

Duality in Diversity: How Intrapersonal and Interpersonal Cultural Heterogeneity Relate to Firm Performance*

Matthew Corritore¹, Amir Goldberg², and Sameer B. Srivastava³

¹Desautels Faculty of Management, McGill University

²Stanford Graduate School of Business

³Haas School of Business, University of California, Berkeley

Forthcoming, Administrative Science Quarterly
February, 2019

Abstract

How does cultural heterogeneity in an organization relate to its underlying capacity for execution and innovation? Existing literature often understands cultural diversity as presenting a trade-off between task coordination and creative problem-solving. This work assumes that diversity arises primarily through cultural differences between individuals. In contrast, we propose that diversity can also exist within persons such that cultural heterogeneity can be unpacked into two distinct forms: interpersonal and intrapersonal. We argue that the former tends to undermine coordination and portends worsening firm profitability, while the latter facilitates creativity and supports greater patenting success and more positive market valuations. To evaluate these propositions, we use unsupervised learning to identify cultural content in employee reviews of nearly 500 publicly traded firms on a leading company review website and then develop novel, time-varying measures of cultural heterogeneity. Our empirical results lend support for our two core propositions, demonstrating that a diversity of cultural beliefs in an organization does not necessarily impose a trade-off between operational efficiency and creativity.

*We thank Glassdoor for providing the employee review data and Lee Fleming for sharing the patent data that we used in this study. We also thank seminar participants at University of Illinois School of Labor and Employment Relations; University of Chicago, Booth School of Business; IESE Business School; Kellogg School of Management, Northwestern University; Arison School of Business at IDC Herzliya, Israel; and participants at the Lugano Organizations Conference, the Stanford/Berkeley Organizational Behavior Student Conference, the Consortium on Competitiveness and Cooperation Conference for Doctoral Student Research at the Wharton School, University of Pennsylvania, the Carnegie School of Organizational Learning Conference, the International Conference on Computational Social Science, the American Sociological Association Annual Meeting, and the Wharton Corporate Strategy and Innovation Conference. We are also grateful to Lara Yang for the research support she provided, and to Marc-David Seidel and three anonymous reviewers for their helpful comments and suggestions. The usual disclaimer applies.

Whether deliberately cultivated or naturally arising, every organization develops a culture—a system of meanings and norms shared by its members. An organization’s culture can influence the success of its members and the organization as a whole through its effects on individual motivation and commitment, interpersonal coordination, and group creativity and innovation (Chatman and O’Reilly, 2016). Although organizational scholars often ask how the content or intensity of culture relates to performance—for example, how shared beliefs and norms about the importance of cross-functional collaboration can boost or diminish firm profitability—a growing literature has focused instead on the consequences of *cultural heterogeneity* for organizational productivity and vitality. Research in this vein asks: When is a diversity of ideas and beliefs conducive to organizational success and when is it instead detrimental?

Different literatures have provided varied and inconsistent answers to this question. One line of work sees cultural heterogeneity as an obstacle to firm performance. Research on cultural strength, for example, emphasizes the importance of cultural agreement among organizational members. This perspective argues that incompatibilities in employees’ beliefs and normative expectations can impede their ability to coordinate tasks (Denison and Mishra, 1995; Weber and Camerer, 2003; Kotter and Heskett, 1992), thereby producing a negative relationship between cultural heterogeneity and firm performance. In contrast, research by economic and cultural sociologists typically views cultural diversity as an advantage. This perspective conceptualizes heterogeneity as a reflection of the cultural “toolkit” available to individuals (Swidler, 1986). When organizational members have access to a broad array of cultural resources, the organization is assumed to have greater capacity for creativity and innovation to address changing, uncertain, and competing environmental demands (Stark, 2011; Hallett and Ventresca, 2006).

This tension is also manifest in work on demographic diversity. Although it has primarily focused on diversity in ascribed characteristics such as sex or age and in functional experiences, this literature has also called attention to deep-level differences in how people think and what they believe about group and organizational culture (Harrison et al., 1998). Whereas some studies

highlight demographic diversity’s negative implications for performance through the facilitation of conflict and coordination inefficiencies, others focus on its positive effects on collective learning and creativity (Fiol, 1994; Van Knippenberg and Schippers, 2007; Page, 2007; Lau and Murnighan, 1998).

Taken together, these literatures suggest that cultural heterogeneity presents a fundamental trade-off: Culturally diverse firms are better at creative problem-solving, but this capability comes at the cost of decreased coordination and efficiency. Empirical support for such a trade-off, however, remains meager and unsettled. We reconcile these divergent theoretical perspectives and mixed empirical findings by drawing on the core insight that cultural heterogeneity is not a unidimensional construct. Existing literature has conceptualized this diversity exclusively through the prism of differences between individuals. Yet heterogeneity can also exist within persons (Bunderson and Sutcliffe, 2002; Goldberg et al., 2016; Morris et al., 2015). Building on this insight, we propose that there are two distinct forms of cultural heterogeneity in organizations: that stemming from misalignment among members’ cultural perceptions of the organization and that arising from the breadth of cultural resources they use to understand and make sense of the organization.

As illustration, imagine a stylized world in which there exist only two possible beliefs, *A* and *B*, about how work ought to be done. Imagine further two hypothetical organizations. In the first organization, half of the employees espouse only belief *A* and the other half espouse only belief *B*. In the second organization, *all* employees espouse *both* beliefs. While the two organizations appear to be equally culturally heterogeneous—the beliefs account for the same proportion in both—the nature of this diversity varies between them. In the first organization, heterogeneity stems from divergent cultural beliefs between people, which we refer to as *interpersonal heterogeneity*. In the second organization, heterogeneity stems not from divergent beliefs, but instead from individuals having a multiplicity of cultural beliefs about the organization. We refer to this form as *intrapersonal heterogeneity*. Although interpersonal and intrapersonal heterogeneity are interrelated, they are analytically distinct.

Integrating insights from demographic diversity, group learning, and cultural sociology, we argue that interpersonal and intrapersonal heterogeneity are related to different organizational outcomes. Consistent with work on the organizational consequences of cultural strength (e.g., Kotter and Heskett, 1992), we theorize that interpersonal heterogeneity will be linked to a firm’s coordination and execution capabilities and thus to indicators of performance such as firm profitability. In contrast, drawing on group learning research (e.g., Page, 2007) and the toolkit theory of culture (e.g., Swidler, 1986), we posit that intrapersonal heterogeneity will be tied to a firm’s capacity for creative exploration and therefore to patenting success and market expectations of future growth. Decomposing cultural heterogeneity into its interpersonal and intrapersonal components reveals that there need not be a trade-off between organizational coordination and innovation—for example, a culture with low interpersonal but high intrapersonal heterogeneity can facilitate coordination without necessarily undermining creative problem-solving and innovation.

Empirically, the methods most commonly used to study organizational culture—chiefly self-reports (O’Reilly et al., 1991) and participant-observation (Kunda, 2009; Turco, 2016)—can yield rich insight but are not well-suited to generating dynamic measures of cultural heterogeneity for a large, diverse sample of firms. In part for this reason, prior work examining the link between culture and firm performance has tended to rely on cross-sectional designs and, for the most part, side-stepped questions about the dynamic interplay between the two. Moreover, survey-based methods highlight variation between respondents (O’Reilly et al., 1991) but are less well-suited to detecting differences in intrapersonal heterogeneity across organizations.

To overcome these limitations, we apply the tools of computational linguistics to derive novel, time-varying measures of interpersonal and intrapersonal heterogeneity for a sample of nearly 500 publicly traded companies on Glassdoor (www.glassdoor.com)—a career intelligence website that allows employees to evaluate and write reviews about their firms. Given that cultural content can appear in a wide variety of reviews, we use unsupervised learning to identify distinct cultural topics in the nearly one million sentences that contain the word “culture” and its synonyms. We then

train a topic model, which we fit to all employee reviews in our sample, and derive our cultural measures based on these identified topics. To move closer to causal estimates of the relationship between these two forms of cultural heterogeneity and firm performance, we apply coarsened exact matching (Iacus et al., 2012) to identify pairs of firms that vary on the cultural dimensions of interest but are otherwise observationally equivalent. Our results demonstrate that, counter to prevailing wisdom, cultural heterogeneity does not necessarily embody a conflict between execution efficiencies and creative capacity in organizations. We find that firms with more interpersonal heterogeneity have a lower return on assets, while those with more intrapersonal heterogeneity have a higher Tobin’s Q, produce more patents, and produce patents of higher quality. Our study provides the first evidence, to our knowledge, of a dynamic relationship between cultural heterogeneity and different performance outcomes in a comprehensive, longitudinal dataset of the largest firms in the U.S.

CULTURAL HETEROGENEITY AND FIRM PERFORMANCE

Why Cultural Heterogeneity Matters

Culture is often understood by organizational scholars as a “system of publicly and collectively accepted meanings” (Pettigrew, 1979, p. 574) that a group—including a formal organization—develops in response to challenges of external adaptation and internal integration. These meanings manifest both in the form of deeply rooted assumptions and beliefs about the world, as well as in the normative and behavioral expectations that these beliefs and assumptions prescribe (Schein, 2010). Individuals learn to recognize, internalize, and conform to the organization’s cultural code through the ongoing process of socialization and enculturation (Ashford and Nurmohamed, 2012; Srivastava et al., 2018; Van Maanen and Schein, 1979).

What aspects of culture matter for organizational success? Organizational culture has been characterized along multiple dimensions—for example, content, or which specific beliefs and be-

havioral norms are prevalent; and intensity, or the degree to which members are willing to sanction non-conforming and reward normatively compliant behavior—that have varying implications for firm performance (Chatman and O’Reilly, 2016). The content of cultural values, beliefs, and norms is important because it influences how people work to accomplish tasks within firms. For example, a company feeling a sense of urgency to introduce a disruptive technology might prioritize speed, autonomy, and tolerance of mistakes over other considerations. Although we acknowledge that the content of such norms and the intensity with which they are reinforced can have important consequences, our investigation focuses instead on a different cultural feature: heterogeneity, or the variety and distribution of ideas, beliefs, and normative expectations held by organizational members. Returning to the hypothetical disruptive company example, a cultural heterogeneity perspective would focus less on specific values such as speed or autonomy and more on the variety and distribution of *all* prevailing ideas and beliefs in the organization.

Our focus on heterogeneity is consistent with the distributive approach to analyzing organizational culture (Harrison and Carroll, 2006). This perspective acknowledges that culture’s role in organizational success is often idiosyncratic such that the cultural content associated with success varies by industry and a firm’s choice of competitive strategy. Moving fast may be conducive to success for an organization competing in a fast-paced and undetermined technology market but not, for example, for a hospital in which minor errors can have devastating implications. Although the effects of cultural heterogeneity on performance may also vary across competitive contexts, we anticipate that the relationship between heterogeneity and performance will be more stable and consistent across empirical settings than will the link between specific beliefs or norms and organizational performance. Thus, consistent with previous work that focuses on the distribution of culture, we propose that cultural heterogeneity has implications for firm performance that are independent of cultural content and intensity.

Heterogeneous Perspectives on Cultural Heterogeneity

Organizational theorists who study cultural heterogeneity see it as both a blessing and as a curse. Although work on organizational culture is vast and fragmented (Chatman and O'Reilly, 2016), two conflicting themes on the relationship between cultural heterogeneity and performance prevail. The first, most strongly associated with research on cultural strength, sees cultural heterogeneity as an impediment to organizational performance. This line of work conceptualizes culture as a solution to a complex coordination problem. Heterogeneous cultures, it is argued, are detrimental to organizational performance because they undermine interpersonal integration and erode internal cohesion.

Two specific mechanisms undergird the theorized link between cultural strength and organizational performance. First, cultural homogeneity is assumed to promote interpersonal coordination by facilitating goal alignment and behavioral consistency (Kreps, 1996; Gordon and DiTomaso, 1992; Weber and Camerer, 2003). A lack of such alignment can produce coordination failures. For example, an employee working in accordance with a norm that encourages speed, autonomy, and a willingness to make mistakes will coordinate less well with another employee behaving in accordance with a norm emphasizing deliberateness, caution, and precision. Second, the absence of a unified and shared culture can generate fragmentation and a sense of personal estrangement. This can lead to conflict, as well as a decline in morale and a dampening of individual commitment and motivation (Martin, 1992; Jehn et al., 1999).

In contrast, the benefits of cultural heterogeneity for organizational performance become apparent when culture is conceptualized as a set of cognitive resources that members deploy in adapting to external changes and competitive pressures. This perspective sees cultural heterogeneity as an advantage rather than an obstacle. Core to this view is the assumption that creativity—the application of novel and useful solutions to problems (Amabile, 1996, 1988)—stems from the ability to recombine existing ideas in unconventional ways (Fleming, 2001; Uzzi et al., 2013; de Vaan et al.,

2015). Drawing on these insights, researchers who emphasize the learning benefits of cultural diversity for performance argue that it does so by promoting a capacity for creative problem solving and the exploration of a broader solution space. This capacity derives not only from the fact that culturally diverse teams draw on a breadth of ideas and interpretative lenses but also from the superadditive effects of this breadth: the novelty that emerges when ideas intersect and recombine (Page, 2007; Tadmor et al., 2012b). The combination, for example, of a profit-oriented banking culture and a development-oriented social mission enabled the banks in Battilana and Dorado’s (2010) study to pioneer commercial microfinance in Bolivia in the early 1990s.

Culturally heterogeneous organizations learn more effectively by fusing different schemas, scripts, and interpretive understandings to generate novel solutions. Cultural homogeneity, in contrast, can be detrimental to a firm’s capacity for creativity for at least two reasons. First, employees in culturally uniform organizations are slower to recognize the need for change than their counterparts in organizations with diverse cultures (Lant and Mezias, 1992). Second, whereas cultural strength can foster first-order learning—for example, determining how to more efficiently execute tasks that are known to be important for marketplace success, it can inhibit second-order learning—identifying which new tasks to take on in response to a new or changing competitive landscape (Denison, 1984).

These competing perspectives on the relationship between cultural heterogeneity and performance are also echoed in research on relational demography. This line of work has mostly examined the consequences of differences in ascribed traits (such as sex, race, and age) and in backgrounds (such as functional experience, education and tenure) on group cohesion and performance (Tsui et al., 1992; Pfeffer, 1983). Research in this vein often assumes that differences at the surface level, such as race and age, also reflect disagreements at a deeper level of attitudes and beliefs. When these attitudes and beliefs relate to shared meanings, assumptions, and normative expectations, they become the cognitive material that makes up organizational culture. As Harrison et al. (1998) point out, demographic differences relate to attitudinal differences early in a group’s life but tend to wane as members coordinate their cultural orientations over time.

Demographic diversity, in other words, is not necessarily tantamount to cultural diversity. Nevertheless, work on the relationship between demographic diversity and performance is consistent with the literature on culture. On one hand, demographic differences, especially when different dimensions of categorical dissimilarity reinforce one another, foster subgroup conflict and erode group cohesion and performance (Lau and Murnighan, 1998). At the same time, under certain conditions—such as functional interdependence—cultural heterogeneity can allow group members to bring together varying perspectives and ideas in ways that can enhance creativity and innovation (Jehn et al., 1999; Van Knippenberg and Schippers, 2007).

INTERPERSONAL AND INTRAPERSONAL CULTURAL HETEROGENEITY

Taken together, these literatures suggest that cultural heterogeneity presents a fundamental trade-off: Culturally diverse firms are better at creative problem-solving, but this capability comes at the cost of decreased coordination and efficiency. Yet empirical support for the existence of such a trade-off is scant and inconclusive. Sørensen (2002), for example, theorizes that the negative effects of cultural heterogeneity on firm performance will attenuate in volatile contexts, which ought to favor firms with a greater capacity for adaptation; however, he finds inconsistent support for this contention. Similarly, Kotrba et al. (2012) report that the relationship between cultural heterogeneity and firm performance is contingent on various other cultural attributes and varies by performance indicator (e.g., market-to-book ratio versus return on assets), while Burt et al. (2000) identify additional network-based market contingencies of strong cultures.

We argue that these mixed empirical findings in part reflect a theoretical shortcoming—the assumption that cultural heterogeneity is unidimensional and arises predominantly from dissimilarities *between* people. The positive effects of heterogeneity on creativity and problem-solving, for example, are often assumed to emerge from the combination of disparate beliefs that are otherwise distributed across individuals (Page, 2007). Similarly, prevailing organizational culture constructs

tend to emphasize between-person differences. Cultural consensus, for example, refers to the degree of alignment between group members on which normative expectations matter (Chatman and O'Reilly, 2016).

Yet heterogeneity can also exist *within* individuals. For example, in team diversity research, Bunderson and Sutcliffe (2002) show that heterogeneity derives not only from differences between people—the variety of functional experts on a team—but also within individuals—the aggregate functional breadth of team members. Extending this insight to the realm of organizational culture and drawing on the idea that culture is a toolkit of multiple and potentially inconsistent cognitive resources (Swidler, 1986), we propose that heterogeneity can arise in two conceptually distinct ways. The first is the familiar between-person route: it can surface from cultural misalignment between organizational members. The second is a heretofore unexamined within-person pathway: when organizational members subscribe to multiple and potentially incompatible beliefs and values.

Although within-person heterogeneity has been mostly overlooked by the literature on organizational culture, there is considerable evidence to support its existence. Firms commonly encourage their employees to adopt a broad and potentially inconsistent set of values. Take Netflix, the on-line DVD rental and streaming service, as an example. The company's influential 126 page-long culture statement includes value statements ranging from freedom and autonomy to curiosity and responsibility.¹ Some of these values appear to be at odds with one another. For instance, Netflix emphasizes collective outcomes such as selflessness and teamwork but also the relentless pursuit of individual performance. The culture slide deck states, "We're a team...." and "You seek what is best for Netflix, rather than best for yourself..." but then goes on to issue the stern warning that, "...adequate performance gets a generous severance package."

