

Cultural Diversity and Team Performance: A Field Experiment*

Sander Hoogendoorn

Mirjam van Praag

Abstract

One of the most salient and relevant dimensions of team heterogeneity is culture. We measure the impact of a team's cultural diversity on their business performance using a field experiment. Companies are set up by teams of undergraduate students in international business in realistic though similar circumstances. We vary the cultural composition of otherwise randomly composed teams in a multi-cultural student population. Our results indicate that a moderate level of cultural diversity has no effect on the teams' business outcomes (sales, profits and profits per share). However, if at least the majority of team members is diverse then more diversity seems to affect the performance of teams positively. Our data suggest that this might be related to the more diverse pool of relevant knowledge facilitating (mutual) learning within diverse teams.

JEL-codes: J15, L25, C93, L26, M13, M14, D83

Keywords: cultural diversity, ethnic diversity, team performance, field experiment, entrepreneurship

*We are grateful to the department of international business studies of the Amsterdam College of Applied Sciences for their support in carrying out this research. We thank Diane Burton, Alexander Coad, Hessel Oosterbeek, Suncica Vujic and seminar participants in Amsterdam (ABEE), Barcelona (DRUID), Berlin (CASE Distinguished Lectures, Humboldt University), Bonn (EALE), Boston (AOM), Oslo (National Science Foundation), Potsdam (IZA Workshop), University of Aalborg, Utrecht (School of Economics) and The Hague (CPB) for their comments and suggestions. We acknowledge financial support from the Dutch Ministry of Economic Affairs. The usual disclaimer applies. Sander Hoogendoorn is affiliated with the CPB Netherlands Bureau for Economic Policy Analysis (2508 GM The Hague). Mirjam van Praag is the Maersk Mc-Kinney Møller Professor of Entrepreneurship at Copenhagen Business School (2000 Frederiksberg). Besides, both authors are affiliated with the University of Amsterdam School of Economics (1018 WB Amsterdam) and the Tinbergen Institute (3062 PA Rotterdam). Emails: s.m.hoogendoorn@cpb.nl; mvp.ino@cbs.dk (corresponding author). Phone: +31 6 5273 0756.

1 Introduction

It is impossible to pick up a business publication these days without reading about the wonders of teamwork. [...] Once teamwork is accepted as a basic business principle, it is not much of a stretch to think about teams that are comprised of diverse individuals, coming from different countries and cultures (Lazear, 1999, p. 15).

Human resources increasingly determine organizational performance in information-based economies. Teams have a central role in the human resources of modern organizations. One of the salient characteristics of a team is its diversity (Alesina and La Ferrara, 2005). Populations and workforces have become more diverse: the share of women in the workforce and cultural diversity are increasing in Western populations. Thus, the challenge and possibility have been created to bring together the necessary diversity of information, backgrounds, and values to create value (Jehn et al., 1999). In other words, team diversity can be employed as an organizational tool to promote efficiency or profitability (e.g., Cox et al., 1991; Horwitz and Horwitz, 2007; McLeod et al., 1996; Van Knippenberg and Schippers, 2007; Watson et al., 1993). Yet, the benefits and costs and precise impact of diversity in teams on their performance may be subtle (Van Knippenberg and Schippers, 2007; Williams and O'Reilly, 1998), dependent on the context (e.g., Joshi and Roh, 2009; Nielsen and Nielsen, 2013) and dimension of diversity (Ely and Thomas, 2001; Horwitz and Horwitz, 2007; Stahl et al., 2010).

We contribute to the literature on team diversity by measuring the effect of cultural diversity in business teams on their performance. To this end, we conducted a field experiment where randomization took place based on cultural diversity in teams of (student) entrepreneurs.¹ The choice for this topic is easily motivated.

Cultural diversity is highly relevant in an increasingly globalized world. Local Western populations - especially in big cities - become more mixed and multicultural (Alesina and La Ferrara, 2005; Lazear, 1999; Ottaviano and Peri, 2006; Ozgen et al., 2011; Stahl et al., 2010). In the United States, for example, the share of minorities is expected to rise from about one-third nowadays to roughly the majority in 2042 (Bernstein and Edwards, 2008). Moreover, teams within (multinational) firms are increasingly often staffed from a global base of potential team members (Lazear, 1999). As a consequence, it is likely that any

¹Another (orthogonal) dimension on which randomization took place in this 'double experiment' is gender. The gender experiment has been discussed in Hoogendoorn et al. (2013).

team will become more and more culturally diverse.

Cultural diversity implies differences between individuals in terms of their shared attitudes, values, goals, knowledge, beliefs and behavior (Hofstede, 1980, 1990). “Culture is the collective programming of the mind distinguishing the members of one group or category of people from others” (Hofstede, 1980, p. 25), Hofstede has demonstrated that cultural differences are substantial, especially across nationalities, even for countries that are geographically close to each other. Indeed, most academic studies have measured cultural diversity in terms of diversity in nationalities and/or ethnicities.² Cultural differences may be influential in the workplace (Hofstede, 1983, 1985; Guiso et al., 2009). Jehn et al. (1999) argue: “No theory suggests that a workgroup’s diversity on outward personal characteristics such as race and gender should have benefits except to the extent that diversity creates other diversity in the workgroup, such as diversity of information or perspective” (p. 742). Cultural diversity, serving as a proxy for diverse perspectives, beliefs and networks (Cox et al., 1991; McLeod et al., 1996), is also associated with diverse human resources (Richard, 2000).

In our study, we define cultural diversity in terms of nationality. We follow a common practice by measuring the extent of diversity in terms of the proportional representation of team members with a non-local (in this case, non-Dutch) origin.³ We base the definition of a non-Dutch origin on parent’s nationality: a team member is non-Dutch if at least one of their parents is born abroad. We use *parents’* country of birth because culture is a stable characteristic of nationalities (and ethnicities). Cultural changes require multiple generations since values and preferences are mainly acquired in childhood and via parents (Hofstede, 1980; Guiso et al., 2009).

Cultural diversity has been demonstrated to be meaningful in multiple dimensions of behavior and performance and at various levels of organizations ranging from nations to workteams.⁴ Research at the level of nations, regions, cities and to a bit lesser extent

²Examples of studies measuring cultural diversity in terms of nationalities are Bellini et al. (2013); Earley and Mosakowski (2000); Guiso et al. (2009); Hofstede (1980), Hofstede (1985); Kahane et al. (2013); Niebuhr (2010); Ottaviano and Peri (2006); Ozgen et al. (2011); Parrotta et al. (2010). Studies using diversity in ethnicity as a measure of cultural diversity include Cox et al. (1991); Ely and Thomas (2001); Hansen et al. (2006); Jehn et al. (1997); McLeod et al. (1996); Østergaard et al. (2011); Richard (2000), Richard et al. (2002), Richard et al. (2004) and Vogel et al. (2014). Nathan and Lee (2011) and Watson et al. (1993) are examples of studies using measures of cultural diversity based on both dimensions.

³Examples of studies using this measure of proportional representation are Cox et al. (1991); Earley and Mosakowski (2000); Hansen et al. (2006); Kahane et al. (2013); McLeod et al. (1996); Nathan and Lee (2011); Ottaviano and Peri (2006); Ozgen et al. (2011); Watson et al. (1993); Richard (2000), Richard et al. (2004); Stahl et al. (2010) and Vogel et al. (2014).

⁴Empirical studies have addressed diversity at the level of nations (Guiso et al., 2009), regions (Niebuhr, 2010), cities (Ottaviano and Peri, 2006; Nathan and Lee, 2011; Bellini et al., 2013), firms (Ely and Thomas,

the workforce in organizations seems to support the value of diversity in various ways. The outcomes from research at the team level, however, are far from univocal. It is often claimed that the lack of congruent results is due to contextual differences (e.g., team size or task complexity) and/or might relate to differences in the employed measures of diversity and performance (e.g., Stahl et al., 2010; Van Dijk et al., 2012; Van Knippenberg and Schippers, 2007; Williams and O'Reilly, 1998).⁵ The lack of consistent results so far is an additional motivation for our topic choice, besides the relevance and potential impact of cultural diversity in teams.

