: Youth and Crisis* (Norton, New York).
Fayol H (1916) Administration industrielle et generale. *Bull. Soc. Indust Minér.* 10:5–164.
Fiske S, Taylor S (1991) *Social Cognition* (McGraw-Hill, New York).
Geertz C (1973) *The Interpretation of Cultures: Selected Essays* (Basic Books, New York).
Ghemawat P (2001) Distance still matters: The hard reality of global expansion. *Harvard Bus. Rev.* 79(8):137–147.
Gibbons R, Henderson R (2012) Relational contracts and organizational capabilities. *Organ. Sci.* 23(5):1350–1364.
Gibbons R, Henderson R (2013) What do managers do? Exploring persistent performance differences among seemingly similar enterprises. Gibbons R, Roberts DJ, eds. *The Handbook of Organizational Economics* (Princeton University Press, Princeton, NJ), 680–731.
Goldberg A, Srivastava SB, Manian VG, Monroe W, Potts C (2016) Fitting in or standing out? The tradeoffs of structural and cultural embeddedness. *Amer. Sociol. Rev.* 81(6):1190–1222.
Grant RM (2016) *Contemporary Strategic Analysis* (John Wiley & Sons, Hoboken, NJ).
Groysberg B, Lee J, Price J, Cheng JY-J (2018) The leader's guide to corporate culture. *Harvard Bus. Rev.* 96(1):44–52.
Guiso L, Sapienza P, Zingales L (2006) Does culture affect economic outcomes? *J. Econom. Perspect.* 20(2):23–48.
Guiso L, Sapienza P, Zingales L (2015) The value of corporate culture. *J. Financial Econom.* 117(1):60–76.
Hamel G, Prahalad C (1996) *Competing for the Future* (HBR Press, Brighton, MA).
Heckathorn DD (1989) Collective action and the second-order freerider problem. *Rationality Soc.* 1(1):78–100.
Hofstede G (1980) *Culture's Consequences: International Differences in Work-Related Values* (Sage Publications, London, UK).
Huston TL, Levinger G (1978) Interpersonal attraction and relationships. *Annual Rev. Psych.* 29:115–156.
Kotter JP, Heskett JL (1992) *Corporate Culture and Performance* (Free Press, New York).
Kreps DM (1990) Corporate culture and economic theory. Alt JE, Shepsle KA, eds. *Perspectives on Positive Political Economy* (Cambridge University Press, Cambridge, UK), 90–143.
Kreps D, Milgrom P, Roberts J, Wilson R (1982) Rational cooperation in the finitely-repeated prisoners' dilemma. *J. Econom. Theory* 27(2):245–252.
Kuhn S (2019) Prisoner's dilemma. Zalta EN, eds. *The Stanford Encyclopedia of Philosophy* (Metaphysics Research Laboratory, Stanford University, Stanford, CA).
Lazear EP (1995) Corporate culture and diffusion of values. Siebert H, ed. *Trends in Business Organization: Do Participation and Cooperation Increase Competitiveness?* (Mohr Publishing Company, Tubingen, Germany), 89–133.
Legros S, Cislaghi B (2020) Mapping the social-norms literature: An overview of reviews. *Perspect. Psych. Sci.* 15(1):62–80.
Levine SR (2019) A superb strategy without the culture to support it leads to failure. *Forbes.* https://www.forbes.com/sites/forbesinsights/2019/06/17/a-superb-strategy-without-theculture-to-support-it-leads-to-failure/?sh=40b34fba4075.
Li H (2016) Developing shared knowledge in growing firms. *J. Law Econom. Organ.* 33(2):332–376.
Li J, Matouschek N, Powell M (2017) Power dynamics in organizations. *Amer. Econom. J. Microeconom.* 9(1):217–241.
McPherson M, Smith-Lovin L, Cook JM (2001) Birds of a feather: homophily in social networks. *Annual Rev. Sociol.* 27:415–444.
O'Reilly CA, Chatman JA (1996) Culture as social control: Corporations, cults, and commitments. *Res. Organ. Behav.* 18:157–200.
Parks L, Guay RP (2009) Personality, values, and motivation. *Perspest. Individual Differences* 47(7):675–684.
Parsons T (1951) *The Social System* (Routledge, New York).
Prasad S, Tomaino M (2020) Resources and culture in organizations. *J. Econom. Management Strategy* 29:854–872.

Roethlisberger FJ, Dickson WJ (1939) *Management and the Worker: An Account of a Research Program Conducted by the Wester Electric Company, Hawthorne Works, Chicago* (Harvard University Press, Cambridge, MA).

Roy D (1952) Quota restrictions and goldbricking in a machine shop. *Amer. J. Sociol.* 57(5):427–442.

Saloner G, Shepard A, Podolny J (2001) *Strategic Management* (John Wiley & Sons, Hoboken, NJ).

Schein EH (1985) *Organizational Culture and Leadership* (Jossey-Bass Publishers, San Francisco, CA).

Siegel JI, Licht AN, Schwartz SH (2013) Egalitarianism, cultural distance, and foreign direct investment: A new approach. *Organ. Sci.* 24(4):1174–1194.

SRI (2009) *Strategy Reader* (Strategy Research Initiative, Toronto), 1348–1364.

Srivastava SB, Goldberg A, Manian VG, Potts C (2018) Enculturation trajectories: Language, cultural adaptation, and individual outcomes in organizations. *Management Sci.* 64(3):983–1476.

Tajfel H, Turner JC (1979) An integrative theory of inter-group conflict. Austin W, Worchel S, eds. *The Social Psychology of Inter-Group Relations* (Brooks/Cole, Monterey, CA), 33–47.

Tirole J (1996) A theory of collective reputations (with applications to the persistence of corruption and to firm quality). *Rev. Econom. Stud.* 63(1):1–22.

Turner JC, Hogg MA, Oakes PJ, Reicher SD, Wetherell MS (1987) *Rediscovering the Social Group: A Self-Categorization Theory* (Basil Blackwell, Oxford, UK).

Van den Steen E (2010a) Culture clash: The costs and benefits of homogeneity. *Management Sci.* 56(10):1718–1738.

Van den Steen E (2010b) On the origin of shared beliefs (and corporate culture). *RAND J. Econom.* 41(4):617–648.

Van den Steen E (2017) A formal theory of strategy. *Management Sci.* 63(8):2616–2636.

Van den Steen E (2018) The strategy in competitive interactions. *Strategy Sci.* 3(4):574–591.

Weber RA, Camerer CF (2003) Cultural conflict and merger failure: An experimental approach. *Management Sci.* 49(4):400–415.

Willer R, Kuwabra K, Macy MW (2009) The false enforcement of unpopular norms. *Amer. J. Sociol.* 115(2):451–490.

Young HP (2015) The evolution of social norms. *Annual Rev. Econom.* 7:359–387.

Hongyi Li is a senior lecturer at University of New South Wales Business School. He is an applied theorist whose research has focused on how organizations change (or fail to change), and on communication and coordination problems within organizations.

Eric Van den Steen is a professor of business administration at Harvard Business School, where he holds the Royal Little Chair. His research studies the fundamentals of strategy and competitive advantage, the role of leaders in strategy, and the interaction between strategy and organization, with particular attention to vision and culture.