Formal statements and espoused values are, of course, often aspirational and do not necessarily represent employees' lived experiences. No one in Enron, at least as far as we know, officially en-

¹Downloaded more than 10 million times, the slide deck was hailed by Facebook's Chief Operating Officer, Sheryl Sandberg, as the "most important document ever to come out of the valley."

dorsed malfeasance. There are nevertheless reasons to expect that, like formal mission statements, enacted culture can also encompass a multiplicity of ideas, including ones that depart from official doctrine. Support for this view comes from two interrelated bodies of research. The first, hailing from cultural sociology, argues that people are cognitively equipped to internalize and selectively deploy multiple, coinciding cultural frames. This approach conceives of culture as a loosely-held repertoire or cultural “toolkit” (Swidler, 1986). Research in cognitive and cultural psychology generally supports this conceptualization. People often hold multiple and often inconsistent cognitive schemas (DiMaggio, 1997) and are capable of identifying with multiple cultural identities (Morris et al., 2015). Different situations invoke the deployment of different cultural lenses. Participants in Swidler’s (2001) study of romantic relationships, for example, at times described their bonds through a prism of love and selfless commitment and at others emphasized their relationships’ rational and instrumental foundations. Individuals, in other words, do not necessarily subscribe to a single and internally coherent cultural order. Rather, they embody multiple cultural models that are invoked by different institutional contexts (Hallett and Ventresca, 2006; Harding, 2007).

A second line of work demonstrates that people, especially in complex modern societies, acquire a capacity for polyculturalism stemming from their chronic exposure to multiple and incongruent institutional orders. The normative assumptions governing relationships in the family, for example, are very different from those governing market transactions (Friedland and Alford, 1991). Consequently, people habitually draw symbolic boundaries between familial and economic relationships to resolve this incongruence (Zelizer, 2007).

Extending these arguments across levels of analysis, institutional sociologists have proposed that organizations, like individuals, often operate in multi-institutional environments (Boltanski and Thévenot, 2006). For example, many companies cultivate a family-like ethos but draw on a market logic to manage labor relations. Employees intersect the different cultural orders upon which such organizations are founded. Medical professionals, for example, must navigate the tensions between competing cultural logics that understand medicine either through the lens of science or care-giving.

The former emphasizes scientific authority and diagnostic success, whereas the latter conceptualizes quality health care as compassionate and preventive. Each prescribes different criteria for evaluating the legitimacy and desirability of behaviors and outcomes (Dunn and Jones, 2010).

Institutional plurality begets friction and fragmentation in organizations when different individuals subscribe to distinct institutions, adopt different identities, and see the organizational mission through internally consistent but interpersonally incongruent lenses. Organizations can overcome this tension when their members cognitively fuse these different cultural components (Battilana and Dorado, 2010; Zilber, 2002; Besharov, 2014). In fact, cultural plurality can be a source of advantage. Work by cultural sociologists finds that multivocal actors (Padgett and Ansell, 1993) and cultural objects (Griswold, 1987) derive power from their capacity to engender multiple meanings while retaining coherence.

Bringing together the two broad perspectives—one that emphasizes heterogeneity’s roots in group composition and the other that focuses on the breadth of cultural repertoires available to individual actors—we argue that cultural heterogeneity comprises two analytically distinct dimensions: interpersonal and intrapersonal. By *interpersonal heterogeneity* we refer to misalignment in cultural perceptions among the individuals who make up the organization.² By *intrapersonal heterogeneity*, we mean the breadth of cultural beliefs to which those individuals subscribe.

To illustrate this distinction, we return again to our stylized example in which there exist only two possible cultural beliefs, *A* and *B*. Figure 1 illustrates the demographic makeup of two hypothetical organizations, each represented by a circle. The organization on the left, comprising individuals who either adopt belief *A* or belief *B*, exhibits high interpersonal heterogeneity but low intrapersonal heterogeneity. Its culture exhibits low alignment across individuals, and individuals

²Interpersonal heterogeneity is related to but conceptually distinct from the construct of organizational subcultures (i.e., clusters of individuals who share similar beliefs that differ from those held by people in other clusters). In the extreme case of an organization with no interpersonal heterogeneity—i.e., where all members hold exactly the same cultural beliefs—it is, of course, not possible for subcultures to exist. Yet in organizations whose members hold divergent cultural beliefs, subcultures may or may not exist. For example, in the extreme case of an organization in which each member holds a single culture belief that no one else holds, interpersonal heterogeneity is high, but there are no subcultures. In the analyses that appear in Online Supplement B, we nevertheless include a robustness check that accounts for the presence of organizational subcultures.

themselves tend to subscribe to a smaller set of cultural beliefs. The organization on the right, in contrast, is characterized by high intrapersonal heterogeneity and low interpersonal heterogeneity. Individuals on average maintain a wide variety of beliefs about how to accomplish work in the organization and exhibit high alignment about the importance of those beliefs. Distinguishing between these two components of heterogeneity helps to uncover an important insight: organizations with broad within-person cultural repertoires need not be characterized by high levels of between-person differences.

Netflix, for example, has developed a variety of human resource practices that complement its broad cultural mission. The company places a strong emphasis on hiring and dismissal on the basis of cultural fit and, at the same time, institutes formal procedures and behavioral norms that are consistent with the mission and breadth of values it espouses (McCord, 2014). Netflix, in other words, invests in cultivating low interpersonal heterogeneity and high intrapersonal heterogeneity. Insofar as these practices are effective, they should produce a culture that is both consensual and broad.

Performance Implications of the Two Forms of Cultural Heterogeneity

Seen in this light, the trade-off between organizational coordination and problem-solving capacity no longer seems inescapable. If cultural heterogeneity comprises two dimensions—interpersonal and intrapersonal—then culture can facilitate coordination without necessarily undermining creative problem-solving and innovation. Consistent with this logic, we argue that each dimension of heterogeneity should promote different organizational outcomes.

Drawing on the literature on cultural strength, we propose that interpersonal heterogeneity will weaken an organization’s coordination and cohesion and will therefore undermine its capacity for effective execution. We therefore anticipate:

HYPOTHESIS 1 (H1): All else equal, interpersonal cultural heterogeneity will be negatively related to a firm’s capacity for efficient execution.

Interpersonal heterogeneity and its negative effects on individual and organizational performance have been extensively researched. Hypothesis 1 is broadly consistent with prior work on cultural strength and consensus in organizations. To our knowledge, however, no prior studies have systematically compared—as we do here—differences in interpersonal heterogeneity across organizations over time and studied the implications of these differences for performance.

Unlike interpersonal heterogeneity, the intrapersonal form of cultural heterogeneity has not been previously examined in organizational research. Our core argument about intrapersonal heterogeneity is that, because creativity and innovation stem from the recombination of previously unrelated ideas, a wide cultural toolkit should—all else equal—be conducive to individual creativity (Amabile, 1988). Previous work has conceived of such recombination as arising from interpersonal exchange—that is, as a function of teams rather than of individuals (de Vaan et al., 2015). Our conceptualization of intrapersonal heterogeneity instead shifts attention to the cultural resources available to individuals and the recombinant innovation they can engage in as a result. Creative recombination is, ultimately, something that individuals do (Amabile, 1988; Sauermann and Cohen, 2010), although organizations create contexts that can facilitate or inhibit this creativity (Taggar, 2002; Amabile et al., 1996).

Prior work on cultural heterogeneity has assumed that organizations enable creativity by forming culturally varied teams (e.g., Pieterse et al., 2013). We propose that organizations can also facilitate creativity—whether deliberately or organically—by increasing the representation of members who espouse a broad set of beliefs. Evidence for a link between intrapersonal heterogeneity and individual creativity comes from cultural psychology. Multi-cultural individuals, for example, exhibit a capacity for high integrative complexity and creative output by virtue of chronic exposure to different national cultures (Tadmor et al., 2012a).

We argue that the recombinant advantages of intrapersonal cultural heterogeneity should also manifest in organizations whose members have access to broad cultural toolkits. Performing complicated tasks often requires individuals to draw upon and integrate multiple and sometimes con-

flicting cultural ideas about how work is or should be done. Indeed, historical accounts of inventors and entrepreneurs often place them in cultural crossroads. For example, the hackers of the early computer industry intersected the seemingly antithetical worlds of cold war military research culture and 1960s counterculture, ushering in new technologies and organizational forms to translate these technologies into products (Turner, 2010). This ability to combine multiple perspectives was presumably also conducive to creative entrepreneurship for the machine operators in Stark’s (2011) study of a Hungarian factory in the late 1980s. Operating in a cultural context that valued technical skills but also promoted an anti-bureaucratic, relationship-oriented ethos, these factory workers were successful in forging innovative partnerships and pursuing semi-private enterprise under Hungary’s late-communist “second-economy” legislation.

As both of these examples illustrate, organizations whose members draw on broader cultural toolkits tend to exhibit a greater capacity for creativity and innovation. Thus, we hypothesize:

HYPOTHESIS 2 (H2): All else equal, intrapersonal cultural heterogeneity will be positively related to a firm’s capacity for creativity and innovation.

How organizations transform innovative ideas into positive organizational outcomes differs from how they facilitate individual creativity. The generation of novel ideas requires divergent thinking and complex integration, which, we contend, can be catalyzed by intrapersonal heterogeneity. In contrast, selecting which of these ideas to act on and doing so effectively requires other forms of information processing and problem solving capacities (Berg, 2016). We do not hypothesize about the group processes that enable teams to translate creative ideas into organizational outcomes, as those have been studied elsewhere (e.g., Woodman et al., 1993; Taggar, 2002). Rather, we posit that intrapersonal cultural heterogeneity is, *ceteris peribus*, conducive to innovative output through its relationship with individual creativity.³

³Considering how innovation relates to execution naturally raises questions of how intrapersonal and interpersonal heterogeneity might jointly, rather than independently, shape organizational outcomes. We discuss this possibility and present supplementary analyses in Online Supplement A.

LANGUAGE AS A WINDOW INTO CULTURAL HETEROGENEITY

In operationalizing the two constructs—interpersonal and intrapersonal heterogeneity—we begin with the premise that organizational culture can be detected in the language used by members (Pinker, 2007; Crémer et al., 2007). The relationship between language and culture is complex. A useful way of conceptualizing this relationship draws on the distinction between culture’s behavioral and cognitive dimensions (Mobasseri et al., forthcoming). In the former view, language can be thought of as a set of norms that facilitate interpersonal coordination and that people adhere to when they want to fit in to an organization. Weber and Camerer (2003), for example, experimentally demonstrate that linguistic conventions formed by individuals solving a coordination task increase group efficiency but also serve as an impediment when groups with different conventions are combined. In more recent work, Srivastava et al. (2018) and Goldberg et al. (2016) develop a language-based measure of cultural fit and, using an email corpus and personnel records from a mid-sized firm, demonstrate that compliance with linguistic norms is positively related to individual attainment.

Yet language can also reveal a person’s underlying beliefs and assumptions. For example, the language activists and civil society organizations used in public discourse reflected deep-current cultural shifts in Americans’ perceptions of nuclear energy in the 1980s (Gamson and Modigliani, 1989) and of Muslims in the period following the September 11th terrorist attack (Bail, 2012). Building on these insights, we propose that organizational culture can not only be detected by observing the degree of linguistic compliance that members exhibit when communicating with each other—for example, in emails or text messages—but also in the language they use to describe the organization as a whole. In particular, we focus on the topics that members use when describing their culture to each other and to outsiders. When explicitly talking about culture, organizational members consciously articulate the assumptions and beliefs they perceive are prevalent in their organization.

Whether measuring culture based on self-reports, implicit measures, or expressions of language, prior work has focused on specific categories, such as innovation, transparency, or collaboration that were predefined by researchers or informants (e.g., senior leaders in the firm) (Ehrhart and Naumann, 2004; O’Reilly et al., 1991; Srivastava and Banaji, 2011). For example, Luo et al. (2016) identify cultural topics such as “innovation” and “quality” in employee reviews of firms on Glassdoor (the same site from which our data are drawn) and show that different categories are correlated with employee satisfaction and corporate performance in different industries.

Although we acknowledge that certain cultural topics may matter more than others for success in a given industry sector, we propose that the *distribution* of cultural topics between and within individuals can nevertheless be independently related to firm profitability and innovation. The novelty of our approach is that it neither privileges one set of cultural topics over others nor assumes that researchers and informants understand the culture better than the typical organizational member does. Instead, we assume that *all* topics used in discourse about the organization’s culture are potentially informative.

Given a set of topics that organizational members use to describe culture in a given period, we define interpersonal heterogeneity as the dissimilarity of topics that group members mention in their characterizations. In other words, organizations exhibit greater interpersonal heterogeneity when their members diverge from one another in describing the culture. We define intrapersonal heterogeneity as the breadth of topics used in individual members’ cultural descriptions. Organizations exhibit greater intrapersonal heterogeneity when their members have access to and draw from a more diverse cultural toolkit.

METHOD

Data Sources and Sample

The data include all employer reviews written by employees in the United States from January 2008 to July 2015 on the website Glassdoor.⁴ Glassdoor is a career intelligence website that attracts a diverse audience primarily as a job search platform. It has an estimated 17 million unique users per month. While their identities as employees are authenticated by Glassdoor, reviewers are anonymous, thus making the reviews less susceptible to bias stemming from fear of employer retribution. Reviews are either unsolicited or contributed by users searching for jobs in exchange for unlimited site access (see Online Supplement B for details). Popadak (2013) used similar employee review text to construct longitudinal culture measures.⁵

We restricted the firm sample to: 1) publicly-traded companies for which we have access to performance data from Compustat; and 2) firms with at least 50 employee reviews in one or more quarters to ensure that there were a sufficient number of reviews to calculate our culture measures. A small number of reviews were later dropped from this sample because they did not contain at least five words that were weighted by the LDA culture model. Only firm/quarters with at least 25 reviews were used in estimated models. The resulting sample contains 512,246 reviews across 492 organizations. We lagged all predictors by one quarter to partially alleviate concerns of reverse causality and standardized the culture measures.

Measures—Dependent Variables

Our hypotheses focus on the link between cultural heterogeneity and firms' capacity for efficient execution (H1) and creativity and innovation (H2). We link the capacity for efficient execution

⁴ Accessible at www.glassdoor.com.

⁵ Popadak's (2013) measures of organizational culture focused on specific cultural features, such as integrity, detail-orientation, and adaptability, rather than the distribution of cultural content within and between employees. Nevertheless, she demonstrated that cultural attributes of organizations derived from the text of employee reviews of firms on Glassdoor are highly correlated with popular, survey-based assessments of the workplace culture. This finding helps validate the use of Glassdoor reviews as a means to assessing aspects of a firm's changing organizational culture.

to firm profitability, as measured by Return on Assets (ROA)—i.e., income before extraordinary items over total assets. We link the capacity for creativity and innovation to both the market’s expectation of a firm’s growth potential and to patenting success.

ROA is commonly understood as a measure of profitability that is indicative of a firm’s ability to effectively capitalize on its assets. We therefore use it as a reflection of a firm’s capacity for efficient execution. Creativity and innovation are more challenging to measure, and we consequently rely on two types of performance indicators: market valuation and patenting output.

We measure the market’s expectation of a firm’s growth potential using Tobin’s Q (TQ)—i.e. the market value of a firm’s assets relative to their book value. Given the forward-looking nature of the market, this measure is commonly understood as an indicator of intangible capabilities and assets—from marketing prowess to innovative IT use (Bharadwaj et al., 1999). Creativity and innovation are inherently linked to Tobin’s Q. Investors tend to reward companies that exhibit creative business strategies and that pursue technological innovation (Kogan et al., 2017). Indeed, R&D intensity and patenting success are generally positively associated with Tobin’s Q (Hall et al., 2005). Formally:

$$TQ = \frac{\text{market value of assets}}{0.9 * \text{book value of assets} + 0.1 * \text{market value of assets}} \quad (1)$$

Where market value of assets is defined as:

$$MV = \text{book assets} + (\text{market value of common equity} - \text{common equity} - \text{deferred taxes}) \quad (2)$$

While Tobin’s Q broadly captures a firm’s potential for growth and innovation as judged by the market, we also measure innovation more directly via patenting success. Patents mark the creation of new knowledge and are well established as important innovation outcomes (Lanjouw and Schankerman, 2004). First, we measure a firm’s ability to successfully produce patents by counting

the total number of patents a firm applies for within a given quarter that are later approved. Second, we measure the mean number of backward citations made by the patents in a firm’s portfolio in a given quarter, which is a common indicator of the technological importance and market value of the patents (Lanjouw and Schankerman, 2004).⁶ Patents are required to cite relevant prior art, or backward citations to existing patents upon which the focal patent builds. We logged the mean backward citations measure to account for skew.

Although patenting is a direct measure of innovation, it is not a perfect one. First, patenting captures technological innovation that can be protected as intellectual property, but a large proportion of innovation, such as marketing creativity or business strategy, is not reflected in patenting output. Second, industries vary significantly in patenting volume. While patenting is extremely common in the pharmaceutical industry, for example, it is not in retail. Our measures of patenting, particularly patenting quality, therefore only apply to a relatively small subset of companies. Overall, we evaluate H2 by examining both Tobin’s Q and patenting outcomes. Whereas the former is a broad but noisy measure of innovation and creativity, the latter provides a more direct but narrower indication of innovation.

Measures—Interpersonal and Intrapersonal Cultural Heterogeneity

We develop language-based measures of cultural heterogeneity to capture variation in interpersonal and intrapersonal heterogeneity. We measure interpersonal heterogeneity and intrapersonal heterogeneity using free response text written by employees reviewing the firm. Following prior text analysis work, we treat each review as a “bag of words,” which assumes that we can identify topical content even after discarding word order. We then represent each review as a vector of unigram counts, which identifies how many times the review includes individual words. Together, these individual words comprise a set of the most popular words that appear across the entire text corpus.

⁶We lack adequate data on forward citations, another measure of technological importance that counts the number of times that focal patents are cited by subsequent patents, given the recency of the data. This is an avenue for future research.

Our empirical strategy consists of two primary steps: 1) training a linguistic topic model to identify distinct dimensions of organizational culture mentioned in employee reviews across the entirety of the Glassdoor data; and 2) fitting that model to our analytic sample to identify the cultural dimensions mentioned in each employee review for the companies that we can track over time. We use a Latent Dirichlet Allocation (LDA) topic model (see Appendix for technical details). LDA inputs a document-term matrix, for which the rows are reviews and the columns are unigram counts, and identifies distinct topics across the corpus by observing words that tend to co-occur frequently within each review. LDA then outputs a document-topic matrix, for which each review is assigned to a probabilistic mixture of topics, or a probability distribution giving the percentages across all topics $c \in C$ that the model estimates comprise the review. (For further background and empirical examples of LDA in use, see DiMaggio et al. 2013; Mohr and Bogdanov 2013; Jha and Beckman 2017; and Kaplan and Vakili 2015.)