Our approach of a field experiment is motivated by the scarcity of studies identifying the causal effect of team diversity on performance in realistic circumstances. Identification of the causal impact of team diversity in real situations is hindered by the non-random nature of team composition (Boisjoly et al., 2006) and little spread in the distribution of team diversity.⁶ Laboratory studies, on the contrary, do not suffer from selection bias or a lack of spread in the diversity distribution of teams in the sample. Their disadvantage is a limited resemblance to real-world situations, typically measuring short-term effects (Boisjoly et al., 2006). This may be detrimental to finding effects of diversity, which are not likely to become evident instantaneously (Earley and Mosakowski, 2000; Harrison et al., 2002; Hoogendoorn et al., 2013; Watson et al., 1993).⁷

Few studies measuring the effect of cultural diversity have combined the advantages of studies in real organizations with experimental studies by carrying out field or 'quasi' experiments. Watson et al. (1993) show in a randomized experiment that culturally homogeneous student teams outperform heterogeneous teams on a problem solving task in

2001; Østergaard et al., 2011; Ozgen et al., 2011; Parrotta et al., 2010; Richard, 2000; Richard et al., 2002) and workgroups or (management) teams (Cox et al., 1991; Earley and Mosakowski, 2000; Hansen et al., 2006; Jehn et al., 1997; McLeod et al., 1996; Kahane et al., 2013; Richard et al., 2004; Vogel et al., 2014, Watson et al., 1993).

⁵Most studies so far have found insignificant or non-robust results (Carter et al., 2010; Hamilton et al., 2012; Hansen et al., 2006; Harrison et al., 2002; Jehn et al., 1997; Kilduff et al., 2000; Pelled et al., 1999; Watson et al., 1993). Positive conclusions about the role of cultural diversity on team performance can be drawn from Kahane et al. (2013), McLeod et al. (1996) and Vogel et al. (2014) and negative conclusions from Thomas (1999). Moreover, some studies find significant results that are non-monotonous (Earley and Mosakowski, 2000; Nielsen, 2010; Richard et al., 2004) or can be (partly) attributed to mediating variables such as various sorts of conflicts (Earley and Mosakowski, 2000; Jehn et al., 1997).

⁶The majority of studies on team diversity and performance is executed on real world data where team composition is non-random (e.g., Carter et al., 2010; Harrison et al., 2002; Jehn et al., 1997; Kahane et al., 2013; Kilduff et al., 2000; Nielsen, 2010; Pelled et al., 1999; Richard et al., 2004).

⁷A number of laboratory experiments have been performed and produced mixed results. For instance, positive results are found by McLeod et al. (1996) of ethnic diversity on brainstorming (compared to all Anglo-American teams). Earley and Mosakowski (2000) show that moderately diverse teams of managers assigned to a management team outperform other teams that participate in a management game. Thomas (1999) shows that culturally homogeneous teams of students in New Zealand outperform heterogeneous teams when evaluating business cases in a laboratory.

the short run, but not in the longer run, after 17 weeks. A recent study by Vogel et al. (2014) shows that (culturally) diverse teams of startup entrepreneurs get higher ratings from people acting as investors in an online experiment. Hansen et al. (2006) measure the impact of demographic diversity (age, gender and ethnicity) in student groups on the team's academic performance and find no effect.

Thus, randomized controlled trials measuring the causal impact of a team's cultural diversity on its performance are scarce. Studying a wide variety of culturally (more and less) diverse teams and their performance during a longer time in which people develop real interactions can be a fundamental contribution. Especially if the experiment takes place among business teams that have arguably similar tasks and objectives as real venture and management teams. Venture and management teams have much power in organizations and are, obviously, influential for value creation. Hence, obtaining insight in the effect of cultural diversity on performance in this kind of setting may be important. It is the objective of our field experiment in the context of a compulsory entrepreneurship program for undergraduate students in international business at the Amsterdam University of Applied Sciences.

The field experimental setting of our study allows us to draw team members from a diverse population of students from 53 nationalities (only 45% has a Dutch cultural background). The resulting sample of teams is extremely diverse in terms of the extent of cultural diversity, i.e., between 20% and 90%. This spread enables obtaining results along the entire distribution of diversity and thus creates the opportunity to study teams in which the majority of its members is diverse. In this sense, the results may be indicative of what may take place in reality in a number of years, given the sharp trend of increasing diversity. The wide spread may also reveal non-monotonicities that would otherwise remain undercover (Jehn et al., 1997).

The field experiment implies a level of controlled circumstances comparable to the laboratory: exogenously composed teams with different levels of cultural diversity start up companies simultaneously and under equal circumstances. All members come from a pool of students, minority or Dutch, of the same caliber. We follow 550 students who set up and manage 45 real companies as part of their curriculum in the academic year of 2008-2009. The teams' business performance is measured at the same moment (end of year), in a uniform way at a level playing field.

At the same time, the entrepreneurship program and the truly joint task with strong

incentives to maximize business performance in a real market resemble realistic circumstances. People have time to establish roles and observations of other team members. Since the entire firm consists of the team we study, our approach addresses concerns about the difficulties to isolate the impact of diversity on bottom line firm performance (Ely and Thomas, 2001).

Our results indicate a flat or even slightly negative impact of a team’s cultural diversity on its performance for teams of relatively low cultural diversity. This impact turns positive when the team starts to be truly heterogeneous, i.e., when the majority of the team is culturally diverse. The zero effect we find for that part of the distribution where cultural diversity is not so high is consistent with earlier findings, all based on samples including a limited range of the distribution of diversity (Carter et al., 2010; Hamilton et al., 2012; Kahane et al., 2013; Leonard et al., 2010; Parrotta et al., 2010). The positive effect of more diversity on business performance when the majority of the team has culturally diverse backgrounds is a novel finding, possibly due to the novel possibility of studying the very right handside of the diversity distribution. This suggests that there is a tipping point: “the moment of critical mass, the threshold, the boiling point” (Gladwell, 2000, p.12; see also Card et al., 2008).

Finally, in our far from perfect attempt to identify underlying or mediating mechanisms we show that the data are consistent with the idea that more culturally diverse teams tend to have a more diverse pool of relevant knowledge facilitating (mutual) learning. This may lead to better business outcomes. However, like most other studies, our tests of underlying mechanisms cannot provide causal evidence.

In what follows, Section 2 discusses theories that explain why and under which circumstances we might expect effects from a team’s cultural diversity on their performance. Section 3 provides information on the context and design of the field experiment (similar to Hoogendoorn et al., 2013). Section 4 describes the data and randomization checks. Section 5 shows the empirical findings. Section 6 provides a discussion and conclusion.

2 Cultural diversity and performance

2.1 *Cultural diversity and organizations*

Hofstede (1980, 1984) introduced the pervasive idea that the cultural backgrounds of people affect the organizations that people build and thereby their performance. Cul-

tural diversity is associated with both surface level diversity and deeper level diversity like attitudes and values (Constant and Zimmermann, 2008; Stahl et al., 2010). Surface level (demographic) differences are thus correlated with deeper level differences (Pelled et al., 1999) due to dissimilar belief structures (Wiersema and Bantel, 1992) that are often caused by differences in background, training and experience (Hambrick and Mason, 1984) as well as attitudes and cognitions (Kilduff et al., 2000). In other words, people sharing certain demographic characteristics such as nationality often tend to share certain world views, norms, values, and priorities (Ely and Thomas, 2001). On the contrary, differences in nationality are associated with distinct attitudes, values, perspectives and cognitive frameworks.⁸ People from different cultures bring different sources and means of information-processing to a team because of these differences (Hofstede, 2001; Lane et al., 2009). “The generic effects of diversity are thus likely to be magnified when the source of diversity is culture” (Stahl et al., 2010). Cultural diversity may affect teams differently than other diversity sources (Lane et al., 2004, 2009).⁹ Below we discuss theories that argue that cultural diversity in organizations and management teams might have positive and negative consequences. For both the benefits and costs of cultural diversity we revisit the literature to understand the arguments.

2.2 Benefits of cultural diversity

Various theories pose that cultural diversity generates benefits for organizations. Ely and Thomas (2001) argue that the “cultural styles and perspectives of minority people... are in fact valuable assets to workgroups” (p. 233) due to heterogeneous groups’ different points of view (Stahl et al., 2010). This leads to a larger diversity of approaches to tasks and problem solving (Ely and Thomas, 2001; Watson et al., 1993) and to potentially higher levels of creativity in multinational teams than intra-national teams (Stahl et al., 2010; Cox et al., 1991; Hambrick et al., 1996). Due to complementarities and learning opportunities (Alesina and La Ferrara, 2005; Lee and Nathan, 2011; Ozgen et al., 2011) and the need

⁸Values associated with cultural backgrounds from countries as close to each other as, for instance, Germany, France or Sweden can be different (Hofstede, 1984, p. 81).

⁹Instead of distinguishing surface level or bio-demographic diversity from deep level diversity, it has been distinguished from task relevant diversity, such as functional and educational background or information and perspectives that are relevant for accomplishing tasks (Horwitz and Horwitz, 2007; Ely and Thomas, 2001). Task related diversity has been distinguished from relation-oriented diversity (Joshi and Roh, 2009; Van Knippenberg et al., 2004). Cultural diversity involves most of these types of diversity. Type of diversity does not optimally map onto the contrasting theoretical perspectives that explain costs and benefits of diversity (Vogel et al., 2014). Van Knippenberg and Schippers (2007) claim that we should not accept that certain types of diversity might be beneficial whereas others are negative.

to integrate diverse perspectives (Van Knippenberg et al., 2004) culturally diverse teams are associated with more information sharing, more mutual learning and higher levels of creativity and innovation.

In the upper echelon theory, the mechanism at play is bounded rationality: managers selectively perceive a limited number of phenomena and can process even a subset of that for (strategic) decision making using filters ‘woven’ by their cognitive resources and values (Hambrick and Mason, 1984).

Vogel et al. (2014) measure that investors’ responses to (cultural) diversity of entrepreneurial teams are positive. Possible underlying mechanisms are the value of diversity for external relationships with stakeholders and the higher likelihood of pursuing strategies and activities of product differentiation and internationalization thanks to understanding the needs of a greater variety of potential customers.

Thus, theories acknowledge that diversity in attitudes, perspectives, values and (processed) knowledge, which are associated with cultural diversity, may impact organizational performance positively. The arguments are closely related to the information/decision making perspective (Pelled et al., 1999; Van Knippenberg and Schippers, 2007): a diverse team covers a broader territory of information, networks and perspectives, and this can lead to enhanced learning, problem-solving, creativity and innovation.¹⁰ They are also connected to the resource based view of the firm (Richard, 2000; Wernerfelt, 1984): more diverse teams have more diverse human capital.

2.3 Costs of cultural diversity

The literature has also indicated costs associated with cultural diversity in organizations, for instance of top management teams (Kilduff et al., 2000). Top management-team demographic diversity has been shown to be associated with higher turnover rates (Jackson et al., 1991; Wiersema and Bird, 1993), lower psychological attachment to the organization (Tsui et al., 1992) and more difficult communication (e.g., Zenger and Lawrence, 1989). While the benefits associated with cultural diversity are related to deep level diversity, some of the costs are related to surface level diversity (Bell, 2007; Harrison et al., 1998, 2002).

¹⁰The influence of diversity in attitudes and values has also been studied from a more general perspective and the results have been mixed (Van Knippenberg and Schippers, 2007). Van Knippenberg and Schippers (2007) conclude from the collection of research and the mixed results that ‘diversity in attitudes and values is worth research attention’ (p. 521).

Visible differences between team members may evoke (subconscious) classifications of each individual by individual team members as either ingroup/similar or outgroup/dissimilar (Van Knippenberg et al., 2004; Van Knippenberg and Schippers, 2007). Deep level diversity may also evoke categorization processes, although probably not instantaneously because it is less salient (Stahl et al., 2010). Categorizations may harm team processes due to favoritism and stereotyping and disturb communication and coordination, according to the social identity and social categorization theory (Tajfel, 1982). Stahl et al. (2010) note that nationality, race, and ethnicity have been among the most common social categories by which to identify people (e.g., Earley and Mosakowski, 2000; Tajfel, 1982).

Diversity may also be costly according to the similarity/attraction perspective (Van Knippenberg and Schippers, 2007). People are attracted to working with team members they find similar in terms of values, beliefs, and attitudes (Stahl et al., 2010; Williams and O'Reilly, 1998). Again surface level diversity may trigger immediate similarity-attraction processes (Stahl et al., 2010).

Another potentially costly aspect of diversity, especially of deep level diversity, is the possibility of conflicts due to lower levels of value congruity (Jehn et al., 1997). Various types of possible conflicts (relationship, process, task) may have different effects on performance (Jehn, 1995; Bendersky and Hays, 2012). Relationship-oriented conflicts are mainly associated with deep level aspects, such as values and norms (Stahl et al., 2010). They are detrimental for team outcomes due to lower satisfaction, intent to remain in the group, and effort.¹¹

Finally, faultlines have often been associated with the diversity-performance relationship (Lau and Murnighan, 1998). Their power depends on the similarity of group members within potential subgroups and the dissimilarity across subgroups. A hump shape in the diversity-performance relation may be caused by it because faultlines are more active at low and moderate levels of diversity. Faultlines are likely to impair team performance and may even lead to the dissolution of a team (Lau and Murnighan, 2005).

2.4 Link between theory and experiment

In what follows, we discuss the various empirical aspects of our study and then the empirical test results. In terms of the theory we can only make modest claims based on these

¹¹Moderate task related conflicts have the potential to affect team outcomes positively, especially with non-routine tasks (Jehn, 1995).

empirical tests. We can only test whether there is a net effect (benefits minus costs) of a team's cultural diversity on its business performance. Moreover, we may be able to see whether this effect is linear or is of another functional form. We will do so based on our particular definition and measurement of cultural diversity and team performance in the particular context of the experiment in terms of tasks, duration, team size and student population.

After we have measured this net effect, we perform some tests of mediating mechanisms that have been discussed in this section. These tests can only establish correlations between each underlying factor and diversity (and between that factor and performance). Hence, they are of limited validity because no causality can be inferred from them. Another limitation of this exploration is that the measures of most of the mediating factors are not precise and rely on self-reports (questionnaires). A final limitation is that we do not have any measures at all for some of the underlying mediating mechanisms proposed by theory.

We will test whether the information-processing explanation possibly generates measurable benefits by addressing the relationship between diversity and knowledge, complementary skills and (mutual) learning. Moreover, we will address the benefits of diversity due to more product differentiation. At the cost side, we are able to address the role of relationship-oriented conflicts and the role of categorization (or similarity/attraction) processes based on visible (surface level) differences between team members. We are not able to assess the role of networks, faultlines, diversity or non-relationship type of conflicts.

3 Context and design

3.1 Context

The teams in our field experiment take part in the Junior Achievement (JA) Young Enterprise Start Up Program, in the US known as the JA Company Program. This is the worldwide leading entrepreneurship program in secondary education (US and Europe) and post-secondary education (only Europe), see Oosterbeek et al. (2010). The specific entrepreneurship program that we study (in 2008-2009) is a year long compulsory and substantial part of the curriculum at the department of international business studies of the Amsterdam College of Applied Sciences, see Hoogendoorn et al. (2013) for a more extensive description. The basis is that 550 students set up and manage real companies

in 45 teams of about 10 to 12 students.¹² Students thus frequently interact, build up relationships, and create routines and processes to achieve their common goal.

The literature has pointed out that the diversity-performance relation of teams is affected by the context in which the team functions (Horwitz and Horwitz, 2007; Joshi and Roh, 2009; Nielsen, 2010; Stahl et al., 2010). The effect of (cultural) diversity depends for instance on the interdependence of the tasks involved and task complexity (Jehn et al., 1999). In particular, diverse information, skills, knowledge, ability and perspectives may be valuable in the case of non-routine and complex tasks (e.g., Stahl et al., 2010; Richard, 2000, 2002, 2004). Hence, cultural diversity may have a stronger positive impact on performance in the case of complex, interdependent and non-routine tasks.

Another important contextual factor that could affect the relation we study is the duration of the collaboration in the team (Watson et al., 1993; Harrison et al., 1998, 2002). Watson et al. (1993) claim (p. 592) that the beneficial effects of diversity are stronger in longer term collaborations. In new groups, members often pick up on surface level diversity, but not on deeper level diversity (Harrison et al., 2002; Zellmer-Bruhn et al., 2008). In the longer run, diverse group members integrate and get to know each other better and will appreciate the different perspectives, knowledge and view points they bring to the table (Harrison et al., 2002). They will learn over time how to exploit each others' networks and information sources better. The possible negative performance effects caused by surface level differences between team members based on categorization and/or similarity/attraction processes become weaker when team members get to know each other better (Watson et al., 1993). Deep level differences become more dominant after some time, in the longer run.