Training the LDA model allows us to learn what topics employees across all organizations in the Glassdoor data collectively consider germane to organizational culture. Our model training approach requires a key assumption: when employees write about firm culture, they sometimes explicitly use the word “culture” or a synonym and sometimes do not. Regardless, we can use the presence of a culture synonym as a label that indicates a given phrase contains content relevant to culture. Training the LDA model on text with these explicit references allows the model to identify a set of cultural topics. The model is then fit to reviews in our analytic sample to identify the cultural topics in text containing either explicit or implicit culture references (see Appendix for details).

The LDA model requires the researcher to choose the number of topics to output. Our goal in choosing the number of topics is not maximizing the coherence or distinctiveness of the topics. This is because we are not interested in the cultural content per se, but rather the distribution of content between and within reviewers. As such, we output a large number of topics—i.e., 500—to ensure we tease apart conceptually meaningful distinctions between cultural topics. This decision is informed

by methodological research on LDA models—above some threshold, new topics only “nibble away” at existing topics rather than fundamentally alter the topic distribution. This suggests that more topics are preferable to fewer ones given our focus on the distribution of content (Wallach et al., 2009, p. 8). Our cultural heterogeneity measures are highly-correlated and our results consistent using different numbers of topics (in particular 25, 50, 100, and 250 topics).

The topics identified by the LDA model have face validity as cultural dimensions that capture the linguistic signatures of ideas, beliefs, and normative expectations possessed by organizational members. One way to validate LDA topics is to examine the words that are most highly weighted within each topic (DiMaggio et al., 2013). Table 1 shows the highest-weighted words for four hand-picked and four randomly selected LDA topics, as well as simple labels we chose that generally capture the underlying meanings of the topics. The first set was hand-picked based on its highly distinctive culture content, and the randomly selected set is representative of average LDA topics. The LDA topics appear to be germane to organizational culture, lending support for our unsupervised learning approach.

After identifying cultural topics using this training set of phrases with explicit cultural references, we fit the LDA model to the reviews in our sample. In contrast to clustering methods, LDA is a mixed membership approach, which assigns each document to a probability distribution over multiple topics. Figure 2 illustrates LDA’s assignment of each review in the analytic sample to a mixture of multiple culture topics, represented as a topic probability distribution over the set of cultural topics. The model predicts that two reviews with similar topic probability distributions contain similar content.

We measure interpersonal heterogeneity by assessing the degree to which a firm’s employees in a given quarter characterize the firm using dissimilar cultural topics. (The Appendix provides a series of measurement validation checks.) After fitting the LDA model to the reviews in our analytic sample, each review i is represented as a probability distribution p indicating the relative proportion of each cultural topic c estimated as present in the review text.

We define interpersonal heterogeneity for a given firm/quarter as the mean Jensen-Shannon (JS) divergence between the LDA probability distributions for all unordered pairs of reviews i, j for that firm/quarter, formally:

$$A = \frac{\sum_{i,j} JS(p_i, p_j)}{\sum_{i,j}}, \text{ for all } \{i, j \mid i < j\} \quad (3)$$

where the JS-divergence between the two probability distributions is defined as:

$$JS(p_i, p_j) = \frac{1}{2}KL(p_i, M) + \frac{1}{2}KL(p_j, M) \quad (4)$$

and where $M = \frac{1}{2}(p_i + p_j)$ and $KL(p_i, M)$ is the Kullback-Leibler divergence of M from p_i :

$$KL(p_i, M) = \sum_{c \in C} p_i(c) \log_2 \frac{p_i(c)}{M(c)} \quad (5)$$

JS-divergence is a symmetric measure of the dissimilarity of two probability distributions. It is well-suited for comparing sparse, power-law distributions of words observed in natural language and has been used previously to measure the similarity of organizational members' language use (Goldberg et al., 2016; Srivastava et al., 2018).

We measure intrapersonal heterogeneity by assessing the degree to which a firm's employees discuss a broad versus narrow set of cultural topics (see Appendix). Each review i is represented as a probability distribution p indicating the relative proportion of each cultural topic c estimated as present in the review text. We apply the Herfindahl index, a popular measure of concentration, to these probability distributions, and calculate the mean Herfindahl score across all reviews for a given firm/quarter. Formally:

$$\bar{H} = \frac{\sum_i \sum_{c \in C} (p_i^c)^2}{\sum_i} \quad (6)$$

After taking the inverse, higher values indicate that employees discuss a broader range of cultural

topics, while lower values indicate a narrower, concentrated set of topics. We take the natural log of the mean Herfindahl index because the measure has a highly skewed distribution. Intrapersonal heterogeneity for a firm/quarter is formally defined as:

$$B = 1 - \ln \bar{H} \quad (7)$$

Figure 3 provides a stylized example of how our measures of interpersonal and intrapersonal heterogeneity capture systematic differences in the LDA topic probability distributions across individual reviews within a given firm-period. Panel A illustrates that firm-periods with low interpersonal heterogeneity feature reviews with more similar topic distributions. Conversely, high interpersonal heterogeneity firm-periods have reviews with more dissimilar topic distributions. Panel B illustrates that low intrapersonal heterogeneity firm-periods have reviews with more concentrated topic distributions on average, while high intrapersonal heterogeneity firm-periods have reviews with on average more uniformly distributed topic distributions.

Our language-based model of cultural heterogeneity has two advantages over survey-based culture measures. First, it allows us to measure dimensions of organizational culture longitudinally for a large, diverse set of organizations, which would be difficult using more expensive and time-intensive survey methods. Second, the model inductively identifies topics that employees consider germane to organizational culture. We do not require the researcher to make *a priori* assumptions about the cultural topics that broadly characterize organizations.

Given that the employees who wrote Glassdoor reviews were not selected through random sampling from the population of firm employees, Online Supplement B includes robustness checks to address the impact of non-random selection of employees into writing Glassdoor reviews that could bias our findings. We find no evidence that either the number or composition of reviewers systematically changes with firm performance and no evidence that the cultural heterogeneity measures themselves vary with the number or composition of reviewers.

Analytical Strategy and Estimation

We test our hypotheses with two types of models. The first is a conventional OLS model with lagged independent variables. (When estimating patent counts, we also use a negative binomial model.) In addition to our main independent variables, we control for firm size (measured as logged value of assets) and number of reviews (logged, to account for systematic variation in cultural heterogeneity measures attributable to the number of reviews with which they were produced). We also include industry and quarter fixed effects, given that performance outcomes vary significantly by industry and over time. We cluster standard errors by firm to allow for correlation between observations of the same firm over time.

A second type of model relies on Coarsened Exact Matching (CEM) (Iacus et al., 2012). We use CEM for two reasons. First, CEM allows us to (partially) address concerns about endogeneity. CEM identifies firm observations that vary on the culture variable of interest but have the same or very similar values for each control variable. Conditional on identifying all variables that affect the relationship between cultural heterogeneity and performance, CEM helps to correct for selection bias. We acknowledge, however, that the relationship between culture and firm performance is likely to be complicated and potentially bi-directional. CEM moves us closer to causal estimates; however, in the absence of exogenous variation in cultural heterogeneity or a compelling instrumental variable, we stop short of making a strong causal claim.

Second, we assume that a substantial proportion of variation in the effects of cultural heterogeneity on performance is attributable to differences between firms (for example, to organizations' different emphases on cultivating cultural consensus). Moreover, culture is known to change slowly within firms. Although our data are comprehensive, they only afford systematically observing within-firm cultural change over an extensive period of time for a subset of companies. To properly estimate between-firm differences it is therefore important that our modeling strategy account for other factors, such as industry, firm size and period, which may result in non-linear relationships

between cultural heterogeneity and performance. CEM provides such a modeling approach by allowing us to match firms that are culturally different but otherwise similar.

We implemented CEM as follows. For each culture variable (interpersonal and intrapersonal heterogeneity), we first divided observations into high and low categories on the culture measure, defined by a value above or below the industry median.⁷ We then matched these firms identified as high or low on the culture variable with others that were the same or very similar on observed characteristics to achieve covariate balance.

CEM allows us to match exactly on some covariates and coarsely on other covariates when it is infeasible to produce exact matches. We matched exactly on industry, year, and quarter. We matched coarsely on firm assets as a firm size control and on number of reviews to account for the level of coverage on the Glassdoor website. For the firm performance outcomes and patent count, coarse matches were identified using the binning algorithm default for the **cem** command in Stata.⁸ However, the default binning algorithm did not produce enough matches for the mean backward citations outcome to run a meaningful statistical analysis since that outcome is conditional on a firm having patented in a given quarter. As such, we manually instructed the command to identify matches using five equally-spaced bins with respect to firm assets, and three equally-spaced bins with respect to the number of reviews.

This procedure produced enough matches for analysis while also achieving balance on the covariates. An attractive feature of CEM is that it can produce matched strata with an unequal number of high and low culture observations so as to maximize the total number of matched observations and thus increase estimation efficiency in the subsequent analysis. The CEM algorithm produces simple weights to adjust for these differences during estimation, which we apply in all models. We also include strata fixed effects in all models—in other words, we model variation in the

⁷We identify high and low culture groups using industry medians because the distributions of interpersonal and intrapersonal heterogeneity vary substantially across some industries, where industry is defined by two-digit SIC code.

⁸Sturge’s rule is the default algorithm, which is commonly used to determine the bin width when representing a probability distribution as a histogram.

performance outcomes between high and low culture observations within each stratum of matched observations. Finally, we once again cluster standard errors by firm to allow for correlation between observations of the same firm over time.

RESULTS

Main Results

Table 2 reports univariate statistics and bivariate correlations for the final analytical sample. Interpersonal heterogeneity and intrapersonal heterogeneity have a moderately high negative correlation. Consistent with Hypothesis 1, interpersonal heterogeneity has a significant negative association with ROA. Consistent with Hypothesis 2, intrapersonal heterogeneity has a significant positive correlation with Tobin’s Q and patenting. Additionally, interpersonal heterogeneity has a moderately high positive correlation with firm size, and intrapersonal heterogeneity has a negative correlation with the number of GlassDoor reviews.

We report both OLS and CEM model results as our main findings. The matching strategy is successful in that it eliminates statistically significant differences in the observed covariates for observations with non-missing data across the four outcome variables. Table 3 shows t-tests on covariate means before versus after matching for both the interpersonal heterogeneity and intrapersonal heterogeneity matching across the four outcomes. Any large t-statistics for firm assets and number of Glassdoor reviews that exist before matching are sharply reduced to the point of non-significance after matching.

Table 4 shows regression results for models of ROA on interpersonal heterogeneity as tests of Hypothesis 1. Model 1 shows that interpersonal heterogeneity has a significantly negative association with ROA in an OLS specification with quarter and industry fixed effects. Model 2 shows that the negative association holds in the matched sample. Thus, we find support for Hypothesis 1: firms higher in interpersonal heterogeneity exhibit lower profitability.

Table 5 shows regressions results for models of Tobin’s Q and the patenting outcomes on intrapersonal heterogeneity. Model 1 shows that intrapersonal heterogeneity has a significantly positive association with Tobin’s Q in an OLS specification with quarter and industry fixed effects. This positive association holds in Model 2 using the matched sample. Model 3 shows that higher intrapersonal heterogeneity predicts greater patenting volume in an OLS specification. We also modeled the raw patent count in a negative binomial specification in Model 4, which produces consistent results. Model 5 shows that this effect of intrapersonal heterogeneity on patenting volume holds with the matched sample. We find in Model 6 that higher intrapersonal heterogeneity is associated with higher mean backward citations among the patents in a firm portfolio. This association holds in Model 7 using the matched sample. Together, these results provide support for Hypothesis 2: firms higher in intrapersonal heterogeneity exhibit a greater capacity for creativity and innovation as reflected in their increased market valuation, larger patenting volume and higher patent quality.

Figure 4 provides a visualization of the predicted effects of interpersonal and intrapersonal heterogeneity on the firm performance and patent outcomes for our CEM models. A one standard deviation increase in interpersonal heterogeneity reduces ROA by approximately 0.35%, which is equivalent to 0.13 standard deviations (Panel A). A one standard deviation increase in intrapersonal heterogeneity increases Tobin’s Q by approximately 0.25 points, or 0.15 standard deviations (Panel B); increases patent count by roughly 0.2, or approximately 0.07 standard deviations (Panel C); and increases mean backward citations by approximately 0.28, or 0.25 standard deviations (Panel D).

Online Supplement A describes a series of alternative models that we use to explore the interrelationships between the two heterogeneity measures. We show that our main results for each heterogeneity measure hold when we control for the other measure and find no evidence of cross-effects (i.e., interpersonal heterogeneity affecting growth/innovation or intrapersonal heterogeneity influencing productivity). Second, although we do not find evidence of a linear interaction effect between the heterogeneity measures, we present some preliminary findings based on indicator vari-

ables that point to possible non-linear interrelationships, suggesting that this may be a promising avenue for future research.

Robustness Checks

We conducted four additional analyses to assess the robustness of our findings. First, to ensure that our measure of interpersonal heterogeneity is not simply a proxy for the presence of organizational subcultures, we used a clustering procedure to assess the extent to which each organization-quarter observation can be divided into cultural subgroups. The clustering procedure uses the K-means algorithm, based on our measure of cultural distance, to divide the population into clusters and then uses the gap statistic (Tibshirani et al., 2001) to assess the optimal division into subgroups. Including a control for the presence of subgroups did not materially change our results.

Second, given that both of our heterogeneity measures could be related to the overall level of cultural breadth in the organization, we used a variant of our intrapersonal heterogeneity measure that assesses the dispersion of topics discussed across *all* employee reviews. Specifically, we calculate the (logged) mean Herfindahl index across all employee reviews in a given firm/quarter and then take the inverse. High values indicate that employees discuss a broad range of cultural topics, while low values suggest a narrow range. Controlling for firm-level cultural breadth does not materially alter our results.

Third, to establish that our results are not driven by specific cultural categories and that our measure is robust to the selection of fewer than 500 topics, we ran a simulation analysis, which reveals a high degree of correspondence between measures based on randomly selected subsets of topics and measures based on our 500 topic model. These three analyses are presented in Online Supplement B.

Finally, we undertook a manual coding exercise using a 100 topic model and demonstrate in Online Appendix C that our algorithmically-derived measures and results are nearly identical to those obtained via manual coding.

DISCUSSION AND CONCLUSION

This article has sought to bring conceptual clarity and empirical evidence on a longstanding question in organizational research: when is a diversity of ideas and beliefs beneficial for organizational success and when is it instead detrimental? From a perspective that understands organizations as solutions to complex coordination problems, cultural heterogeneity is mostly seen as a source of dissonance and friction. In contrast, through the lens of economic and cultural sociology, it is viewed as a necessary condition for creativity and innovation.

We propose that this tension can be resolved when we consider the duality in diversity—that it can take two analytically distinct forms. Prior research has emphasized heterogeneity between organizational members but mostly overlooked the heterogeneity that can exist *within* individuals. We conceptually separate the two forms and demonstrate that interpersonal heterogeneity—the extent to which organizational members diverge in their understanding of firm culture—is negatively associated with effective coordination and execution, whereas intrapersonal heterogeneity—the breadth of cultural beliefs about the organization that are held by members—is positively linked to creativity and the capacity for recombinant innovation. We further propose that the two forms of heterogeneity will be tied to different organizational outcomes and report evidence of a negative link between interpersonal heterogeneity and a firm’s subsequent return on assets, as well as a positive relationship between intrapersonal heterogeneity and a firm’s Tobin’s Q and patenting activity.

Organizational Culture and Firm Performance

Findings from this investigation make several noteworthy contributions to the extensive and multidisciplinary literature that examines the link between organizational culture and firm performance (e.g., Sørensen, 2002; Van den Steen, 2010; Kotter and Heskett, 1992). First, whereas much of this work has examined how specific cultural content—for example, norms of adaptability—relate to outcomes such as cash flow growth (Chatman et al., 2014), we demonstrate that the distribution

of culture within an organization can have performance consequences that are independent of the organization’s specific cultural content (cf. Carroll and Harrison, 1998). Our findings neither invalidate nor downplay the importance of studies that focus on specific organizational culture features. Yet the cultural content that matters for organizational success is known to vary considerably across competitive contexts (such as industry, geography, or regulatory environment). In contrast, a distributive approach to measuring culture such as ours uncovers linkages between culture and performance that would appear to be more generalizable across competitive contexts and over time.

Drawing on the sociological conceptualization of culture as toolkit, we introduce to this literature a novel organizational culture construct: intrapersonal heterogeneity. Prevailing approaches to studying organizational culture have exclusively explored cultural heterogeneity through an interpersonal lens. Constructs such as cultural strength focus on the degree of normative consensus between organizational members and the intensity with which these norms are enforced interpersonally (Chatman and O’Reilly, 2016). In other words, cultural heterogeneity is assumed to inhere in differences between individuals. Heterogeneity within individuals has been largely overlooked by this literature. Indeed, existing measures and analytical models of cultural heterogeneity are not designed to provide the data that would be needed to measure within-person heterogeneity. For example, the Organizational Culture Profile—a widely used instrument for assessing cultural features—requires respondents to array cultural norms on a normal-like scale (O’Reilly et al., 1991). It therefore assumes a fixed within-person distribution of culture. Similarly, formal models of cultural distribution in organizations (e.g. Harrison and Carroll, 2006; March, 1991) often focus on interpersonal transmission, reducing culture to a unidimensional scale.

We identify a complementary dimension of heterogeneity—that which exists within people—and demonstrate that considering the two forms in tandem helps to resolve the ambiguity that previously existed in the literature about how cultural heterogeneity relates to firm performance. Indeed, our findings suggest that the heterogeneity tradeoff assumed in previous research—that heterogeneity harms execution and productivity but helps creativity and innovation—may, in fact,

be escapable. Insofar as firms can increase intrapersonal heterogeneity while holding interpersonal heterogeneity constant, our theory predicts that they will achieve greater innovation output without sacrificing short-term profitability. Conversely, if firms can decrease interpersonal heterogeneity without eroding its intrapersonal counterpart, we anticipate that they can improve profitability without compromising innovativeness. Exploratory findings reported in Online Supplement A seem to point in this direction.