All in all, the specific context of our experiment might lead to more positive effects of cultural diversity in this sample than on average. The entrepreneurial task is arguably complex, non-routine and involves interdependencies to a large extent. Moreover, the teams in our study are rather large (10-12 members), which makes coordination more complex than in typical teams of 4-6 people (e.g., Stahl et al., 2010). The teams work together for a full year which can be considered very long compared to other studies (e.g., Watson et al., 1993). Finally, because the sample consists of students who self-select in a program of international business studies, with possibly an entrepreneurial

¹²Teams were allowed to choose the product or service themselves, after conducting market research. Hence, the college did not put restrictions on the type of business activity. Indeed, the products and services of teams varied greatly (see Table 1 for the various ventures).

Table 1. Team characteristics

#	Name	Share of minorities	Team size	Sales (euros)	Profits (euros)	Profits/share	Description of product/service
1	A-Card	0.50	16	1236	-848	-11.78	Discount card Amsterdam nightlife
2	A'dam Gadgets	0.45	12	534	-41	-0.47	USB hot plate for coffee, tea, etc.
3	Appie	0.40	11	455	150	3.00	Apple-shaped box to preserve apples
4	Aqua de Coctail	0.58	12	1130	-306	-3.12	Comprehensive cocktail shaker set
5	ArtEco Bags	0.60	11	912	-402	-7.44	Durable give-away bag clothes stores
6	BubbleMania	0.70	11	503	-62	-1.34	Multi-purpose protective key chain
7	D'Wine	0.63	9	740	-55	-1.62	Bottles of wine
8	Eastern Green	0.69	14	513	106	2.93	Engravable text bean growing a plant
9	Escapade Inc	0.22	9	593	-111	-3.09	Tube clip for sealing food, toiletry, etc.
10	eyeBMA	0.50	16	558	125	3.90	Package with easy-to-use eye shadow
11	Firefly	0.20	12	2226	294	3.67	Ascending fire lantern for celebrations
12	Fl!pthat	0.64	13	455	215	9.77	Redecorating already existing websites
13	Ginger	0.58	12	977	-107	-2.14	Multi-purpose solar energy charger
14	Himitsu	0.86	10	775	36	0.86	n/a
15	I-Care	0.54	15	1204	477	11.36	Dead Sea minerals beauty products
16	iJoy	0.64	14	1953	94	1.44	Wristband with USB capacity
17	I-Juice	0.54	13	1255	-39	-0.42	Pocket-size lightweight mobile charger
18	IMSC	0.55	11	625	-390	-7.41	n/a
19	iShield	0.50	11	4209	130	2.20	Invisible protective shield for iPhones
20	KISBag	0.57	9	205	-117	-3.90	Tiny foldable bag replacing plastic bags
21	Laservibes	0.40	11	130	-229	-4.32	Organizing lasershow for companies
22	Mengelmoes	0.71	10	942	63	1.24	Easy-to-wear telephone charger device
23	My-Buddy	0.45	12	297	-58	-2.65	USB doll for kids reflecting emoticons
24	Nine2Five	0.60	12	235	-1016	-12.87	USB hot plate for coffee, tea, etc.
25	Picture Perfect	0.54	15	260	-51	-1.45	Customized shirts for men and women
26	Pietje Plu	0.40	12	n/a	n/a	n/a	Trendy umbrellas
27	Pocket Memory	0.73	16	979	103	1.20	Business cards with USB capacity
28	Pro'Lux	0.54	14	378	-395	-9.18	Promotional gifts with USB capacity
29	Qwinlok	0.42	13	340	35	0.91	Boxer shorts for female adolescents
30	Reflection	0.36	11	890	45	0.84	Cosmetics mirror incl. mascara clip
31	SAME	0.36	11	1618	152	2.15	Comfortable unisex earwarmer
32	Sappho	0.50	8	980	n/a	n/a	n/a
33	Sharity	0.67	12	265	-241	-8.04	Peace sign necklace for teenagers
34	ShoeTattoo	0.77	13	270	88	1.21	Shoe customization by graphic artists
35	Student Promotion	0.42	13	571	235	15.64	Promotional activities for companies
36	StuPill	0.31	14	731	-1011	-15.48	Convenient Indonesian anti-RSI pillow
37	Test-a-Holic	0.45	11	728	220	4.88	Alcohol breath tester for nightlife
38	We-Do Solutions	0.56	10	604	-267	-6.06	Multi-purpose trendy key chain
39	We 'R U	0.33	13	1041	50	0.89	Compact wallet in several colors
40	XNG	0.90	12	1088	258	7.60	Shirts 'Chicks on Kicks' community
41	YEN Empowered	0.83	13	1267	33	0.71	n/a
42	YET's Wear	0.79	16	789	-247	-2.47	Customized shirts own YET-brand
43	YOU	0.64	12	0	-242	-6.55	Hotel door hanger to store keys, etc.
44	Young Legends	0.67	9	400	59	0.84	n/a
45	YUVA	0.70	16	1153	294	12.79	Engravable rice grain in glass covering

Note: Share of minorities is based on a dummy indicator for students of Dutch and non-Dutch cultural background (excluding students whose cultural background is unknown). The number of students whose cultural background is unknown amounts to 52 (out of 550 students). Team size reflects the size of teams at baseline. Missing or incomplete descriptions of a team's product or service are indicated by 'n/a' (not available).

orientation (Crant, 1996; Richard et al., 2004), the group’s ‘diversity perspective’, i.e., their normative beliefs and expectations about cultural diversity, is arguably more positive (Ely and Thomas, 2001). Table 1 lists team characteristics such as their cultural composition, size, business performance and product or service.

Cultural diversity

Students are defined as native or Dutch if both parents are born in the Netherlands, whereas students are considered immigrant, minority or non-Dutch if at least one parent is not born in the Netherlands (see Hartog, 2011). Our definition is thus based on the country of birth of the parents. Cultural values, attitudes and cognitions only change slowly over time from generation to generation (Hofstede, 1980), because these are sticky and acquired while young, mainly through the parents (Guiso et al., 2008). In turn, parents tend to teach to their children what they learned themselves from their own parents (Bisin and Verdier, 2000).

Panel A of Table 2 shows that the share of students with a non-Dutch cultural background is 0.55 and that the percentage of students with a non-Dutch cultural background varies per team from 20% to 90%.¹³ The 550 students in our sample are from 53 different countries of origin with an average number of approximately 6 different countries of origin per team. Students with a non-Dutch cultural background are born abroad in 38% of the cases, with 82% of the fathers and 87% of the mothers born outside the Netherlands. In 69% of the cases both the father and the mother are born abroad. For the purpose of presentation, the descriptive statistics in panel B are collapsed into groups (see Parrotta et al., 2010). They indicate that cultural variation among minority students is substantial.

We measure the extent of cultural diversity in terms of the proportional representation of team members with a non-local (in this case, non-Dutch) origin. A disadvantage of this common practice is that one may end up with a ‘diverse’ team with the majority of members still from the same country. We do not face that problem because non-Dutch team members within teams are almost invariably from different countries of origin, thanks to the great diversity in the sample. Hence, they represent true cultural diversity (Kahane et al., 2013).

Panel C indicates that there are virtually no clusters of cultural subgroups at the

¹³The average share of minorities in Dutch higher vocational schools is approximately 0.20, which is close to the population average (Hartog, 2011). The considerably higher share here reflects the international character of this program in the multicultural capital of Amsterdam.

Table 2. Descriptive statistics of cultural diversity

A: Cultural diversity (team level)				
	<u>Mean</u>	<u>SD</u>	<u>Min</u>	<u>Max</u>
Share of minorities	0.55	0.16	0.20	0.90
Number of different countries of origin	6.04	1.82	3.00	12.00
B: Cultural diversity (collapsed into groups of countries)				
	<u>Fraction</u>			
Netherlands (native)	0.45			
North America / Oceania	0.01			
Central and South America	0.15			
Formerly Communist Countries	0.02			
Mediterranean Countries	0.17			
East Asia	0.06			
Asia	0.04			
Africa	0.03			
Other European Countries	0.07			
C: Minority students with number of team members from the same country of origin				
	<u>Fraction</u>			
0	0.70			
1	0.19			
2	0.05			
≥ 3	0.06			
D: Numbers of students and teams, and share of minorities by field of study				
	<u>Students</u>	<u>Teams</u>	<u>Share of minorities (per team)</u>	
			<u>Average</u>	<u>Minimum</u> <u>Maximum</u>
Business management	240	18	0.52	0.20 0.79
Management	60	5	0.54	0.31 0.64
Trade management Asia	105	9	0.66	0.42 0.90
Business languages	118	11	0.50	0.22 0.77
Financial management	27	2	0.66	0.58 0.73
Total	550	45	0.55	0.20 0.90
<i>Note:</i> (The parents of) Minority students are from the following countries of origin: Afghanistan, Algeria, Angola, Antilles, Argentina, Australia, Barbados, Brazil, Cameroon, Canada, Cape Verde, China, Colombia, Costa Rica, Cuba, Dominican Republic, Egypt, Eritrea, Gabon, Germany and other Western European countries, Ghana, Guinea, Guyana, Honduras, Hong Kong, India, Indonesia, Iran, Iraq, Israel, Jamaica, Japan, Kenya, Laos, Malaysia, Morocco, Nigeria, Pakistan, Palestine, Peru, Philippines, Russia and other Eastern European countries, Singapore, Somalia, South Africa, Surinam, Tunisia, Turkey, United States and Vietnam.				

team level. The vast majority of non-Dutch students has no team members from the same country of origin (70%), whereas 19% of the non-Dutch students has only one team member from the same country of origin (the average number of students with a non-Dutch background per team is 6.8 at baseline). Hence, we can safely assume that the share of students with a non-Dutch background is a true indicator of the team’s cultural diversity. This varies exogenously and substantively. Moreover, the average and range of the share of minorities are similar across fields of study (panel D). We will test for possible pretreatment differences more formally in Section 4.

3.2 Design

One week before the start of the entrepreneurship program the school provided administrative data that we used for assigning students to teams. The cultural background of students was determined based on students’ names, Dutch versus non-Dutch (see Mateos, 2007 for a motivation of this name-based procedure). Nearly 90 percent of these name-based classifications matched with students’ actual national background that we retrieved through a pretreatment survey one week later. As expected, most of the mismatches in the category of minority students occurred among those with Dutch fathers.

Subsequently, we determined and varied the fraction of minority students for each team within fields of study and assigned students with Dutch and non-Dutch cultural backgrounds randomly to teams in accordance with these fractions. The program coordinators enforced this assignment successfully (only 6 students managed to switch teams during the year). Students and business coaches were uninformed, while professors only knew that a research project was conducted which required to stick to the imposed team assignment. Interviews with students corroborate their ignorance regarding our field experiment. All in all, we are confident that our results are not obtained due to conscious responsiveness to the experiment itself, possibly in a self-fulfilling way.

The design of the field experiment was such that non-Dutch students are randomly allocated across teams in varying proportions. Thus, the exogenous variation in teams’ diversity is based on the proportional measure of diversity. As a consequence, we can only use that measure convincingly to study the causal effect of cultural diversity on the performance of teams. Nevertheless, when we adapt our diversity measure by constructing a Blau’s diversity index (1977) that categorizes team members into different ethnic groups, i.e., Blacks, Whites, Latinos, Asians and Mediterraneans, we find similar results as with

our proportional measure of cultural diversity.¹⁴ In addition, if we distinguish between first and second generation minority students (and employ their proportional representations in the team), the results show that the impact of cultural diversity we find is not only driven by minority students of the first generation. This finding provides suggestive evidence consistent with the idea that culture is transmitted through parents (Hofstede, 1980; Guiso et al., 2009). As a final adaptation of our diversity measure, we have tried to identify whether the effect of cultural diversity on team performance differs when we restrict our definition of a non-Dutch origin from at least one parent to two parents born outside the Netherlands. Team members with only one parent born abroad would perhaps not self-identify as being different and this might lead to a difference in perceived diversity (Harrison et al., 2002). The results suggest that there might be such a difference, but we lack sufficient variation between teams to draw firm conclusions if we express cultural diversity in terms of the share of minority students with two parents born abroad.

Composition dynamics

Composition dynamics such as dropouts during the entrepreneurship program or the appointment of students to management teams may compromise the design of our field experiment if these dynamics alter the exogenously imposed cultural composition of teams.

During the year 104 students (19%) dropped out.¹⁵ This reduced the average team size from approximately 12 at the start to about 10 at the end of the program. The design of our field experiment was not contaminated by dropouts. First, the cultural composition of teams is largely unaffected. The correlation between teams' share of minorities at baseline and at the end of the program is 0.86. Second, dropout rates do not vary across teams in relation to their cultural composition. Regression coefficients of the share of non-Dutch students at baseline (and its square) when explaining dropout decisions are insignificant.

Companies appoint a management team twice, for the first and second half of the program. Tests indicate that the cultural diversity of the entire team - on which the treatment is based - is a realistic reflection of the management team (possibly the more influential part of the team). Regressing the share of minorities in the management team

¹⁴Blau's diversity index (1977) is defined by $1 - \sum_{i=1}^N p_i^2$, where p is the share of individuals in a (cultural) category i and N the number of categories.

¹⁵Dropout rates for first-year students in Dutch higher vocational schools are about 30% on average, including students that switch study and/or school (ref. HBO-raad, 2010). Lower dropout rates than the national average at the department of international business studies of the Amsterdam College of Applied Sciences may be explained by the fact that international programs generally attract students that are more motivated.

on the share of minorities in the entire team returns a coefficient of 0.95 for the first half of the program and 0.97 for the second half.¹⁶ Moreover, students of non-Dutch cultural background are not significantly more or less likely to be part of the management team than Dutch students in each of the semesters.

In sum, we are confident that the composition dynamics that take place after the initial assignment of students to teams did not compromise the design of our field experiment.

4 Data

Data about teams and individuals was collected through administrative data, teams' annual reports and three extensive surveys. Students filled out a survey at the first day of the academic year (pretreatment, in September 2008), halfway (in January 2009) and at the end of the entrepreneurship program (posttreatment, in May 2009). Response rates are respectively 88%, 86% and 78%. Based on the first survey that administers individual characteristics, we check the validity of the name-based assignment of students to cultural backgrounds and the random assignment of students to teams given their cultural background. The collection of questionnaires also enables us to relate our findings to and control for individual and team characteristics and (some of the) underlying mediating mechanisms discussed in Section 2.

Table 3. Descriptive statistics of individual and team characteristics

	Scale	Mean	SD	Min	Max
<u>Individual level</u>					
Age	years	19.37	1.99	15.98	30.92
Gender (dummy = 1 if female)	0/1	0.44	0.50	0.00	1.00
Grade point average	1-10	6.46	0.24	6.05	7.23
Business knowledge	1-5	2.66	0.88	1.00	5.00
Entrepreneurship knowledge	1-5	2.71	0.98	1.00	5.00
Leadership knowledge	1-5	3.14	0.98	1.00	5.00
<u>Team level</u>					
Size (at baseline)	persons	12.22	2.09	8.00	16.00

Note: Through standard batteries of questions we also obtain (non-tabulated) validated measures of 'softer' individual characteristics that are associated with entrepreneurship such as the 'big five' factor model of personality structure (see Goldberg, 1990; Zhao and Seibert, 2006). These are left out from the table for the purpose of presentation and not used in any of the analyses; there are no pretreatment differences.

¹⁶These coefficients are similar for median regressions that minimize the effect of possible outliers. F-tests further indicate that none of these coefficients are significantly different from 1.

Table 3 reports descriptive statistics of individual and team characteristics. The pretreatment survey administers background characteristics such as age and gender. In addition, we measure scholastic achievements of students just before entering the college (indicated by 'grade point average'). All three surveys also include self-assessments of the knowledge that students have in seven areas relevant for their studies. From these, we select the three most relevant knowledge areas for successful entrepreneurship (see Karlan and Valdivia, 2011; Minniti and Bygrave, 2001) and include them in our analyses of the relationship between diversity and knowledge, complementary skills and (mutual) learning.¹⁷ The levels reported in Table 3 are knowledge levels at baseline and also form the basis for our constructed measures of complementarities and learning. These and other mediating mechanisms such as product differentiation, relationship-oriented conflicts and categorization (or similarity/attraction) processes are more thoroughly discussed in Section 5.