The identification of two analytically distinct forms of cultural heterogeneity naturally raises questions about how deliberative organizational change efforts influence each dimension and how the two relate to each other over time. An extensive literature has explored the organizational determinants of strong, interpersonally homogeneous cultures. In contrast, because intrapersonal heterogeneity is a novel construct, its antecedents remain unclear. Nevertheless, previous work suggests that disparate cultural logics can be fused together when practices embody multiple meanings (e.g., Zilber, 2002) and that organizations can influence this process through hiring and socialization (e.g., Battilana and Dorado, 2010). It remains to be studied which organizational processes are effective at facilitating intrapersonal heterogeneity and how intentional attempts to broaden organizational members cultural toolkits affect interpersonal heterogeneity. Might a deliberate focus on increasing within-person heterogeneity provide greater scope for shared understanding between people or instead sow interpersonal cultural discord? Conversely, how do managerial efforts aimed at diversifying the cultural makeup of a company's workforce affect intrapersonal heterogeneity?

Intrapersonal cultural heterogeneity, however, is not exclusively the purview of intentional managerial policies inside organizations. Rather, we expect that the breadth of individuals' cultural toolkits is shaped by their social experiences such that the diversity of social worlds an individual inhabits—whether by virtue of her personal history or network position—is positively related to her intrapersonal cultural diversity. This suggests that organizations play an important societal role in bridging or reinforcing cultural boundaries between different social groups, especially if increased labor market mobility is leading employees toward greater exposure to cultural multiplicity. Under

some circumstances such exposure begets intrapersonal cultural breadth (Morris et al., 2015), but under others it serves to narrow cultural repertoires and entrench cultural divisions (Goldberg and Stein, 2018; Bail et al., 2018).

In the future, questions such as these will be easier to answer because of the methodological innovation we introduce: using unsupervised learning to identify cultural content in the language employees use to describe their organizations and deriving time-varying measures of organizational culture based on this language. Unlike traditional survey-based measures of culture, language-based measures of culture can be produced unobtrusively on a continuous basis and at scale. Whereas recent work in this vein has mined internal employee communications to characterize how individuals fit in culturally within their organization (Goldberg et al., 2016; Srivastava et al., 2018; Doyle et al., 2017), the approach we develop here enables us to characterize the culture of the organization as a whole, to make comparisons across organizations, and to systematically track distributive cultural change within and across organizations.

Culture as Toolkits

Our work also injects greater theoretical precision to, and provides empirical support for, the theory of culture as toolkit (Swidler, 1986), which is one of the most influential perspectives in contemporary cultural sociology but also one that has been criticized as vague and susceptible to slippage in terminology (Lamont, 1992; Small et al., 2010). First, our intrapersonal heterogeneity construct represents a concrete manifestation of the somewhat hazy “symbols, stories, rituals, and worldviews” that constitute toolkits in Swidler’s (1986: 273) theory. Although we acknowledge that toolkits may include a broader set of implements, we propose that a useful way to conceptualize toolkits—at least in the organizational context—is to focus on the breadth of topics employees draw upon in describing their organization’s culture.

Second, our approach offers a way to extend cultural toolkits to organizations and the fields in which they are embedded (Weber, 2005). Our method can be readily adapted to characterizing

the (changing) distance between firms in the space of cultural topics. Complementing cultural compatibility analyses based on surveys (Stahl and Voigt, 2008), formal models (Van den Steen, 2010), and laboratory experiments (Weber and Camerer, 2003), language-based measures of cultural similarity between firms could be used to examine how culture influences success or failure at the interorganizational level—for example, in mergers and acquisitions, (Bauer and Matzler, 2014; Stahl and Voigt, 2008; Van den Steen, 2010; Weber and Camerer, 2003), joint ventures, and alliances (Park and Ungson, 1997; Pothukuchi et al., 2002). Such measures could also be used to examine how the cultural distinctiveness or similarity of an organization relates to its strategic positioning and competitive advantage. Moreover, our measures can be aggregated to the field or industry level as a means to explore how toolkit breadth and composition relate to field-level dynamics.

Finally, whereas toolkit theory focuses on the repertoire of cultural resources individuals draw upon to construct strategies of action, we make the micro-to-macro link between individual toolkits, group creativity, and organizational-level outcomes. While the image of culture as toolkit has been widely influential in sociology, few studies have examined the relationship between the breadth of cultural toolkits and group-level outcomes. We argue that individuals with broad cultural repertoires are more likely to engage in recombinant innovation and demonstrate that this breadth is linked to organizational creativity and innovation. While our empirical examination focuses on for-profit organizations, we see no reason why it should not extend to other domains. To the same extent that community-level ethnic integration relates to entrepreneurship and innovation (e.g. Samila and Sorenson, 2017), our findings suggest that communities—whether at the local or national level—whose members draw on a wide range of cultural elements will exhibit, all else equal, a greater capacity for innovation. Communities high in intrapersonal but low in interpersonal cultural heterogeneity might enjoy diversity’s benefits, such as enhanced resilience, while avoiding the risk of coordination breakdowns and inefficiencies (Page, 2010).

More broadly, high intrapersonal but low interpersonal heterogeneity might allow nations to leverage the benefits of multiculturalism, while still preserving an overarching sense of cultural

cohesion and unity. For example, American society might not face a strict trade off between being a cultural “melting pot,” in which assimilation forces foster cohesion but stifle diversity, versus being a multicultural “salad bowl” that preserves diversity but reifies cultural boundaries (Fischer and Mattson, 2009). The realization of cultural diversity’s potential for macro-level innovation, however, might overall be stymied if people are increasingly engaging with cultural echo chambers that limit their exposure to cultural multiplicity, thereby narrowing their cultural toolkits (DellaPosta et al., 2015).

Cultural Heterogeneity and Group Diversity

Although our conceptual arguments are not directly derived from research on relational demography and group diversity, we identify a number of important parallels between the two bodies of work. First, our distinction between the interpersonal and intrapersonal forms of cultural heterogeneity echoes Bunderson and Sutcliffe’s (2002) separation of dominant function diversity (a between-person comparison of functional expertise on a team) from intrapersonal functional diversity (based on within-person functional breadth). Our theory applies a similar distinction to cultural diversity in organizations.

At the same time, the constructs we have developed at the organizational level have implications for research on group effectiveness. Prior work focusing on cultural diversity in groups and teams has examined surface-level traits such as race and nationality, as well as deep-level values that are measured using proxies such as the national-cultural distance between group members (Stahl et al., 2010). As this work has demonstrated, categorical differences between individuals on dimensions such as gender or occupational background can, but often do not, relate to differences in underlying beliefs (Harrison et al., 1998). The constructs we introduce offer the potential to more directly assess cultural diversity in groups. It remains to be explored how intrapersonal and interpersonal heterogeneity in work groups might relate to mechanisms such as group conflict, creativity, and social integration through which diversity—both ascribed and internalized—affects

team performance.

Limitations and Future Directions

Given the nature of the data we analyze, this study has certain limitations, which also point to avenues for future research. First, the coarsened exact matching approach we use (Iacus et al., 2012) achieves balance between our treatment and control groups by matching on observed firm attributes. It does not, however, address potential threats to causal identification stemming from unobserved heterogeneity. Future research in this vein—especially studies that draw on data sets spanning longer time horizons and thus affording a window into changing firm cultures—could account for time-invariant unobserved heterogeneity by estimating within-firm models. Over the time horizon of our data set, employee descriptions of firm culture in Glassdoor reviews simply do not exhibit sufficient temporal variance to support the use of within-firm estimates.

Also, although we report robustness checks that help to dispel concerns that our findings can be accounted for by compositional shifts in the kinds of employees who choose to comment about firm culture prior to changes in firm performance, we cannot fully rule out the potentially confounding role of selection effects. We leave to future research the task of more thoroughly accounting for selection dynamics in employee reviews. For example, researchers could draw on national survey panels to identify a representative set of employees at firms included in the Glassdoor data and ask them to rate their firm using the same pro and con questions used by Glassdoor.

In addition, any method of measuring culture necessarily makes simplifying assumptions about a complex, multifaceted phenomenon in order to gain analytical tractability. Our LDA topic modeling approach assumes that word order is negligible and infers the existence of cultural topics based on the co-occurrence of words within employee reviews. This method may obscure some more nuanced cultural meanings residing in the ordering of words and the relationships between specific words in sentences. Future research might look to alternative natural language processing tools, such as word embedding models, to detect cultural meanings overlooked by topic modeling.

Conclusion

This study paves the way for novel investigations of the role of culture in organizational performance. Drawing on language as a window into organizational culture, it demonstrates that cultural heterogeneity can be a double-edged sword, with its interpersonal form foreshadowing a decline in profitability and its intrapersonal form heralding heightened market expectations of future firm growth and innovative output. At the same time, it highlights that the coordination-creativity trade-off that culture presumably embodies is not inexorable. Understanding the duality in cultural diversity may help uncover organizational practices and interventions that simultaneously promote efficiency and innovation.

References

- Amabile, T. M.
1988 "A model of creativity and innovation in organizations." *Research in Organizational Behavior*, 10: 123–167.
- Amabile, T. M.
1996 *Creativity in Context: Update to the Social Psychology of Creativity*. Boulder, CO: Westview.
- Amabile, T. M., R. Conti, H. Coon, J. Lazenby, and M. Herron
1996 "Assessing the work environment for creativity." *Academy of Management Journal*, 39: 1154–1184.
- Ashford, S. and S. Nurmohamed
2012 "From past to present and into the future: A hitchhiker's guide to the socialization literature.", In Connie R. Wanberg (Ed.), *The Oxford Handbook of Organizational Socialization*, pp. 8–24 Oxford: Oxford University Press.
- Bail, C. A.
2012 "The fringe effect: Civil society organizations and the evolution of media discourse about Islam since the September 11th attacks." *American Sociological Review*, 77: 855–879.
- Bail, C. A., L. P. Argyle, T. W. Brown, J. P. Bumpus, H. Chen, M. B. F. Hunzaker, J. Lee, M. Mann, F. Merhout, and A. Volfovsky
2018 "Exposure to opposing views on social media can increase political polarization." *Proceedings of the National Academy of Sciences*, 115(37): 9216–9221.
- Battilana, J. and S. Dorado
2010 "Building sustainable hybrid organizations: The case of commercial microfinance organizations." *Academy of Management Journal*, 53: 1419–1440.
- Bauer, F. and K. Matzler
2014 "Antecedents of M&A success: The role of strategic complementarity, cultural fit, and degree and speed of integration." *Strategic Management Journal*, 35: 269–291.
- Berg, J. M.
2016 "Balancing on the creative highwire: Forecasting the success of novel ideas in organizations." *Administrative Science Quarterly*, 61: 433–468.
- Besharov, M. L.
2014 "The relational ecology of identification: How organizational identification emerges when individuals hold divergent values." *Academy of Management Journal*, 57: 1485–1512.
- Bharadwaj, A. S., S. G. Bharadwaj, and B. R. Konsynski
1999 "Information technology effects on firm performance as measured by Tobin's Q." *Management Science*, 45: 1008–1024.

Blei, D. M., A. Y. Ng, and M. I. Jordan

2003 “Latent dirichlet allocation.” *Journal of Machine Learning Research*, 3: 993–1022.

Boltanski, L. and L. Thévenot

2006 *On Justification: Economies of Worth*. Princeton, NJ: Princeton University Press.

Bunderson, J. S. and K. M. Sutcliffe

2002 “Comparing alternative conceptualizations of functional diversity in management teams: Process and performance effects.” *Academy of Management Journal*, 45: 875–893.

Burt, R. S., M. Guilarte, H. J. Raider, and Y. Yasuda

2000 “Competition, contingency, and the external structure of markets.”, In Ingram, Paul and Silverman, Brian S. (Eds.), *The New Institutionalism in Strategic Management*, pp. 167–217 Emerald Group Publishing Limited.

Carroll, G. R. and J. R. Harrison

1998 “Organizational demography and culture: Insights from a formal model and simulation.” *Administrative Science Quarterly*, 43: 637–667.

Chatman, J. A., D. F. Caldwell, C. A. O’Reilly, and B. Doerr

2014 “Parsing organizational culture: How the norm for adaptability influences the relationship between culture consensus and financial performance in high-technology firms.” *Journal of Organizational Behavior*, 35: 785–808.

Chatman, J. A. and C. A. O’Reilly

2016 “Paradigm lost: Reinvigorating the study of organizational culture.” *Research in Organizational Behavior*, 36: 199–224.

Crémer, J., L. Garicano, and A. Prat

2007 “Language and the theory of the firm.” *The Quarterly Journal of Economics*, 122: 373–407.

DellaPosta, D., Y. Shi, and M. Macy

2015 “Why do liberals drink lattes?” *American Journal of Sociology*, 120(5): 1473–1511.

Denison, D. R.

1984 “Bringing corporate culture to the bottom line.” *Organizational Dynamics*, 13: 5–22

Denison, D. R. and A. K. Mishra

1995 “Toward a theory of organizational culture and effectiveness.” *Organization Science*, 6: 204–223.

DiMaggio, P.

1997 “Culture and cognition.” *Annual Review of Sociology*, 23: 263–287.

DiMaggio, P., M. Nag, and D. Blei

2013 “Exploiting affinities between topic modeling and the sociological perspective on culture: Application to newspaper coverage of U.S. government arts funding.” *Poetics*, 41: 570–606.

Doyle, G., A. Goldberg, S. Srivastava, and M. Frank

2017 “Alignment at work: Using language to distinguish the internalization and self-regulation components of cultural fit in organizations.” In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics*. vol. 1, pp. 603–612.

Dunn, M. B. and C. Jones

2010 “Institutional logics and institutional pluralism: The contestation of care and science logics in medical education, 1967–2005.” *Administrative Science Quarterly*, 55: 114–149.

Ehrhart, M. G. and S. E. Naumann

2004 “Organizational citizenship behavior in work groups: A group norms approach.” *Journal of Applied Psychology*, 89: 960–974.

Fiol, C. M.

1994 “Consensus, diversity, and learning in organizations.” *Organization Science*, 5: 403–420.

Fischer, C. S. and G. Mattson

2009 “Is America Fragmenting?” *Annual Review of Sociology*, 35: 435–455.

Fleming, L.

2001 “Recombinant uncertainty in technological search.” *Management Science*, 47: 117–132.

Friedland, R. and R. R. Alford

1991 “Bringing society back in: Symbols, practices and institutional contradictions.”, In Walter W. Powell and Paul J. DiMaggio (Eds.), *The New Institutionalism in Organizational Analysis* Chicago: University of Chicago Press.

Gamson, W. A. and A. Modigliani

1989 “Media discourse and public opinion on nuclear power: A constructionist approach.” *American Journal of Sociology*, 95: 1–37.

Goldberg, A., S. B. Srivastava, V. G. Manian, and C. Potts

2016 “Fitting in or standing out? The tradeoffs of structural and cultural embeddedness.” *American Sociological Review*, 81: 1190–1222.

Goldberg, A. and S. K. Stein

2018 “Beyond Social Contagion: Associative Diffusion and the Emergence of Cultural Variation.” *American Sociological Review*, 83(5): 897–932.

Gordon, G. G. and N. DiTomaso

1992 “Predicting corporate performance from organizational culture.” *Journal of Management Studies*, 29: 783–798

Griswold, W.

1987 "The fabrication of meaning: Literary interpretation in the United States, Great Britain, and the West Indies." *American Journal of Sociology*, 92: 1077–1117.

Hall, B. H., A. Jaffe, and M. Trajtenberg

2005 "Market value and patent citations." *RAND Journal of Economics*, pp. 16–38.

Hallett, T. and M. J. Ventresca

2006 "Inhabited institutions: Social interactions and organizational forms in Gouldner's Patterns of Industrial Bureaucracy." *Theory and Society*, 35: 213–236.

Harding, D. J.

2007 "Cultural context, sexual behavior, and romantic relationships in disadvantaged neighborhoods." *American Sociological Review*, 72: 341–364.

Harrison, D. A., K. H. Price, and M. P. Bell

1998 "Beyond relational demography: Time and the effects of surface- and deep-level diversity on work group cohesion." *Academy of Management Journal*, 41: 96–107.

Harrison, J. R. and G. R. Carroll

2006 *Culture and Demography in Organizations*. Princeton, NJ: Princeton University Press.

Iacus, S. M., G. King, G. Porro, and J. N. Katz

2012 "Causal inference without balance checking: Coarsened exact matching." *Political Analysis*, pp. 1–24

Jehn, K. A., G. B. Northcraft, and M. A. Neale

1999 "Why differences make a difference: A field study of diversity, conflict and performance in workgroups." *Administrative Science Quarterly*, 44: 741–763.

Jha, H. K. and C. M. Beckman

2017 "A patchwork of identities: Emergence of charter schools as a new organizational form.", In Seidel, Marc-David L. and Greve, Henrich R. (Eds.), *Emergence (Research in the Sociology of Organizations, Volume 50)*, pp. 69–107 Emerald Publishing Limited.

Kaplan, S. and K. Vakili

2015 "The double-edged sword of recombination in breakthrough innovation." *Strategic Management Journal*, 36: 1435–1457.

Kogan, L., D. Papanikolaou, A. Seru, and N. Stoffman

2017 "Technological innovation, resource allocation, and growth." *The Quarterly Journal of Economics*, 132: 665–712.

Kotrba, L. M., M. A. Gillespie, A. M. Schmidt, R. E. Smerek, S. A. Ritchie, and D. R. Denison

2012 "Do consistent corporate cultures have better business performance? Exploring the interaction effects." *Human Relations*, 65: 241–262.