Business performance metrics are gathered or calculated from the companies' annual reports that we obtained from 43 out of 45 teams (at the end of the program the teams were required to hand in their annual reports).¹⁸ We measure sales, profits and profits per share in euros. We also add a binary indicator of positive profits because students tend to view as the bottom line result whether or not they are able to satisfy their shareholders. Column (1) of Table 4 shows that average sales for all teams amount to 838 euros with a standard deviation of 707 euros. Profits are on average negative at -69 euros varying from a loss of 1016 euros to a profit of 477 euros. 22 teams make positive profits, while 21 teams run a loss. Profits per share vary between -15 and +15 euros.¹⁹

Columns (2) through (7) of Table 4 present descriptive statistics of business outcomes for different degrees of cultural diversity. We tabulate two pairs of cutoff levels around the average share of minorities in the sample of 0.55. Columns (2) through (4) with cutoffs at

¹⁷The knowledge areas that we include are related to business, entrepreneurship and leadership (thereby excluding administration, strategy, organization and management). Besides, the questionnaires administer self-assessments of entrepreneurial skills such as creativity, market awareness and networking (see Hoogendoorn et al., 2013; Oosterbeek et al., 2010; Parker, 2009). We left them out from further analyses. The entrepreneurial skills turned out not to be developed significantly during the entrepreneurship program and we observed no pretreatment differences forcing us to use them as controls.

¹⁸Each annual report (and its calculation of sales and profits) needed to be approved by an outside accountant. Profits were calculated as the difference between sales and costs, including tax and social security payments but excluding wage costs (students had the opportunity to determine their own wages).

¹⁹Shares have a nominal value of 20 euros. The mean number of shares issued is 52 (s.d. 21.5), while the minimum and maximum numbers of shares sold are respectively 15 and 100. Roughly half of the shares are owned by team members themselves (approximately 50 euros per student on average); the remaining shareholders are usually family members, friends or acquaintances. The number of shares and, hence, funding of companies is unrelated to the cultural diversity of teams.

Table 4. Descriptive statistics of business outcomes

	All	Cultural diversity (0.45 and 0.65)				Cultural diversity (0.40 and 0.60)			
		Low		Moderate		Low		Moderate	
		share<0.45 (2)	0.45≥share≤0.65 (3)	High share>0.65 (4)	High share>0.60 (7)	share<0.40 (5)	0.40≥share≤0.60 (6)	share>0.60 (7)	
Sales (euros)	Mean	838	859	881	745	1183	826	735	
	SD	707	634	891	349	622	895	468	
	Min	0	130	0	265	593	130	0	
	Max	4209	2226	4209	1267	2226	4209	1953	
Profits (euros)	Mean	-69	-39	-146	41	-97	-78	-51	
	SD	318	375	346	164	468	294	305	
	Min	-1016	-1011	-1016	-247	-1011	-848	-1016	
	Max	477	294	477	294	294	477	294	
Positive profits (0/1)	Mean	0.51	0.70	0.29	0.75	0.67	0.37	0.61	
	SD	0.51	0.48	0.46	0.45	0.52	0.50	0.50	
	Min	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
	Max	1.00	1.00	1.00	1.00	1.00	1.00	1.00	
Profits per share (euros)	Mean	-0.51	0.42	-2.07	1.46	-1.84	-0.58	0.01	
	SD	6.42	7.75	6.33	5.07	7.05	6.64	6.29	
	Min	-15.48	-15.48	-12.87	-8.04	-15.48	-11.78	-12.87	
	Max	15.64	15.64	11.36	12.79	3.67	15.64	12.79	
N	43	10	21	12	6	19	18		

45% and 65% of minorities indicate that teams of high cultural diversity tend to perform better than teams of moderate cultural diversity in terms of profits, the probability of profits being positive and profits per share. Performance differences are less pronounced in a comparison between teams of low and high cultural diversity, although the descriptive statistics suggest that on average business outcomes are slightly higher for teams of high cultural diversity. Columns (5) through (7) show a similar pattern for more symmetric cutoff levels at a share of minorities of 0.40 and 0.60, although the average levels of the business outcomes in the relatively small subgroups of low, moderate and high diversity are (insignificantly) different compared to the average levels in columns (2) through (4). We lack support among teams of low cultural diversity due to the limited number of observations in the range below 40% of minorities (only 6 teams). In what follows, our main focus is therefore on teams with a moderate or high share of minorities, i.e., teams of moderate or high cultural diversity.

Randomization

Before presenting the results in Section 5, we examine at the team level whether (average) characteristics of students correlate with the cultural composition of teams. Table 5 shows that cultural diversity is not systematically related to any of the (average) team characteristics except for the likelihood of studying Trade management Asia. This field of study attracts a disproportionate number of students with a non-Dutch (in this case, Asian) cultural background. Teams with higher shares of minorities are no different from teams with lower shares of minorities with respect to age, gender, GPA or initial levels of entrepreneurship related knowledge. Indeed, the average levels of these ability-related variables are no different for students with Dutch and non-Dutch cultural backgrounds (and the admission criteria for the program are the same for both groups). Hence, teams of different levels of cultural diversity are not composed of students of worse or better quality on average. The 'omnibus' p-value of the Chi squared test of the joint significance of all these relevant predetermined characteristics in explaining the share of minority students in teams is 0.62.

To further examine whether students are randomly assigned to teams of different cultural composition, we have also regressed individual students' characteristics on the share of minorities in their team, both jointly and separately for students of Dutch and non-Dutch cultural background. For these regressions at the individual level, the 'omnibus'

Table 5. Randomization checks

	Share of minorities	
<u>Personal characteristics</u>		
Age	-0.093	(0.473)
Gender	-0.193	(0.166)
Grade point average	-0.052	(0.058)
<u>Knowledge of</u>		
Business	0.380	(0.283)
Entrepreneurship	-0.039	(0.276)
Leadership	-0.308	(0.282)
<u>Field of study</u>		
Business management	-0.562	(0.432)
Management	-0.075	(0.272)
Trade management Asia	0.815**	(0.354)
Business languages	-0.328	(0.408)
Financial management	0.150	(0.148)
Team size	1.579	(1.820)

Note: Based on information from 43 teams. Each coefficient comes from a regression at the team level of the row variable on the share of minorities (bootstrapped standard errors in parentheses; 1000 replications). ***/**/* denotes significance at the 1%/5%/10%-level.

p-value of a Chi-squared test of the joint significance across all of the relevant predetermined student characteristics included in Table 5 are 0.25 for the total sample of students ($n = 434$, for which all regressor values are available), 0.27 for the subsample of Dutch students ($n = 182$) and 0.08 for the subsample of minority students ($n = 252$). In these joint regressions, the only coefficients that are significant belong to the dummy variable distinguishing students in the field of study Trade management Asia. Based on these analyses, all analyses in the next section include a team level dummy to control for the field of study Trade management Asia.²⁰

5 Results

5.1 Main findings

In this subsection we show the results from our explorative search for robust regularities in the relationship between cultural diversity and team performance. Table 6 shows the relationship between the share of minorities in a team and four measures of business performance: sales, profits, the probability of profits being positive and profits per

²⁰ As a robustness check, we have also obtained the main results while including each of the characteristics listed in Table 5 as a control (one-by-one to maintain sufficient degrees of freedom). Upon including these variables, such as the share of women or diversity in age, the estimates for the impact of cultural diversity remain similar (as expected due to the random assignment of students to teams).

share. Note that a larger share of minorities implies a more culturally diverse team due to the limited presence of cultural clusters within teams. In panel A the performance measures are regressed on the share of minorities in the team (using OLS, median and robust M-estimation regression). These specifications testing for a linear effect of the share of minorities on business performance turn out insignificant. We also test polynomial specifications in panel B and these turn out being (largely) insignificant too, although the point estimates consistently suggest a U-shaped relationship between the share of minorities and business outcomes with the minimum at a share of about 0.55.

Panels C1 through C3 of Table 6 measure the effect of cultural diversity on business performance using more flexible spline functions (that were also employed in Hoogendoorn et al., 2013). Using these, we further explore the possibility described in Section 1 that the relationship between team performance and diversity is nonlinear. Spline functions allow distinct slopes for various parts of the distribution of cultural diversity, unlike the linear specifications in Panel A. Moreover, they do not impose equal curvatures of opposite signs below and above the fraction of minority members at which business outcomes is lowest (given a U-shaped pattern), as is the case with a quadratic specification (see Panel B). In search for consistent patterns in the relationship between cultural diversity and team performance we explore different cutoff levels around the sample average of 55% of minorities.