- Kotter, J. P. and J. L. Heskett
1992 *Corporate Culture and Performance*. New York: Free Press.
- Kreps, D. M.
1996 “Corporate culture and economic theory.”, In Peter J. Buckley and Jonathan Michie (Eds.), *Firms, Organizations and Contracts*, pp. 221–275 Oxford, UK: Oxford University Press.
- Kunda, G.
2009 *Engineering Culture: Control and Commitment in a High-Tech Corporation*. Philadelphia, PA: Temple University Press.
- Lamont, M.
1992 *Money, Morals, and Manners: The Culture of the French and the American Upper-Middle Class*. Chicago, IL: University of Chicago Press.
- Lanjouw, J. O. and M. Schankerman
2004 “Patent quality and research productivity: Measuring innovation with multiple indicators.” *The Economic Journal*, 114: 441–465.
- Lant, T. K. and S. J. Mezias
1992 “An organizational learning model of convergence and reorientation.” *Organization Science*, 3: 47–71
- Lau, D. C. and J. K. Murnighan
1998 “Demographic diversity and faultlines: The compositional dynamics of organizational groups.” *Academy of Management Review*, 23: 325–340.
- Luo, N., Y. Zhou, and J. Shon
2016 “Employee satisfaction and corporate performance: Mining employee reviews on Glassdoor.com.” Unpublished Working Paper.
- March, J. G.
1991 “Exploration and exploitation in organizational learning.” *Organization Science*, 2: 71–87.
- Martin, J.
1992 *Cultures in Organizations: Three Perspectives*. Oxford, UK: Oxford University Press.
- McCord, P.
2014 “How Netflix reinvented HR.” *Harvard Business Review*, 92: 70–76.
- Mobasser, S., A. Goldberg, and S. B. Srivastava
forthcoming “What is cultural fit? From cognition to behavior (and back).”, In Wayne Brekhus and Gabe Ignatow (Eds.), *The Oxford Handbook of Cognitive Sociology* Oxford: Oxford University Press.

Mohr, J. W. and P. Bogdanov

2013 “Introduction—topic models: What they are and why they matter.” *Poetics*, 41: 545–569.

Morris, M. W., C.-y. Chiu, and Z. Liu

2015 “Polycultural psychology.” *Annual Review of Psychology*, 66: 631–659.

O’Reilly, C. A., J. Chatman, and D. F. Caldwell

1991 “People and organizational culture: A profile comparison approach to assessing person-organization fit.” *Academy of Management Journal*, 34: 487–516.

Padgett, J. F. and C. K. Ansell

1993 “Robust action and the rise of the Medici, 1400-1434.” *American Journal of Sociology*, 98: 1259–1319.

Page, S. E.

2007 *The Difference: How the Power of Diversity Creates Better Groups, Firms, Schools, and Societies*. Princeton, NJ: Princeton University Press.

Page, S. E.

2010 *Diversity and complexity* Princeton University Press.

Park, S. H. and G. R. Ungson

1997 “The effect of national culture, organizational complementarity, and economic motivation on joint venture dissolution.” *Academy of Management Journal*, 40: 279–307 0001–4273.

Pettigrew, A. M.

1979 “On studying organizational cultures.” *Administrative Science Quarterly*, 24: 570–581.

Pfeffer, J.

1983 “Organizational demography.”, In Cummins, Larry L. and Staw, Barry M. (Eds.), *Research in Organizational Behavior*, vol. 5, pp. 299–357 JAI Press, Inc.

Pieterse, A. N., D. Van Knippenberg, and D. Van Dierendonck

2013 “Cultural diversity and team performance: The role of team member goal orientation.” *Academy of Management Journal*, 56: 782–804.

Pinker, S.

2007 *The Stuff of Thought: Language as a Window into Human Nature*. New York: Viking.

Popadak, J. A.

2013 “A corporate culture channel: How increased shareholder governance reduces firm value.” Unpublished Working Paper.

Pothukuchi, V., F. Damanpour, J. Choi, C. C. Chen, and S. H. Park

2002 “National and organizational culture differences and international joint venture performance.” *Journal of International Business Studies*, 33: 243–265

Samila, S. and O. Sorenson

2017 “Community and capital in entrepreneurship and economic growth.” *American Sociological Review*, 82: 770–795.

Sauermann, H. and W. M. Cohen

2010 “What makes them tick? Employee motives and firm innovation.” *Management Science*, 56: 2134–2153.

Schein, E. H.

2010 *Organizational Culture and Leadership*. San Francisco, CA: Jossey-Bass, 4th edition ed.

Small, M. L., D. J. Harding, and M. Lamont

2010 “Reconsidering culture and poverty.” *ANNALS of the American Academy of Political and Social Science*, 629: 6–27.

Sørensen, J. B.

2002 “The strength of corporate culture and the reliability of firm performance.” *Administrative Science Quarterly*, 47: 70–91.

Srivastava, S. B. and M. R. Banaji

2011 “Culture, cognition, and collaborative networks in organizations.” *American Sociological Review*, 76: 207–233.

Srivastava, S. B., A. Goldberg, V. G. Manian, and C. Potts

2018 “Enculturation trajectories: Language, cultural adaptation, and individual outcomes in organizations.” *Management Science*, 64: 1348–1364.

Stahl, G. K., M. L. Maznevski, A. Voigt, and K. Jonsen

2010 “Unraveling the effects of cultural diversity in teams: A meta-analysis of research on multicultural work groups.” *Journal of International Business Studies*, 41: 690–709.

Stahl, G. K. and A. Voigt

2008 “Do Cultural Differences Matter in Mergers and Acquisitions? A Tentative Model and Examination.” *Organization Science*, 19: 160–176

Stark, D.

2011 *The Sense of Dissonance: Accounts of Worth in Economic Life*. Princeton, NJ: Princeton University Press.

Van den Steen, E.

2010 “Culture clash: The costs and benefits of homogeneity.” *Management Science*, 56: 1718–1738.

Swidler, A.

1986 “Culture in action: Symbols and strategies.” *American Sociological Review*, 51: 273–286.

Swidler, A.

2001 *Talk of Love: How Culture Matters*. Chicago, IL: University of Chicago Press.

Tadmor, C. T., A. D. Galinsky, and W. W. Maddux

2012a “Getting the most out of living abroad: Biculturalism and integrative complexity as key drivers of creative and professional success.” *Journal of Personality and Social Psychology*, 103: 520–542.

Tadmor, C. T., P. Satterstrom, S. Jang, and J. T. Polzer

2012b “Beyond individual creativity: The superadditive benefits of multicultural experience for collective creativity in culturally diverse teams.” *Journal of Cross-Cultural Psychology*, 43: 384–392.

Taggar, S.

2002 “Individual creativity and group ability to utilize individual creative resources: A multilevel model.” *Academy of Management Journal*, 45: 315–330.

Tibshirani, R., G. Walther, and T. Hastie

2001 “Estimating the number of clusters in a data set via the gap statistic.” *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 63: 411–423.

Tsui, A. S., T. D. Egan, and C. A. O'Reilly III

1992 “Being different: Relational demography and organizational attachment.” *Administrative Science Quarterly*, 37: 549–579.

Turco, C.

2016 *The Conversational Firm: Rethinking Bureaucracy in the Age of Social Media*. New York: Columbia University Press.

Turner, F.

2010 *From Counterculture to Cyberculture: Stewart Brand, the Whole Earth Network, and the Rise of Digital Utopianism*. Chicago, IL: University of Chicago Press.

Uzzi, B., S. Mukherjee, M. Stringer, and B. Jones

2013 “Atypical combinations and scientific impact.” *Science*, 342: 468–472.

de Vaan, M., B. Vedres, and D. Stark

2015 “Game changer: The topology of creativity.” *American Journal of Sociology*, 120: 1144–1194.

Van Knippenberg, D. and M. C. Schippers

2007 “Work group diversity.” *Annual Review of Psychology*, 58: 515–541.

Van Maanen, J. and E. H. Schein

1979 “Toward of theory of organizational socialization.” *Research in Organizational Behavior*, 1: 209–264.

Wallach, H. M., D. M. Mimno, and A. McCallum

2009 “Rethinking LDA: Why priors matter.”, In Y. Bengio and D. Schuurmans and J. D. Lafferty and C. K. I. Williams and A. Culotta (Eds.), *Advances in Neural Information Processing Systems* 22, pp. 1973–1981 Curran Associates, Inc.

Weber, K.

2005 “A toolkit for analyzing corporate cultural toolkits.” *Poetics*, 33: 227–252.

Weber, R. A. and C. F. Camerer

2003 “Cultural conflict and merger failure: An experimental approach.” *Management Science*, 49: 400–415.

Woodman, R. W., J. E. Sawyer, and R. W. Griffin

1993 “Toward a theory of organizational creativity.” *Academy of Management Review*, 18: 293–321.

Zelizer, V. A.

2007 *The Purchase of Intimacy*. Princeton, NJ: Princeton University Press.

Zilber, T. B.

2002 “Institutionalization as an interplay between actions, meanings, and actors: The case of a rape crisis center in Israel.” *Academy of Management Journal*, 45: 234–254.

FIGURES

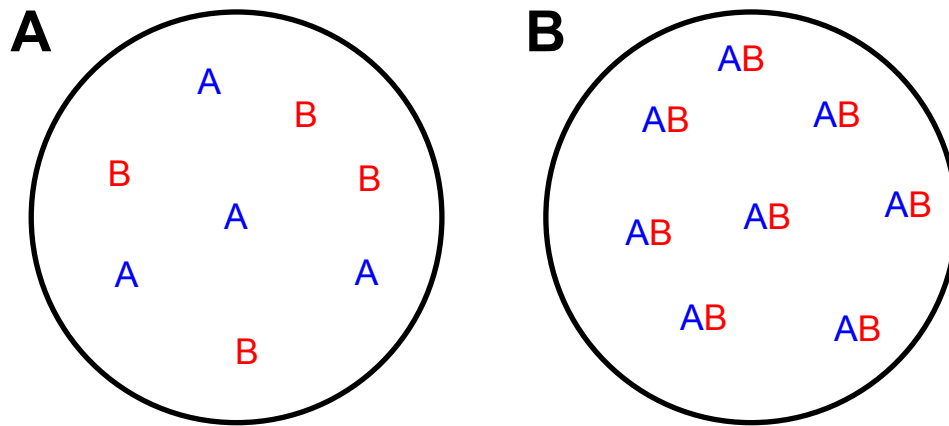


Figure 1: Example of interpersonal and intrapersonal cultural heterogeneity. Two hypothetical organizations are represented by a circle each. Individuals making up the organization are represented by letters, corresponding to the cultural beliefs they espouse. Organization A exhibits high interpersonal but low intrapersonal heterogeneity, whereas organization B exhibits low interpersonal but high intrapersonal heterogeneity.

May 8, 2008



"VMware is a fun place to be; it's aggressively growing, and it strives to provide a comfortable work environment"

★★★★★

Current Employee - Software Engineer in Palo Alto, CA

■ Recommends

■ Approves of CEO

Pros

VMware is a fun place to be during the day. The company is growing, and their technology has lots of room to grow and improve. The employees are diverse in age and culture; and everyone is **friendly**. This is helpful as many of us are new to the Bay Area. Everyone in the company is **smart**, works hard, and pulls their weight. There's a culture of **collaboration** and a degree of freedom in selecting one's work. VMware provides a great set of on-site **perks** free soda & other healthy **beverages**, on-site espresso machines, plenty of healthy **snacks**, **pinball** machines, a friday "beer bash", and plenty of **games** like **pool**, **ping-pong**, and **foosball** tables.

Show Less

Cons

Although virtualization is an interesting and exciting field; some of the periphery work can be burdensome. Because the stock is volatile; recent hires will become very wealthy before other employees who have stock options that are severely underwater. Some employees underestimate the tangible advantages of competitors' products; like Parallels and Microsoft HyperV. Occasionally, **perks** are changed without employee involvement; to the chargin of those who chose VMware for a specific perk. A weakness is that VMware's products built by Unix people are difficult for Windows/Mac-centric users; and products built by Windows people are difficult for Unix/Mac-centric users. The **cubicle** areas can be **noisy**.

Show Less

Topic Probability Distribution

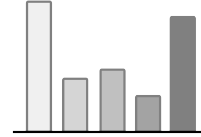


Figure 2: Stylized Example of LDA's Mixed Membership Topic Assignment. For a sample review from our data, we show how words associated with different topics, represented by different shades of grey, can be aggregated to yield a probability distribution over topics (5 in our stylized example).

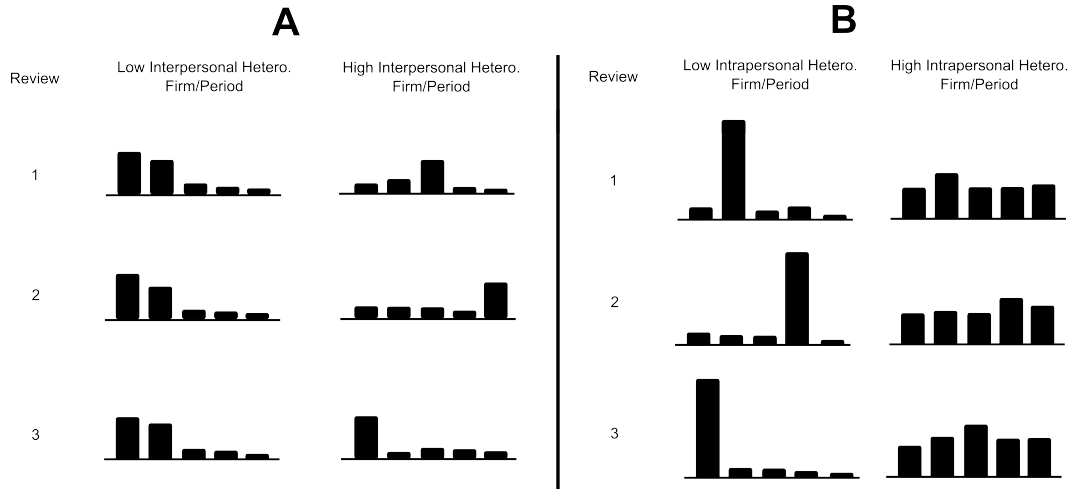


Figure 3: Stylized Illustration of Interpersonal and Intrapersonal Heterogeneity. Panel A illustrates, using the case of a firm that had three employee reviews, how the topic probability distributions might look for firms that are low vs. high in interpersonal heterogeneity. Panel B shows, again using the example of a firm with three employee reviews, how the topic probability distributions might look for firms that are low vs. high in intrapersonal heterogeneity.

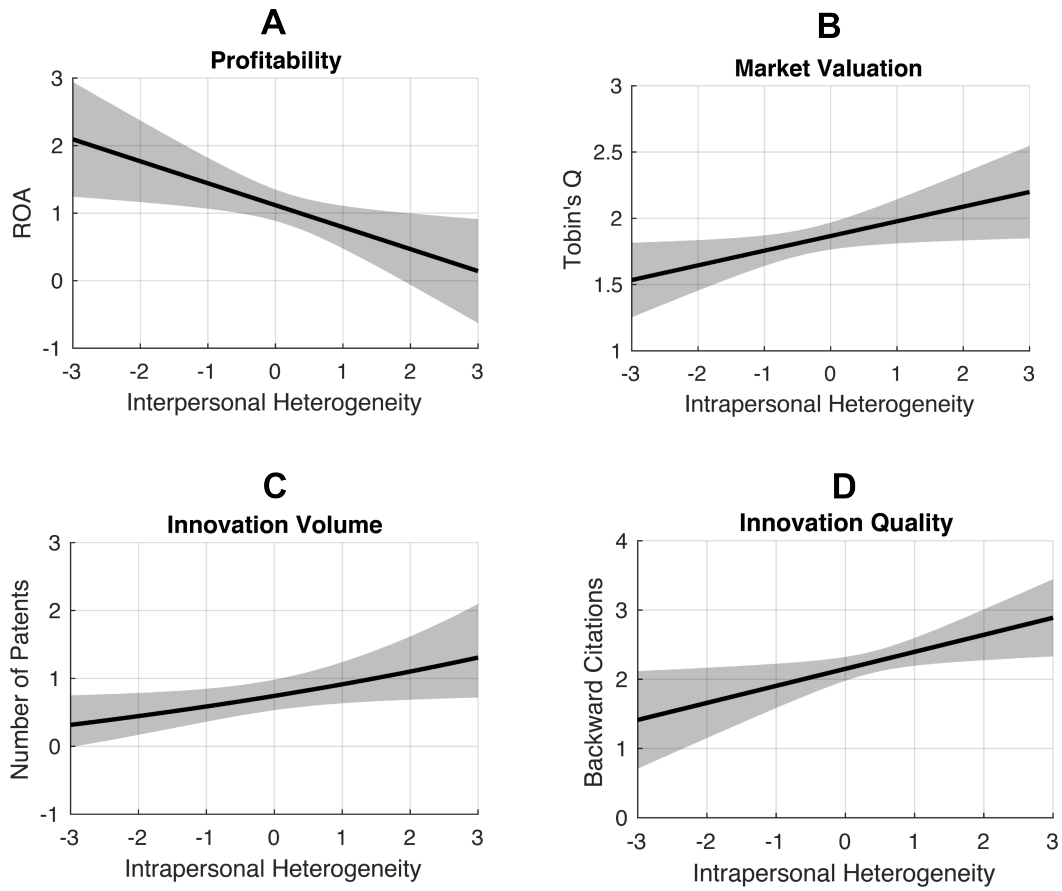


Figure 4: Predicted Effects from CEM Models of Interpersonal and Intrapersonal Heterogeneity on Firm Performance and Patent Outcomes. Panel A shows how a firm's predicted return on assets vary as a function of interpersonal heterogeneity. Panels B, C, and D show how a firm's predicted market valuation (Tobin's Q), number of patents filed, and backward citation count of patents vary as a function of intrapersonal heterogeneity.

TABLES

Table 1: Highest-Weighted Words for Selected LDA Culture Topics

Selected Topic #	Label	Words
10	production quality v. quantity	qualiti product high produc deliv commit counter deliveri emphasi quantiti
16	entrepreneurial norms	entrepreneuri thrive initi spirit dynam motiv collabor starter creativ individu
33	travel and multiculturalism	travel opportun world countri meet experi chanc interact abroad global
35	social interaction	social fun event activ interact lot frequent aspect regular includ
Random Topic #	Label	Words
473	performance recognition	perform reward recognit recogn incent individu consist mediocr contribut resul
415	organizational structure	compani exist virtual anymor bought parent basic built longer sold
399	fun work environment	fun cowork great cool amaz outgo hang toy lightheart easygo brilliant train hip
481	workplace challenges	challeng work interest reward present demand inher tackl workplac

Table 2: Univariate Statistics and Bivariate Correlations

	Mean	S.D.	# Obs.	(1)	(2)	(3)	(4)	(5)	(6)	(7)
1. ROA	1.45	2.51	3251	1						
2. TQ	1.78	0.78	3251	0.41***	1					
3. Patent Count	15.72	64.33	3251	0.14***	0.064***	1				
4. Lag Interpersonal Hetero.	-0.01	0.99	3251	-0.16***	-0.29***	-0.061***	1			
5. Lag Intrapersonal Hetero.	0.02	0.97	3251	0.031	0.098***	0.13***	-0.56***	1		
6. Lag Log Assets	10.21	1.96	3251	-0.12***	-0.46***	0.15***	0.40***	-0.042*	1	
7. Lag Log # Reviews	4.62	0.68	3251	0.076***	0.045*	0.0038	0.081***	-0.18***	0.21***	1

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Log of Mean # of Backward Citations not shown above since non-missing values are conditional on patenting:

Mean: 2.22, S.D.: 1.01, # Obs.: 774

Table 3: t-Tests on Covariate Means Before and After Matching

	Before matching <i>t-test on means</i>		Matched sample <i>t-test on means</i>
ROA	low-high interpersonal hetero.		low-high interpersonal hetero.
Lag Log Assets	-4.52		-0.75
Lag Log # Reviews	-6.49		-0.18
Strata		2515	
Matched strata		265	
Multivariate L1 distance	0.90		0.66
Low interpersonal hetero. obs.	1639		360
High interpersonal hetero. obs.	1644		375
TQ	low-high intrapersonal hetero.		low-high intrapersonal hetero.
Lag Log Assets	0.55		1.01
Lag Log # Reviews	10.40		0.54
Strata		2496	
Matched strata		255	
Multivariate L1 distance	0.91		0.72
Low intrapersonal hetero. obs.	1637		337
High intrapersonal hetero. obs.	1616		362
Log # Patents	low-high intrapersonal hetero.		low-high intrapersonal hetero.
Lag Log Assets	0.74		1.16
Lag Log # Reviews	10.27		0.49
Strata		2537	
Matched strata		259	
Multivariate L1 distance	0.92		0.72
Low intrapersonal hetero. obs.	1668		346
High intrapersonal hetero. obs.	1644		368
# Backward Citations per Patent	low-high intrapersonal heterogeneity		low-high intrapersonal heterogeneity
Lag Log Assets	0.75		1.24
Lag Log # Reviews	6.45		0.77
Strata		579	
Matched strata		111	
Multivariate L1 distance	0.94		0.85
Low intrapersonal hetero. obs.	426		166
High intrapersonal hetero. obs.	523		189

Table 4: ROA on Interpersonal Heterogeneity

	(1)	(2)
	OLS	Matched
Lag Interpersonal Hetero.	-0.38*** (3.71)	-0.33* (2.47)
Lag Log Assets	0.22* (2.08)	
Lag Log # Reviews	0.27 (1.83)	
Constant	-0.57 (0.56)	0.39 (1.24)
Quarter FEs	yes	no
Industry FEs	yes	no
Stratum FEs	n/a	yes
Firm/Quarters	3283	735

Absolute t statistics in parentheses

Standard errors clustered by firm

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 5: Tobin's Q/Patent Outcomes on Intrapersonal Heterogeneity

	(1) TQ <i>OLS</i>	(2) TQ <i>Matched</i>	(3) Log # Pats. <i>OLS</i>	(4) # Pats. <i>Neg. Bin.</i>	(5) Log # Pats. <i>Matched</i>	(6) # Back. Cites <i>OLS</i>	(7) # Back. Cites <i>Matched</i>
Lag Intrapersonal Hetero.	0.10*** (3.69)	0.11* (2.18)	0.14** (2.76)	0.28*** (3.43)	0.094* (2.13)	0.14** (3.05)	0.25* (2.39)
Lag Log Assets	-0.11** (3.15)		0.21** (2.66)	0.83*** (5.89)		0.14* (2.23)	
Lag Log # Reviews	0.16** (3.17)		0.054 (0.54)	0.071 (0.32)		-0.15 (1.87)	
Constant	2.80*** (8.73)	1.01*** (10.20)	-0.98 (1.22)	-22.5* (1.97)	-0.0039 (0.07)	2.10*** (4.32)	1.01 (1.60)
Quarter FEs	yes	no	yes	yes	no	yes	no
Industry FEs	yes	no	yes	yes	no	yes	no
Strata FEs	n/a	yes	n/a	n/a	yes	n/a	yes
Firm/Quarters	3253	699	3296	3296	714	949	355

Absolute t statistics in parentheses

Standard errors clustered by firm

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

APPENDIX: Measuring Cultural Heterogeneity Using Latent Dirichlet Allocation

All analyzed text was first preprocessed according to standard text analysis conventions. We removed common stop words and punctuation, discarded word order, and stemmed the words using the Porter stemming algorithm.

To train the Latent Dirichlet Allocation (LDA) model, we constructed a document-term matrix for which the rows represent distinct sentences observed across all available reviews for all organizations that contain the word “culture” or a close synonym (environment, atmosphere, attitude, climate, value, philosophy, belief). This results in 904,613 sentences. We identify the 4,000 most popular unigrams in these sentences. Less popular words outside of this set were increasingly proper noun references, badly misspelled, or nonsense words. After we manually removed proper nouns, the document-term matrix tracked the frequency of 3,870 words.

This set of training sentences was analyzed using LDA—a model of the probabilistic generation of a text corpus. Documents are represented as random mixtures of topics, and each topic is characterized as a probability distribution over words (Blei et al., 2003). We parameterized LDA to identify 500 topics present in these culture sentences. Each topic is characterized by a weighted set of words that tend to co-occur within documents.

After identifying cultural topics using this training set of sentences with explicit cultural references, we fit the LDA model to the reviews in our analytic sample. In contrast to clustering methods, LDA is a mixed membership approach, which assigns each document to a probability distribution over multiple topics.

Measure Variation

Organizational culture is stable but not invariant over time (Kotter and Heskett, 1992). As such, we examine the sources of variation in our measures of cultural heterogeneity. Figures A1 and A2 plot the within firm variation in interpersonal and intrapersonal heterogeneity, respectively, moving from time $t - 1$ to t . This visual evidence shows that cultural heterogeneity is relatively stable but not invariant over time.

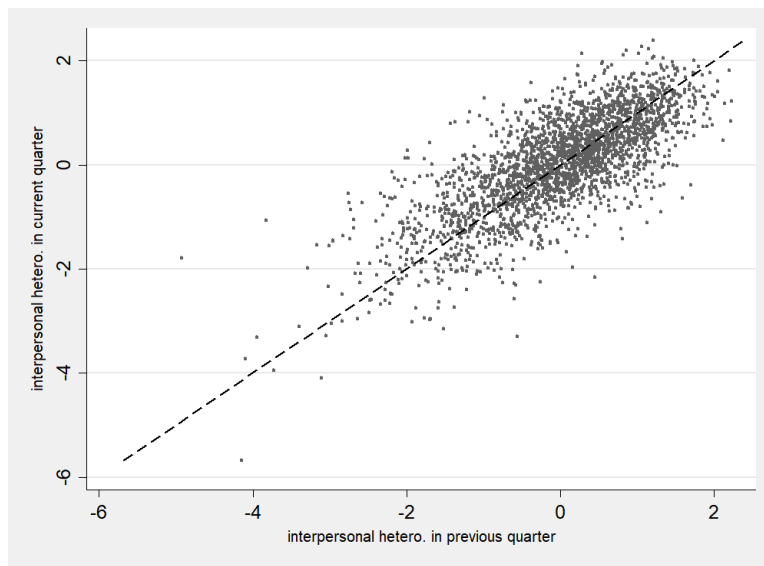


Figure A1: Within Firm Variation in Interpersonal Heterogeneity

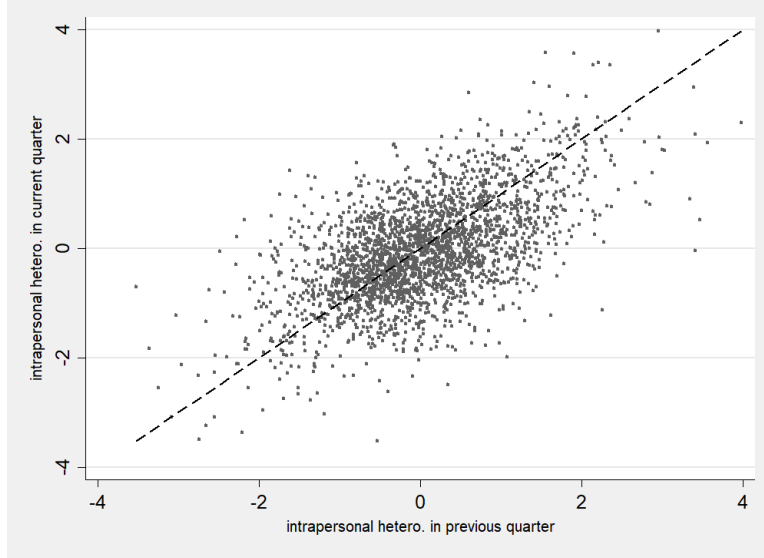


Figure A2: Within Firm Variation in Intrapersonal Heterogeneity

Additionally, we examine the within-firm temporal stability of the cultural heterogeneity measures across the full distributions of the measures. Figures A3 and A4 plot kernel density estimates of the distribution of each culture measure moving within-firm from time $t - 1$ to t . For both measures, Kolmogorov-Smirnov tests fail to reject the null hypothesis that the two distributions are different, providing statistical evidence that the culture measures exhibit relative stability over time.

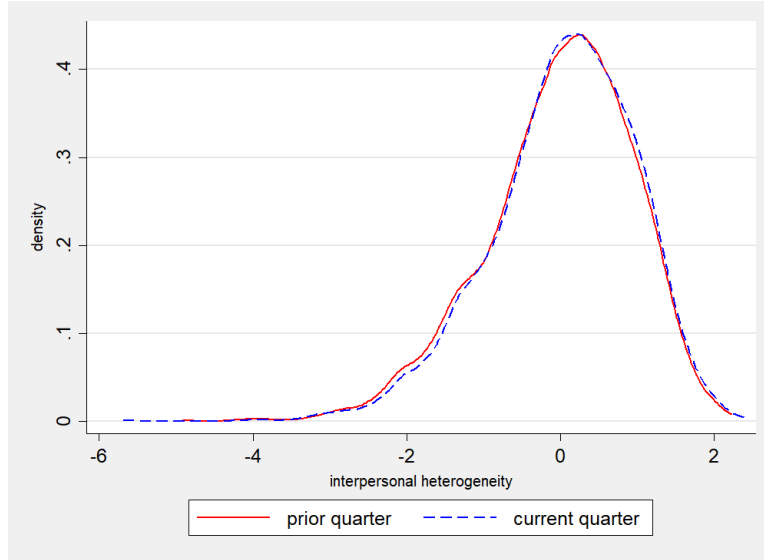


Figure A3: Time Variation in Interpersonal Heterogeneity

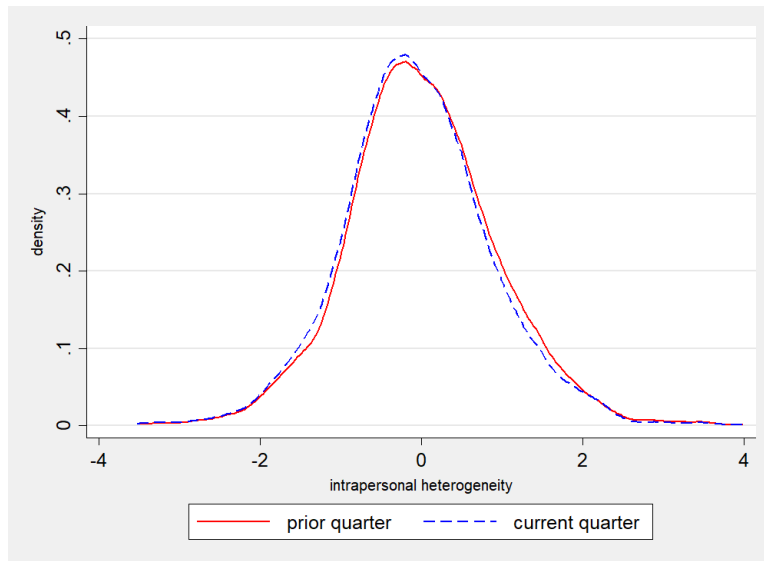


Figure A4: Time Variation in Intrapersonal Heterogeneity

Construct Validity

Beyond the face validity of the cultural topics that we demonstrated in Table 1, our heterogeneity measures themselves have construct validity as capturing variation along these culture dimensions. Table A1 shows the firms in the most represented industry in the data that score highest and lowest

on both cultural heterogeneity measures. Firms are split into large and small firms because the culture measures vary to some degree with firm size. Xerox has high interpersonal heterogeneity, or high disagreement among employees about how to characterize the culture. This accords with lay accounts of Xerox’s culture in the study period, during which a newly appointed CEO vowed to redefine the culture. Conversely, Facebook has low interpersonal heterogeneity, or high agreement about the culture. This is consistent with the company’s well-known emphasis on maintaining a startup culture focused on innovation, autonomy, and open collaboration. The firms high and low on intrapersonal heterogeneity similarly conform to intuition. For example, MicroStrategy has high intrapersonal heterogeneity, meaning its culture is organized about a broad, diverse set of cultural topics. Instead of keeping its engineers behind desks, the company is known to encourage them to work in the field in collaboration with clients so as to expose them to more challenges and potential solutions.

Table A1: Business Service Firms with Highest/Lowest Cultural Heterogeneity Scores, 2008-2015

	Large Firms	Small Firms
<i>Highest Interpersonal Heterogeneity</i>	Xerox SAP Paypal	Kelly Services Convergys TeleTech
<i>Lowest Interpersonal Heterogeneity</i>	Amdocs Facebook Wipro	National Instruments Sapient Cornerstone OnDemand
<i>Highest Intrapersonal Heterogeneity</i>	Microsoft Harris Corp Facebook	MicroStrategy National Instruments Intuit
<i>Lowest Intrapersonal Heterogeneity</i>	Wipro Infosys CGI Group	Virtusa Syntel IGATE

Notes: Restricted to firms with at least 3 quarterly observations. Large and small firms delimited by industry median size.

Additional face validity is demonstrated in the association between the cultural heterogeneity measures and Glassdoor respondents’ subjective assessments of the quality of firm culture and values. Generally speaking, we expect more interpersonally heterogeneous cultures to engender lower approval from members, and more intrapersonally heterogeneous cultures to elicit higher

approval. In supplementary Coarsened Exact Matching models, interpersonal heterogeneity is associated with a lower average culture and values rating and intrapersonal heterogeneity with a higher rating.

**ONLINE SUPPLEMENT to *Duality in Diversity: How Intrapersonal and
Interpersonal Cultural Heterogeneity Relate to Firm Performance***

**ONLINE SUPPLEMENT A: Exploring the Interrelationships Between Interpersonal
and Intrapersonal Cultural Heterogeneity**

In the main body of the paper, we examine the independent relationship between the two cultural heterogeneity types—interpersonal and intrapersonal—and organizational performance. We show that the former predicts profitability, and the latter is associated with growth potential and innovation. A natural extension would be to explore how the two types of heterogeneity relate to one another in producing these outcomes.

Two considerations give us pause in performing such an analysis, however. First, we expect this relationship to be complex and non-linear. Our theoretical framework leads to hypotheses about the independent main effects of interpersonal and intrapersonal heterogeneity on performance, but it does not straightforwardly imply what their joint effects would be. For example, greater intrapersonal heterogeneity implies a greater potential for interpersonal confusion: individuals who espouse multiple and potentially conflicting cultural frames might appear to their peers as behaviorally inconsistent across situations. At the same time, greater within-person cultural diversity mechanically increases, on average, the likelihood that two individuals would have cultural overlap. In other words, intrapersonal heterogeneity might both lead to greater and to diminished capacity for interpersonal coordination.

Second, although the two concepts are analytically distinct, their operationalization leads to a non-linear, mechanical relationship between them. For example, at the extreme positive end of intrapersonal heterogeneity—i.e., when all organizational members equally espouse all possible cultural elements—interpersonal heterogeneity will, by design, be 0. But when intrapersonal het-

erogeneity is low, for example, when all members adopt one cultural element each, interpersonal heterogeneity can range from 0 (when all members espouse the same element) to 1 (when each espouses a different element). This relationship becomes even more complex when there exists variability between members' levels of within-person heterogeneity.

To address the second, mechanical problem, we implement a procedure to adjust interpersonal heterogeneity relative to what would be expected at random, holding intrapersonal heterogeneity constant. Let N_{ot} be the number of individuals reviewing organization o at time t . We can represent the set of reviews for organization o at time t as an $N_{ot} \times K$ matrix X_{ot} (where K is the number of LDA topics). Each row represents the probability distribution of each individual review over the set of K topics. To induce the expected interpersonal heterogeneity, we randomly permute each row in X_{ot} to produce a permuted dataset labeled X_{ot}^i and calculate interpersonal heterogeneity $B(X_{ot}^i)$, as defined in equation 3 in the paper. This procedure preserves intrapersonal heterogeneity as it does not change the shape of the distribution of probabilities over topics within each review. Nevertheless, it affects interpersonal heterogeneity by randomly shifting the topics over which these probability distributions are distributed. We repeat the process $n = 1,000$ times to produce a reference distribution of interpersonal heterogeneity. The mean of this reference distribution is the expected interpersonal heterogeneity, $E(B_{ot}) = \frac{1}{n} \sum_{i=1}^n B(X_{ot}^i)$. We define the adjusted interpersonal heterogeneity as the distance between the observed and expected interpersonal heterogeneity, $\tilde{B}_{ot} = B_{ot} - E(B_{ot})$.