Panel C1 reports the slopes of the estimated linear relationship between cultural diversity in the team and business performance below and above 50% of minorities in the first and second row, respectively. The coefficients are not significantly different from zero. Panel C2 shows the results for the cutoff at a share of minorities of 0.55. All point estimates for the share of minorities, given that this share is above 0.55, are positive (and except for sales) statistically significant. When the share of minorities is lower than or equal to 0.55, all coefficients are negative and in two cases significantly so. The point estimates in columns (4) through (6) of panel C2 imply that profits increase by approximately 100 euros (about one-third of a standard deviation) if the share of minorities is raised from 0.6 to 0.7. Columns (8) through (10) indicate for profits per share that an increase in the share of minorities from 0.6 tot 0.7 again has a positive and substantial impact of roughly one-third of a standard deviation. Panel C3 reveals a similar pattern for the relationship between business performance and cultural diversity above and below 60% of minorities: profits, the probability of profits being positive and profits per share only increase with

Table 6. Cultural diversity and team performance

	Sales			Profits			Pos. profits			Profits per share		
	OLS		Robust	OLS		Median	Robust	OLS		Median	Robust	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)		
A: Linear												
% minorities	-597.8 (624.5)	-267.3 (722.2)	15.3 (664.3)	106.7 (318.0)	3.2 (330.0)	-71.4 (237.8)	0.153 (0.528)	3.779 (6.200)	-0.122 (5.995)	1.751 (5.556)		
R ²	0.02	0.03	0.03	0.02	0.03	0.00	0.03	0.01	0.01	0.00		
B: Polynomial												
% minorities	-3798.5 (3968.5)	-3780.9 (4734.9)	-5009.1 (4322.8)	-1741.8 (2071.2)	-3018.7* (1621.2)	-2215.7 (1567.4)	-4.622 (3.170)	-29.232 (34.954)	-40.341 (33.848)	-48.841 (35.781)		
% minorities ²	2942.8 (3528.0)	3561.9 (4040.7)	4306.0 (3646.3)	1699.5 (1671.1)	2664.7* (1381.7)	1942.6 (1316.9)	4.391 (2.693)	30.350 (28.902)	36.524 (29.662)	45.266 (29.096)		
Minimum	0.65	0.53	0.58	0.51	0.57	0.57	0.53	0.48	0.55	0.54		
R ²	0.04	0.05	0.11	0.05	0.06	0.08	0.11	0.04	0.05	0.07		
C1: Spline (0.50)												
1st segment	-977.5 (2308.9)	-903.9 (2238.4)	-1709.2 (1938.0)	-598.8 (1046.5)	-955.1 (958.6)	-964.8 (947.3)	-1.816 (1.469)	-7.962 (17.040)	-18.541 (18.435)	-19.408 (19.467)		
2nd segment	-359.8 (1347.9)	786.2 (1072.7)	722.2 (834.8)	548.8 (398.8)	405.2 (445.5)	442.1 (401.3)	1.387* (0.811)	11.137 (8.285)	11.402 (10.001)	13.080 (8.892)		
R ²	0.02	0.05	0.08	0.05	0.05	0.08	0.11	0.03	0.05	0.07		
C2: Spline (0.55)												
1st segment	-1205.2 (1404.8)	-903.9 (1521.8)	-1102.3 (1451.2)	-568.9 (766.7)	-955.1 (647.4)	-946.6* (559.9)	-1.801* (1.088)	-9.349 (13.563)	-17.288 (10.851)	-17.494 (12.553)		
2nd segment	35.7 (1199.3)	786.2 (1191.1)	790.7 (921.3)	811.3* (459.4)	1007.6** (486.3)	757.9* (401.8)	2.191** (0.872)	17.471* (9.552)	16.378* (9.696)	19.601** (8.261)		
R ²	0.03	0.05	0.07	0.06	0.08	0.13	0.17	0.05	0.07	0.10		
C3: Spline (0.60)												
1st segment	-1200.1 (1012.2)	-903.9 (1329.7)	-814.2 (1191.7)	-514.4 (630.3)	-718.5 (497.0)	-786.4* (424.9)	-1.422* (0.854)	-8.327 (10.805)	-14.429 (9.575)	-14.789 (10.222)		
2nd segment	464.8 (1236.0)	1733.7 (1609.5)	1079.4 (1155.7)	1202.3* (632.2)	1257.0** (607.1)	1042.3** (517.3)	2.933** (1.143)	25.136** (11.410)	22.331* (12.322)	26.773** (10.950)		
R ²	0.03	0.05	0.06	0.08	0.08	0.14	0.18	0.07	0.08	0.12		

Note: Based on information from 43 teams. All specifications include a team level dummy for the field of study Trade management Asia. OLS, Median and Robust refer to the estimation method. Median and robust specifications for positive profits are excluded since this variable is dichotomous. Bootstrapped standard errors in parentheses (1000 replications). ***/**/* denotes significance at the 1%/5%/10%-level.

the share of minorities in the segment above 60% of minorities.²¹ The magnitude of the significant coefficients in panel C3 is even larger than in panel C2 and corresponds with an increase in these business outcomes of approximately two-fifth of a standard deviation for every 10 percentage point increase in the share of minorities (in the segment above 60% of minorities).

Panels C1 through C3 of Table 6 show that these results are not driven by outliers; they are largely insensitive to using OLS, median or robust M-estimation regression. Moreover, (non-tabulated) results from estimating these spline functions with different cutoff levels such as at a share of minorities of 0.45 or 0.65 give similar results as the ones tabulated for cutoff levels at 50%, 55% or 60% of minorities.²² Table A1 in the appendix further indicates that the results also remain similar when estimating spline functions with three segments (i.e., with teams of low, moderate and high cultural diversity). However, the (more limited) number of teams in each segment limits the precision of these estimates. The business performance of teams in the segments of low and moderate cultural diversity is on average less strongly affected by increasing the percentage of minority team members relative to the segment of high cultural diversity.

The results presented in this subsection indicate that the relationship between team performance and cultural diversity is flat or tends to decline below a certain threshold. However, a substantial and significantly positive marginal effect of cultural diversity on business performance is observed beyond this threshold. The threshold level is around a share of minorities of 0.55. Hence, it seems that only if cultural diversity is sufficiently substantial the net effect of cultural diversity on team performance is positive. This finding provides suggestive evidence consistent with the idea of a tipping point.

5.2 *Costs and benefits of cultural diversity*

This subsection discusses the results of our exploratory analysis of mediating mechanisms that potentially drive the positive effect of more diversity on team performance when the majority of the team has culturally diverse backgrounds. As discussed in Section 2, the analysis is limited by the fact that the relationship between each underlying factor and

²¹F-tests acknowledge that the coefficients of the first and second segment in panels C2 and C3 are not identical for profits, the probability of profits being positive and profits per share.

²²As another robustness check, we have also obtained the main results (see Table 6) while including each of the characteristics listed in Table 5 as a control (one-by-one to maintain sufficient degrees of freedom). Upon including these variables, such as the share of women or diversity in age, the estimates for the impact of cultural diversity remain similar (as expected due to the random assignment of students to teams).

cultural diversity (and between that factor and team performance) is open to multiple interpretations and not indicative of a causal effect. Another limitation is that we lack measures for some of the mediating mechanisms proposed by theory. Nevertheless, we think it is relevant to explore the nature and occurrence of this kind of relationships that may possibly explain our findings. In what follows, we first consider mediating mechanisms that are associated with the benefits of cultural diversity and subsequently the underlying factors that are related to its costs.

For the benefits, we test the relationship between cultural diversity and knowledge, complementary skills and (mutual) learning as suggested by the information-processing perspective (Pelled et al., 1999; Van Knippenberg and Schippers, 2007). We expect a positive relationship between diversity and business performance due to more learning and complementarities in culturally diverse teams. In addition, we address the benefits of cultural diversity as a result of more product differentiation in diverse teams (Vogel et al., 2014).