\tilde{B}_{ot} is the interpersonal heterogeneity that is not mechanically explained by intrapersonal heterogeneity. We find that for all our firm-quarter observations this variable is negative, suggesting, as one would assume, that all firms exhibit a level of interpersonal heterogeneity that is lower than what would be expected at random. In other words, all firms have greater cultural consensus than would be the case if an equally sized set of individuals with random cultural preferences would be assembled, keeping this set of individuals' level of intrapersonal heterogeneity constant. The analyses presented in this online supplement use the adjusted measure of interpersonal heterogeneity

whenever the two heterogeneity variables are included in the same model.

We begin by revisiting the CEM models that test our two hypotheses while including both heterogeneity variables in the models. Model 1 in Table OS-A1 tests Hypothesis 1 and reproduces Model 2 in Table 4. As expected, and consistent with Hypothesis 1, (adjusted) interpersonal heterogeneity remains negatively predictive of ROA, but intrapersonal heterogeneity does not. Models 2 to 4 in Table OS-A1 revisit Hypothesis 2, corresponding to Models 2, 5, and 7 in Table 5. Consistent with Hypothesis 2, intrapersonal heterogeneity is predictive of greater market valuation and innovation output, but interpersonal heterogeneity is not. The only exception is Model 3, which is marginally significant (at $p = 0.06$). In other model specifications (e.g., in an OLS model with quarter and industry fixed effects, corresponding to Model 3 in Table 5) we find that intrapersonal heterogeneity is significantly predictive of the number of patents (at $p < 0.01$) but that interpersonal heterogeneity is not.

Overall, the results in Table OS-A1 indicate that our findings are robust to the inclusion of both heterogeneity variables in our models and that the two different types of heterogeneity are differentially predictive of performance as we hypothesize. Non-hypothesized relationships between cultural heterogeneity and performance are not driving our results. In additional analyses not reported here, we find that, when modeled separately in a CEM specification, intrapersonal heterogeneity does not predict profitability and interpersonal heterogeneity does not predict market valuation or innovation.

Although theoretical predictions about an interaction effect between interpersonal and intrapersonal heterogeneity are not straightforward, we cautiously explore them in Table OS-A2. Our main hypotheses focus on the coordination disadvantages of interpersonal heterogeneity and the creative advantages of intrapersonal heterogeneity. Yet our outcomes are not perfect measures of coordination or creativity. For example, we use ROA as an indication of coordination efficiencies, but we expect that profitability is also affected, to some extent, by creativity. Similarly, patenting output depends on creativity, but the ability to translate creative ideas into useful innovation also depends

on team coordination capabilities. We therefore expect that organizations that simultaneously exhibit low levels of interpersonal heterogeneity and high levels of intrapersonal heterogeneity—that is, firms with consensual and broad cultures—will also exhibit high profitability and innovation relative to other firms. Conversely, we expect that organizations with high levels of interpersonal heterogeneity and low levels of intrapersonal heterogeneity—that is, firms in which individuals espouse different and narrow cultural beliefs—will exhibit lower performance on both dimensions.

To evaluate this proposition, we explore interactions between the two measures in Table OS-A2. Due to expected nonlinearities in effects, we dichotomize the interpersonal and intrapersonal heterogeneity variables into high and low binary variables at their respective medians and examine their interaction, which produces four different quadrants. This exercise leads to a significant reduction in statistical power. We therefore only explore OLS models without coarsened exact matching of observations and include only one quadrant dummy in each model. The results reported in Table OS-A2 are broadly consistent with our assumption about the interaction between the two types of cultural heterogeneity. Firms with low interpersonal and high intrapersonal heterogeneity exhibit greater profitability, higher Tobin’s Q, and higher innovation output than other firms. Culturally fragmented organizations, in contrast, whose employees adopt different and narrow cultural beliefs, exhibit lower profitability and market valuation. These firms also exhibit lower levels of innovation, but at marginally statistically significant levels. Additional analyses wherein we include multiple quadrant dummies in each model do not yield consistent and statistically significant results. We therefore interpret the results in Table OS-A2 as suggestive but inconclusive evidence for the non-linear, joint effects of interpersonal and intrapersonal heterogeneity on performance. We leave to future work further theoretical development and empirical exploration of the interrelationships between the two cultural heterogeneity measures.

Table OS-A1: CEM Models with Adjusted Interpersonal Heterogeneity Measure

	(1) ROA <i>Matched</i>	(2) TQ <i>Matched</i>	(3) Log Patent Count <i>Matched</i>	(4) # Back. Cites <i>Matched</i>
Lag Intrapersonal Hetero.	0.026 (0.21)	0.098* (2.10)	0.085 (1.88)	0.25* (2.46)
Lag Interpersonal Hetero. (adj.)	-41.8* (2.12)	12.5 (1.33)	8.86 (1.51)	8.82 (0.67)
Constant	-0.73 (1.52)	1.31*** (5.30)	0.21 (1.38)	1.21 (1.67)
Matching Weights	yes	yes	yes	yes
Stratum FEs	yes	yes	yes	yes
Firm/Quarters	794	699	714	355

Absolute t statistics in parentheses

Standard errors clustered by firm

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table OS-A2: Performance and Patent Outcomes on Heterogeneity Quadrants

	(1) ROA	(2) ROA	(3) TQ	(4) TQ	(5) Log # Pats.	(6) Log # Pats.	(7) # Back. Cites	(8) # Back. Cites
High Inter., Low Intra.	-0.50*** (3.69)		-0.21*** (5.18)		-0.14 (1.87)		-0.16 (1.92)	
Low Inter., High Intra.		0.26* (2.00)		0.15*** (3.68)		0.17* (2.30)		0.28*** (3.54)
Lag Log Assets	0.18 ⁺ (1.73)	0.15 (1.46)	-0.094** (2.79)	-0.10** (3.04)	0.21** (2.72)	0.21** (2.67)	0.15* (2.29)	0.14* (2.37)
Lag Log # Reviews	0.30* (2.08)	0.31* (2.03)	0.15** (3.12)	0.15** (3.04)	0.036 (0.37)	0.040 (0.40)	-0.19* (2.30)	-0.18* (2.27)
Constant	0.60 (0.65)	0.51 (0.53)	3.01*** (9.99)	2.90*** (9.46)	-0.56 (0.76)	-0.77 (0.99)	2.40*** (4.65)	2.13*** (4.32)
Quarter FEs	yes	yes	yes	yes	yes	yes	yes	yes
Industry FEs	yes	yes	yes	yes	yes	yes	yes	yes
Firm/Quarters	3283	3283	3251	3251	3296	3296	949	949

Absolute t statistics in parentheses

Standard errors clustered by firm

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

ONLINE SUPPLEMENT B: Glassdoor Data Details and Additional Robustness Checks

Employees reviewing their company on the Glassdoor website are required to enter both positive (“pro”) and negative (“con”) comments. Since our objective was to identify the general cultural dimensions mentioned by employees without regard to valence, we combined the pro and con text when analyzing the reviews. Examining the most highly-weighted words for each LDA culture topic reveals many topics that are largely neutral in valence.

Most visitors come to Glassdoor first and foremost to search for jobs rather than to post an employer review. Glassdoor employs a “give to get” model to solicit employer reviews from users. To receive unlimited access to the site’s content, users have to submit an anonymous employer review. Research using the Glassdoor data has found that this method mitigates ratings bias by reducing the prevalence of extremely positive and negative reviews.¹

Since the employees who write Glassdoor reviews were not selected through random sampling from the population of firm employees, a concern is that systematic variation in the number or composition of reviewers is driving the observed associations between the cultural heterogeneity measures and firm performance. We conducted two checks to examine the robustness of our results to potentially non-random selection of employees into writing Glassdoor reviews: 1) modeling within-firm variation in the number and composition of reviews as a function of firm size and performance; and 2) modeling within-firm variation in the cultural heterogeneity measures as a function of number and composition of reviews and firm performance and size. The sample includes firms with at least six quarterly observations so as to have enough within-firm observations to include firm fixed effects.

Table OS-B1 shows within-firm models of the number and composition of reviews used when calculating the cultural heterogeneity measures. These models test whether the number or composition of reviewers systematically changes during periods of high or low firm performance, which could

¹Marinescu, I., N. Klein, A. Chamberlain, and M. Stuart. 2018. Incentives can reduce bias in online reviews. NBER Working Paper No. 24372.

bias our calculations of cultural heterogeneity. We examined reviewer composition by measuring the percentage of reviews in a given firm/quarter written by employees in managerial positions as opposed to lower-level employees, as indicated by non-missing job title information. Models 1 and 2 show that net of firm size, the number of Glassdoor reviews does not vary as a function of either lagged Return on Assets (ROA) or Tobin's Q (TQ). Specifications 3 and 4 model the percentage of managers writing reviews as function of firm performance while controlling for number of reviews and firm size. Reviewer composition is insensitive to lagged ROA, but the percentage of managers decreases with increasing Tobin's Q. This result prompted us to include the percentage of managers as a control in our multivariate Tobin's Q models—inclusion of the control had virtually no impact on the size or significance of the intrapersonal heterogeneity coefficient. Thus, our findings appear to be robust to (observable) changes in reviewer composition.

Table OS-B1: Reviewer Characteristics on Performance

	(1) Log # Reviews	(2) Log # Reviews	(3) % Managers	(4) % Managers
Lag ROA	-0.0075 (1.54)		0.0000018 (0.00)	
Lag TQ		-0.091 (1.12)		-0.017* (2.38)
Log of Number of Reviews			-0.0029 (0.61)	-0.0034 (0.69)
Lag Log of Assets	0.19* (2.53)	0.15 (1.83)	-0.0031 (0.28)	-0.011 (0.97)
Constant	1.98** (2.73)	2.52** (2.86)	0.40*** (3.59)	0.51*** (4.37)
Year FEs	yes	yes	yes	yes
Quarter FEs	yes	yes	yes	yes
Firm FEs	yes	yes	yes	yes
Firm/Quarters	2776	2733	2776	2733

Absolute t statistics in parentheses

Standard errors clustered by firm

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table OS-B2 shows within-firm models of the cultural heterogeneity measures as a function of the number and composition of reviews, as well as firm performance and size. These models directly

test whether, net of firm size and lagged performance, the cultural heterogeneity measures vary with the number or composition of reviews. Models 1 and 2 show that both the interpersonal and intrapersonal heterogeneity measures have strong positive associations with the number of reviews, which reflects the sensitivity of the the measures to the number of reviews as inputs. These results prompted us to explicitly match on number of reviews in our coarsened exact matching models so as to ensure that variation in the number of reviews is not driving the results. In contrast, the percentage of managers is not significantly associated with the culture measures.

Table OS-B2: Cultural Heterogeneity on Reviewer Characteristics and Performance

	(1)	(2)
	Interpersonal Hetero.	Intrapersonal Hetero.
Log # Reviews	0.19*** (4.55)	-0.23*** (5.20)
% Managers	0.30 (1.13)	-0.12 (0.33)
Lag ROA	-0.0023 (0.48)	0.0029 (0.49)
Lag TQ	-0.00016 (0.00)	0.055 (0.65)
Lag Log Assets	-0.050 (0.58)	0.087 (0.91)
Constant	-2.01* (2.21)	1.42 (1.38)
Quarter FEs	yes	yes
Firm FEs	yes	yes
Firm/Quarters	2730	2730

Absolute t statistics in parentheses

Standard errors clustered by firm

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Next, we conducted additional robustness checks to examine whether potentially omitted variables are biasing our estimates of the effects of cultural heterogeneity on firm performance. First, we investigated whether our interpersonal heterogeneity measure is simply picking up on the existence of differentiated subcultures within organizations. It is important to note that interpersonal

heterogeneity need not be indicative of a division into subcultures. In fact, our measure of interpersonal heterogeneity reaches extreme values when people are idiosyncratically different from one another. When they are divided into subcultures, the measure is significantly lower (given that groups of individuals are culturally similar to one another).

Nevertheless, we implemented a supplemental analysis in which we used a clustering procedure to assess the extent to which each organization-quarter observation can be divided into cultural subgroups. The clustering procedure uses the K-means algorithm, based on our measure of cultural distance, to divide the population into clusters and then uses the gap statistic² to assess the optimal division into subgroups. In only about 5% of the cases did we find evidence of multiple subcultures—at least based on Glassdoor reviews. Model 1 in Table OS-B3 reports the results of a robustness check in which we include a control variable for the number of clusters, as measured by this gap statistic. Interpersonal heterogeneity is negatively related to return on assets even when we account for the existence of subcultures (i.e., the number of clusters).

We also tested whether the association between interpersonal heterogeneity and ROA is actually being driven by organizations with subcultures produced by greater structural segmentation. Model 2 and Model 3 show that our results are robust to controlling for structural segmentation, whether operationalized as the number of business segments comprising the firm or the number of 4-digit SIC codes spanned by the firm, respectively. Overall, we believe this provides strong evidence that the link between interpersonal heterogeneity and profitability is not driven by the existence of subcultures.

Further analyses we conducted sought to examine whether our results are driven by particular topics or are otherwise sensitive to culturally meaningless topics recovered by LDA. First, we examine whether our intrapersonal heterogeneity measure, which captures the extent to which reviews on average discuss a broad set of cultural topics, is sensitive to the possibility that some

²Tibshirani, R., Walther, G., & Hastie, T. (2001). Estimating the number of clusters in a data set via the gap statistic. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 63(2), 411-423.

Table OS-B3: ROA on Interpersonal Heterogeneity, Controlling for Subcultures (Number of Distinct Clusters), Number of Business Segments, and Number of SIC Codes

	(1) ROA	(2) ROA	(3) ROA
Lag Interpersonal Hetero.	-0.37*** (3.67)	-0.39*** (3.62)	-0.38*** (3.57)
Lag # Subcultures	0.19 (1.14)		
Lag # Business Segments		0.066 (1.40)	
Lag # SIC Codes			-0.029 (0.36)
Lag Log Assets	0.22* (2.13)	0.22* (2.01)	0.24* (2.23)
Lag Log # Reviews	0.24 (1.78)	0.26 (1.75)	0.30 (1.83)
Constant	-0.69 (0.67)	-0.59 (0.55)	-0.85 (0.80)
Quarter FEs	yes	yes	yes
Industry FEs	yes	yes	yes
Firm/Quarters	3283	2997	3001

Absolute t statistics in parentheses

Standard errors clustered by firm

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

topics may be closer in conceptual meaning than others. To do so, we assumed that two topics that are frequently mentioned within the same review are likely to have lower conceptual distance than two topics that rarely co-occur and therefore applied the cosine distance metric to the review-by-topic matrix to produce distance weights for each topic pair. We then produced an alternative measure of intrapersonal heterogeneity that, for each topic pair, multiplies the estimated prevalence of topic 1 and topic 2, weighted by the pairs conceptual distance. We then sum across all topic pairs to measure the dispersion for an individual review and, as with the original measure, take the mean dispersion across all reviews for a firm/period. This alternative distance-weighted measure is highly correlated (0.90 correlation coefficient) with the original measure, suggesting that there is little incremental value to accounting for the distance between topics and that our results are not skewed by the inclusion of similar cultural topics.

Second, we conducted a robustness check aimed at determining how sensitive our measures are to the inclusion of specific topics. To do so, we reproduced our measures by randomly selecting only a subset of k topics out of the 500 topics we recovered and used only these topics to calculate interpersonal and intrapersonal heterogeneity. We conduct analyses for values of k ranging from 5 to 100 topics. For each value of k , we repeat the process of randomly selecting topics 100 times, examining the correlation between the k -reduced measures and the original measures (based on the full set of 500 topics). The results reported in Figure OS-B1 plot the average correlation between the reproduced and original measures over these 100 random draws, as a function of the number of randomly selected topics. When $k=100$ (i.e., when we base our reproduced measures on only 20% of the original topics), the correlation with the original measures nears 1.

This analysis provides clear evidence that our measures are not determined by a handful of important topics or that culturally meaningless topics are biasing our results. In fact, even if we were to interpretatively determine that up to 80% of topics were culturally meaningless, we would still get effectively identical measures (our measures would retain 0.9 correlation even if we were to randomly select only 10% of topics).

These analyses provides strong evidence that even if some topics are culturally irrelevant or simply reflect linguistically aberrant patterns, these are not biasing our measures. They also helps validate that our choice to rely on a 500 topic model (as opposed to a smaller number of topics) is robust and preferable to models with a lower number of topics, which would likely render our results more sensitive to individual topics.

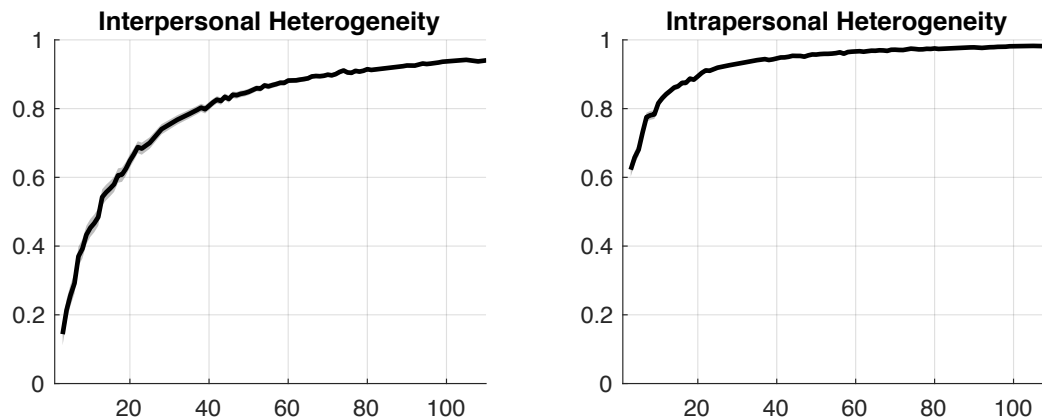


Figure OS-B1: Correlation Between k-Reduced Heterogeneity Measures and Original (Based on 500 Topics) Heterogeneity Measures

ONLINE SUPPLEMENT C: Comparing Selected Empirical Approach to One Based on Manual Coding of Cultural Categories

Our empirical approach relies on unsupervised learning of cultural content in employee reviews based on phrases that contain culture or one of its synonyms. We also choose an arbitrarily large number of topics (500) and consider the *distribution* of these topics within and between individuals in a firm. A virtue of this approach is that it does not rely on subjective human judgment about which topics are or are not culturally meaningful. Instead, it implicitly assumes that topics derived from phrases that contain the word culture or one of its synonyms are culturally meaningful even if a given researcher cannot see this meaning. In other words, it allows for cultural signal to be extracted from text even in cases when researchers lack access to the cultural toolkit needed to detect that signal. As a consequence, our approach is more scalable and easier to port across empirical settings than typical interpretative approaches, and is conducive for studying the distribution of a broad range of cultural beliefs when researchers are not focused on measuring the predefined linguistic signatures of specific cultural content. Moreover, our simulation analyses (see Online Supplement B) demonstrate that the resulting measures of interpersonal and intrapersonal heterogeneity are not sensitive to the specific topics used to derive these distributional measures and that measures based on considerably fewer topics are highly correlated with ones based on 500 topics.

Nevertheless, it remains unclear how our measures relate to ones derived from a more traditional approach that is based on human interpretation and coding of cultural categories. To evaluate the degree of correspondence, we undertook the following procedure. First, we began with a 100 topic model to make the task of manual coding more tractable. Second, we (the three co-authors) independently reviewed each topic and the top 25 word stems associated with it to determine: (1) which topics seemed to lack cultural coherence and should therefore be dropped; (2) which ones seemed to overlap and could therefore be combined; and (3) what label to assign to the remaining topics. Each coder’s independent judgments revealed substantial disagreement over both which

topics should be discarded and which topics should be combined. For illustration, out of the 38 topics that at least one author deemed as lacking coherence, only four topics were deemed so unanimously by all three coders. Similarly, after independently dropping and combining topics from the initial set of 100 topics, one coder proposed 33 topics, another coder 40 topics, and the third coder 57. As such, we sought to reconcile these independent attempts to manually code and interpret the topics using a systematic approach. In doing so, we began by applying the following decision rules: (1) discard topics that at least two coders identified as lacking coherence (20 topics were so identified); and (2) of the remaining 80 topics, combine topics when at least two coders had independently combined them. Applying these decision rules whittled the 80 topic set down to 59.

Next, we reviewed the resulting categorization. Of the 20 topics designated to be discarded, 5 were flagged by one coder as essential, and were reinstated. In 3 additional cases we decided to move three retained topics to the discard list. We then revisited the combined topic list, erring on the side of separating combined topics when, upon further reflection, we could identify subtle differences between them. In the end, we ended up with 65 topics. Table OS-C1 lists the 18 topics that were ultimately discarded, and Table OS-C2 lists the 65 topics (including combined topics) that were retained and the category labels we assigned to each.

Table OS-C1: 18 Discarded Topics

Topic #	Top 25 Words
5	long term time employe compani short work stay sink swim lot period posit employ year train busi longer plan stand tenur promot staf contract
7	good work healthi ethic friendli stabl condit load peopl peer pressur enviro infrastructur benefit enviorn freindli env challeng benifit pleasent or- din paymast remuner amic
9	good work learn salari facil fresher project technolog opportun onsit infrastructur exposur polici hike train hr technic apprais excel chanc transport join brand pressur

18	student research school colleg academ univers educ staff campu teach class graduat teacher fac- ulti administr institut excel facil support world lab kid studi state
23	time work part full good employe flexibl relax friendli lot job enjoy spend stress spent worker hectic wast amount timer flexi pretti money alot
27	good work nice peopl friendli ppl ambienc disen- gag mark adapt christian crap compon sm em- ployess dislik remind section won select smooth brand writer carpet
31	store custom fun discount retail manag employe product easi food cloth worker great sale servic sell restaur cowork deal job merchandis clean de- pend gener
34	compani make money daili corpor basi thing promis employe live run practic word save put heart preach complet show realiti manag absolut dollar true
44	depart depend team work group manag vari offic function divis locat hr cross project commun ex- peri branch area lot collabor gener silo greatli interact
47	call center time day job phone question make custom expect talk meet manag supervisor per- son email answer cold desk constantli sit break stress sale
48	great work peopl team fun experi fab envior sweatshop stagnant hardcor stellar enviorn di- rectli superior hill stodgi clean improp internet squeez american plain king
63	work good bad compani worst peopl profes- sion ethic experienc thing experi absolut appl kind encount pathet recommend load join hor- ribl imagin insid spoil posit
69	compani leav year employe peopl left month start hire ceo fire join stay replac employ due made end complet longer turn realiz desir feel
73	good work benefit salari pay peopl pretti comp- pens infrastructur facil satisfi benifit reput am- bienc cv freindli train prestig multipl mixtur bu- reacraci comapni sum ib
89	great peopl benefit work amaz compani awesom perk fantast product pay cowork locat ton fun terrif fabul phenomen flight stellar beat brand innov adob

90	peopl job talk thing bad make manag lot walk hand end kind put person head feel face fire smile stay pull dead leav boss
96	work compani make employe posit effort put feel creat reward contribut hard part impact step success extra made real individu huge recogn im- port show
98	compani con pro posit part corpor real start find startup found person type feel downsid kind tech biggest mention big import major fit love

Table OS-C2: 65 Selected Topics (Including Combined Topics)

Topic #	Label	Top 25 Words
38	aggressive	top compani line bank aggress corpor manag bottom start firm invest notch employe privat passiv financi industri front busi heavi public account gener client
11	poor leadership	lack manag leadership commun poor direct clear senior vision plan strategi account execut lead due depart process busi weak structur organ top chaotic reactiv
71	hostile management	manag employe hostil unprofession abus favorit hr behavior bulli bad horribl rude disrespect neg treat staff lie harass supervisor practic uneth toxic yell creat
84	blame culture	problem issu manag point blame review solv employe thing neg person fix address respons finger wrong mistak pass real account hr find understand deal
24, 76	work life balance	work life balanc good healthi excel flexibl person compens maintain brown perfect nose home worklif emphasi sheet compens superb strike conduc workahol memori emphas basic
51	fair compensation	good benefit pay decent great compens work fair packag salari competit excel locat perk train stabl offer cowork gener adequ fairli pto master descent
58	competitive compensation	salari benefit competit pay averag compens good industri compar low standard market lower packag offer work fair similar competitor par higher slightli area scale
78	benefits	benefit health great insur plan employe medic pay vacat paid offer match excel packag gener includ care time compens program stock bonu reimburs retir
29	bureacracy	make decis slow process thing risk difficult lot made move busi red polit corpor tape bureaucrat improv conserv bureaucraci progress impact frustrat lead avers
3	caring environment	employe care treat manag famili staff compani patient genuin hospit nurs health owner equal number resid gener employ clinic children show import understand treatment

20	challenge	difficult work time make challeng move thing hard find tough deal person peopl adjust face type bit quickli understand job adapt thrive statu busi
22	dynamism	work challeng project interest client dynam excit lot stimul divers intellectu technic technolog engag reward assign task team varieti involv intern great collabor offer
91	fast pace	fast pace challeng grow dynam move excit slow face learn quickli enjoy fun bore quick thrive demand extrem day constantli adapt handl toe intens
60	community	strong compani sens commun ethic employe teamwork integr corpor leadership focu excel safeti posit divers commit collabor promot pride profession core emphasi organ famili
28	performance	perform base polit promot highli competit reward review system driven recognit rank manag intern merit compens top individu compet peer talent result evalu process
79	cut-throat	cut technolog edg throat due cost constant layoff busi lead econom compani budget year unstabl competit creat recent frequent result continu reduc futur increas
15	safety	work control condit shop clean equip manag safeti air plant hot physic dirti system build area poor cold run qualiti floor manufactur mental offic
33	procedural	corpor polici govern rule union contract employe agenc due state standard oper employ time issu strict contractor procedur polit regul public organ militari ad
88	creative	creativ innov collabor encourag entrepreneuri support foster freedom challeng idea highli dynam motiv product independ talent individu inspir initi creat spirit teamwork design reward
75	consultative	busi firm client consult model market oper real practic partner unit run success understand side strong strategi focu develop area focus invest industri aspect
4	customer focus	custom servic product qualiti employe client focus focu compani base care deliv satisfact excel relationship experi improv result serv centric happi sale interact creat

95	brand focus	compani industri product market brand strong leader mobil innov great reput global world big upward excit media lead competit interest digit field financi stabl
35, 19	learning and training	experi train learn program skill support knowl- edg resourc gain manag technic profession op- portun develop lot excel improv intern set hu- man abil mentor share job industri
54, 56, 21	development and growth	career opportun profession advanc growth limit room person develop lot learn progress great grow path compani support skill potenti plenti excel start move challeng promot
77	growth oriented	compani grow continu improv growth busi start challeng maintain quickli pain rapidli constantli startup evolv success move thing excit expand organ small rapid process
72	gender diversity	divers boy women club group promot corpor school cliqu domin male bit femal network type mental age social men part conserv inclus polit workplac
66	engineering	technolog develop product engin softwar system design process tool project latest tech technic agil date test innov comput market outdat prac- tic learn data exposur
57	team excellence	great work excel team benefit support fan- tast colleagu worker train outstand leadership peer superb solid compens terrif enviro brilliant teamwork positiv ambienc rough etho
37	family oriented	famili orient friendli close team feel small type compani knit busi tight group part warm worker care friend commun cowork run felt owner home
14, 61	flexible	work flexibl home schedul hour friendli time great abil relax option fun remot pay casual easi independ benefit good cowork comfort decent freedom telecommut worker
50	demanding sched- ule	hour work long time flexibl stress expect pay week weekend day retail overtime requir shift sea- son lot job demand holiday busi night hard extra
55	serenity	work nice good friendli clean comfort pleasant worker safe facil profession peopl physic healthi campu modern quiet workplac peac calm neat cowork beauti condit
16, 74	friendly	work good friendli nice posit colleagu gener peopl profession boss enjoy cooper pleasant col- legu relax salari worker helpful support payment upbeat workmat workplac relationship comfort

52, 30	collegial	staff nice friendli good work peopl support offic profession gener colleagu relax member lo- cat warm worker excel pleasant collegi cowork cooper pretti experienc ambianc knowledg
2, 17, 25, 42, 45	fun	great make peopl work love benefit feel lot fun amaz team enjoy smart worker awesom fantast day meet friendli absolut thou friend nice cowork product
10	pet friendly	great work friendli worker fun peopl dog amaz eat love offic team famili cowork posit workplac super bring pet kind beat camaraderi anim bee
43	party	fun activ event compani lot social team em- ploye parti offic sport commun happi build in- volv meet celebr perk outing regular includ en- gag holiday monthli
53	youth/energy	fun young work peopl energet lot profession ex- cit dynam offic vibrant creativ upbeat energi age youth cool enthusiast colleg start motiv workforc make love
8, 39	global	opportun compani intern corpor divers local ex- peri offic world american travel busi exposur countri global japanes wide understand multi in- dia client bank project speak learn
68, 70, 97	laid back	back casual work laid relax friendli pretti dress easi offic code fun stab wear nice fairli busi cool gener comfort cowork peopl watch jean worker
92	exploration	learn lot great opportun thing experi work in- tern respons curv challeng freedom hand stuff exposur network alot resum chanc quickli scope teach steep travel
0	location	offic locat beauti area citi campu live great build park conveni downtown commut facil close site central easi san view access work small town
40	low pay	pay low salari rais wage benefit job increas de- cent promot stress minimum year bonus bonu rate start poor hourli paid averag base worker expect
41, 80	hierarchy	manag level upper senior micro employe middl staff support posit style entri understand execut poor higher approach lower direct director touch top care mid listen
83	merger	compani year past recent corpor improv start acquisit made lost ago shift employe merger wors coupl acquir continu move major due left declin complet

32	sharing	idea open share employe encourag manag improv collabor opinion inform feedback listen innov heard suggest knowledg voic freedom express creativ team implement discuss posit
65	open door	open manag polici door friendli employe commun mind transpar hr approach compani honest support collabor close access revolv senior feedback easili easi leadership inform
81	open space	offic work open space quiet cubicl bore dull bit build moment distract wall floor desk comfort cool small modern cube light depress loud area
64	work hard/hard/play	work hard play reward fun game enjoy mental harder find recogn dedic type favorit worker beat recognit role prepar push notic music acknowledged parti
86	perks	free food lunch fun perk coffe snack gym drink game offic friday beer room break lot park tabl kitchen site cafeteria event parti cater
87	politics	corpor structur organ polit process flat larg hierarchi highli organiz compani intern bit rigid extrem organis hierarch lack bureaucrat system conserv defin big tradit
36	profit	sale commiss sell competit product make custom base goal train pressur market potenti earn money driven hit manag rep number target push retail incent
46	stable	job work secur stabl easi posit stabil stress safe employ satisfact bore enjoy steadi duti task respons find descript repetit part titl fairli comfort
13	size	compani small big larg corpor feel size famili firm smaller startup larger town part start organ busi group bigger number agenc compar typic resourc
94, 82	stress	expect high work stress time pressur demand turnov meet low task school deadlin rate workload energi project level manag perform complet
85	supportive	employe posit compani creat support manag encourag care healthi promot relationship build engag happi maintain trust foster motiv workplac strive teamwork success genuin fellow
6	selection, turnover, and promotion	hire peopl compani fit posit job process promot recruit interview person talent find intern experi role candid type qualifi bring skill fire fill hr

67	individual excel- lence	peopl smart talent great work intellig motiv col- labor passion incred colleagu group surround amaz highli interest individu dedic driven bright creativ hardwork young fantast
49	teamwork	team member support orient collabor strong manag work build spirit posit leadership player leader encourag excel foster true promot execut cohes mate cooper focus
62	goal oriented	team goal success result achiev driven individu orient motiv focus posit set common organ drive succeed perform compani win person reward competit collabor creat
1, 59	hostility	poor work manag creat employe stress moral hostil low neg bad extrem lack toxic terribl make pay tens commun made staff uncomfot horribl unhappi leadership
93	fear	fear creat manag employe toxic ceo neg constant fire lead intimid blame hostil leadership micro- manag distrust base staff bulli trust run senior presid top
99	gossip	polit offic manag corpor extrem toxic bad hor- ribl terribl neg unprofession gossip drama work- plac boss backstab micromanag complet ego im- matur charg petti worst run
26	vacation	day time hour week work paid vacat shift leav holiday year month sick night earli break end start weekend home overtim schedul pm expect
12	vision	compani leadership team strong mission ceo vi- sion leader execut organ product creat posit amaz core great solid senior inspir driven pas- sion commit focus live

Undertaking this exercise reinforced for us the advantages of our chosen empirical approach over manual coding for studying the *distribution* of cultural beliefs between and within individuals. First, it became clear that coders brought different interpretive lenses to the exercise as a function of their idiosyncratic past experiences in organizations and cultural backgrounds. For example, two coders initially excluded Topic 0 on the basis that it appears to be about a firm’s physical location rather than about the firm’s culture. Yet one coder saw in it a culturally meaningful distinction between urban and suburban work cultures. Although there were 18 topics that we chose to drop, we suspect that a different research team might have seen some of these as culturally meaningful and instead chosen to include them. Second, some of the choices we made about which categories to combine or separate seemed arbitrary. For example, we ultimately chose to separate Topic 51 (fair compensation) from Topic 58 (competitive compensation), but we also think it would be completely reasonable to combine these topics. Similarly, although we combined Topics 2, 17, 25, 42, and 45 into a category labeled “fun,” we wondered whether a younger and hipper research team might have seen subtle distinctions among these topics that somehow eluded us.

Remaining mindful of these concerns about our manual coding procedure, we proceeded to create measures of interpersonal and intrapersonal heterogeneity based on the 65 topics identified in Table OS-C2. These measures were highly correlated with their corresponding measures from the 500 topic model (correlation coefficient of 0.76 for interpersonal heterogeneity and 0.83 for intrapersonal heterogeneity). Tables OS-C3 and OS-D4 reproduce our main empirical models in Table 3 and Table 4 using the 65-topic measures. The results are substantively identical to our main results, demonstrating that our empirical strategy recovers cultural signal even in the absence of human interpretation and manual coding.

This supplemental analysis supports our use of measures based on unsupervised learning of cultural topics for studying the *distribution* of cultural beliefs between and within individuals. However, we believe an approach involving human interpretation and coding of cultural categories likely has some advantages for studying particular cultural *content*, such as when researchers have

identified specific cultural themes and their linguistic signatures a priori. We encourage future work measuring culture using text to carefully consider the relative merits of each approach for different applications.

Table OS-C3: ROA on Interpersonal Heterogeneity

	(1)	(2)
	OLS	Matched
Lag Interpersonal Hetero. (65 topics)	-0.67*** (4.34)	-0.53* (2.05)
Lag Log Assets	0.29** (2.78)	
Lag Log # Reviews	0.18 (1.33)	
Constant	-0.48 (0.48)	0.50 (1.14)
Quarter FEs	yes	no
Industry FEs	yes	no
Stratum FEs	n/a	yes
Firm/Quarters	3283	735

Absolute t statistics in parentheses

Standard errors clustered by firm

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table OS-C4: Tobin's Q/Patent Outcomes on Intrapersonal Heterogeneity

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	TQ	TQ	Log # Pats.	# Pats.	Log # Pats.	# Back. Cites	# Back. Cites
	<i>OLS</i>	<i>Matched</i>	<i>OLS</i>	<i>Neg. Bin.</i>	<i>Matched</i>	<i>OLS</i>	<i>Matched</i>
Lag Intrapersonal Hetero. (65)	0.098** (2.77)	0.15* (2.44)	0.22** (3.07)	0.43*** (3.62)	0.13* (2.12)	0.20*** (3.43)	0.25* (2.00)
Lag Log Assets	-0.11** (3.03)		0.22** (2.78)	0.85*** (5.97)		0.14* (2.38)	
Lag Log # Reviews	0.16** (3.13)		0.067 (0.67)	0.084 (0.39)		-0.14 (1.79)	
Constant	2.83*** (8.60)	1.05*** (6.06)	-1.24 (1.48)	1.49*** (8.57)	0.023 (0.20)	1.92*** (4.04)	1.57 (1.14)
Quarter FEs	yes	no	yes	yes	no	yes	no
Industry FEs	yes	no	yes	yes	no	yes	no
Strata FEs	n/a	yes	n/a	n/a	yes	n/a	yes
Firm/Quarters	3253	679	3296	3296	703	949	349

Absolute t statistics in parentheses

Standard errors clustered by firm

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$