We measure learning in terms of the development in three knowledge areas most relevant for successful entrepreneurship: business, entrepreneurship and leadership (see Table 3). For each of these knowledge areas, the indicator of individual learning is the difference between the self-assessed level in the posttreatment and pretreatment questionnaire.²³ Table 7 reports results from least squares regressions (panels A and B) and spline functions (panels C1 through C3) of team-average learning in business, entrepreneurship and leadership knowledge on cultural diversity. Again, note that a larger share of minorities corresponds with a more culturally diverse team. Similarly to Table 6, we present results from using OLS, median and robust M-estimation regression. Panel A shows no significant linear effect of the share of minorities on learning, whereas panel B suggests a U-shaped relationship with the minimum at a share of approximately 0.55. Spline functions in panels C1 through C3 of Table 7 with a cutoff at respectively 50%, 55% or 60% of minorities indicate that the relationship between learning and share of minorities is flat or declines down to a threshold level of about 0.55 and starts increasing beyond this threshold level.

Hence, on average, members of culturally diverse teams learn more than members of homogeneous and moderately heterogeneous teams. Additional (non-tabulated) regres-

²³It may be somewhat uncommon to measure learning based on self-assessment since students' perception of learning could be affected by their perspective on the impact of cultural diversity in teams. However, we are not so concerned about the possible effect of the field experiment on team members' self-assessed learning outcomes. Students and professors were uninformed about the experiment and this was confirmed in a couple of interviews.

Table 7. Learning and cultural diversity

	Development in knowledge of								
	Business			Entrepreneurship			Leadership		
	OLS (1)	Median (2)	Robust (3)	OLS (4)	Median (5)	Robust (6)	OLS (8)	Median (9)	Robust (10)
A: Linear									
% minorities	0.275 (0.436)	0.054 (0.485)	0.241 (0.471)	-0.090 (0.351)	0.061 (0.590)	-0.138 (0.389)	0.237 (0.393)	0.032 (0.660)	0.232 (0.465)
R^2	0.34	0.24	0.42	0.34	0.18	0.32	0.30	0.16	0.27
B: Polynomial									
% minorities	-2.670 (1.970)	-1.874 (2.385)	-2.659 (2.476)	-3.455** (1.727)	-3.835 (2.889)	-3.669 (2.540)	-3.789** (1.799)	-3.825 (2.617)	-3.999** (1.903)
% minorities ²	2.690 (1.656)	1.891 (2.035)	2.514 (2.157)	3.088** (1.557)	3.228 (2.568)	3.192 (2.344)	3.691** (1.650)	3.468 (2.494)	3.891** (1.810)
<i>Minimum</i>	0.50	0.50	0.53	0.56	0.59	0.57	0.51	0.55	0.51
R^2	0.38	0.27	0.47	0.39	0.25	0.37	0.38	0.24	0.33
C1: Spline (0.50)									
1st segment	-1.067 (0.919)	-1.331 (1.182)	-0.897 (1.244)	-1.471 (0.913)	-2.057* (1.088)	-1.741 (1.138)	-1.324* (0.760)	-1.705 (1.029)	-1.607 (1.090)
2nd segment	1.049* (0.623)	0.867 (0.656)	0.779 (0.622)	0.753 (0.530)	0.858 (0.745)	0.752 (0.624)	1.191** (0.549)	1.131 (0.852)	1.374** (0.686)
R^2	0.38	0.26	0.46	0.39	0.26	0.37	0.37	0.26	0.36
C2: Spline (0.55)									
1st segment	-0.786 (0.649)	-0.731 (0.811)	-0.881 (0.798)	-1.270** (0.646)	-1.833** (0.846)	-1.476** (0.693)	-1.142* (0.635)	-1.247 (0.749)	-1.287* (0.712)
2nd segment	1.321** (0.656)	0.908 (0.757)	1.091 (0.685)	1.130* (0.599)	1.219 (0.876)	1.114 (0.711)	1.662*** (0.626)	1.243 (0.869)	1.729** (0.729)
R^2	0.38	0.28	0.49	0.41	0.27	0.41	0.41	0.28	0.37
C3: Spline (0.60)									
1st segment	-0.493 (0.638)	-0.638 (0.686)	-0.731 (0.625)	-0.971* (0.553)	-1.568** (0.764)	-1.179** (0.569)	-0.831 (0.525)	-1.175 (0.756)	-0.896 (0.655)
2nd segment	1.585* (0.873)	1.304 (1.144)	1.459 (1.100)	1.480* (0.793)	1.540 (1.209)	1.499 (1.046)	2.133** (0.883)	2.502* (1.345)	2.129* (1.228)
R^2	0.38	0.28	0.51	0.40	0.25	0.40	0.41	0.25	0.36

Note: Based on information from 43 teams. All specifications include team level controls for size, field of study, Trade management Asia and average knowledge levels at baseline. OLS, Median and Robust refer to the estimation method. Bootstrapped standard errors in parentheses (1000 replications). *** **/* denotes significance at the 1%/5%/10%-level.

sions at the individual level, that we run separately for students of Dutch and non-Dutch origin, show that the learning benefits of more cultural diversity accrue to similar extents to both groups. Moreover, and again consistent with the information-processing explanation, the data show some evidence that culturally diverse teams enter the entrepreneurship program at the start with a more diverse pool of relevant knowledge than less heterogeneous teams. Table A2 in the appendix indicates that complementarities between team members and the coefficients of variation of business, entrepreneurship and leadership knowledge at baseline tend to be larger in diverse teams.²⁴

The relationship between cultural diversity and product differentiation is addressed by analyzing the various products and services of teams (see Table 1). Culturally diverse teams potentially benefit from targeting a more diverse and larger market. To this end, we have categorized the ventures along the following dimensions: high/low technology, product/service and market orientation. The analysis of venture types does not reveal a systematic pattern between cultural diversity and type of venture started. Hence, product differentiation is unlikely to explain why culturally diverse teams perform better than homogeneous and moderately heterogeneous teams.

At the cost side, we investigate the role of relationship-oriented conflicts (e.g., Bendersky and Hays, 2012; Jehn, 1995; Stahl et al., 2010) and the role of categorization (or similarity/attraction) processes based on surface level differences between team members (Bell, 2007; Harrison et al., 1998, 2002; Van Knippenberg and Schippers, 2007).

Relationship-oriented conflicts are identified by asking students to what extent there was (non-task related) conflict or disagreement within their teams on a 5-points scale in the posttreatment survey.²⁵ This kind of conflicts are expected to occur more frequently in culturally diverse teams exerting a negative impact on business performance due to, for example, lower satisfaction, intent to remain in the group, and effort. Indeed, (non-tabulated) results show that cultural homogeneous and moderately heterogeneous teams tend to experience less relationship-oriented conflicts than diverse teams.

²⁴In Table A2 a diverse pool of relevant knowledge at the start of the program is operationalized by: (i) complementarities between the self-assessed knowledge that team members have in business, entrepreneurship and leadership, and (ii) the coefficients of variation of business, entrepreneurship and leadership knowledge in teams at baseline. Complementarities are constructed by first standardizing all three knowledge dimensions, subsequently computing the teams' maximum for each knowledge dimension, and then determining the teams' minimum of the maximums of all three knowledge dimensions. Supposedly, if students of different cultural background complement each others knowledge, these minimums are higher in culturally diverse teams.

²⁵The literal question was: "To what extent was there conflict or disagreement between the team members within your team about personal matters (that did not have anything to do with performing the tasks)?" Possible answers were: (almost) never, sometimes, regularly, often, (almost) always.

We measure categorization (or similarity/attraction) processes in terms of a Blau’s diversity index (1977) that categorizes team members on the basis of visible differences between Blacks, Whites, Latinos, Asians and Mediterraneans (see footnote 14). Note that this measure includes both surface level and deeper level aspects of cultural diversity based on categories of visibly different ethnic groups. As a consequence, Blau’s diversity index does not directly measure the possibly costly impact of categorization processes in isolation. The (non-tabulated) results from regressing the different business outcomes on Blau’s diversity index of visible differences between team members closely resemble the pattern obtained in the previous subsection: only if diversity is sufficiently substantial the net effect of diversity on the performance of teams is positive. This suggests that the potentially costly effect of surface level categorization due to cultural diversity in teams is outweighed by the beneficial impact of deep level differences between team members.

All in all, based on our exploratory analysis of mediating mechanisms we find support for the idea that cultural diversity benefits complementarities and (mutual) learning, in line with the information-processing perspective (see Section 2) and that it is associated with more (costly) relationship-oriented conflicts. On the other hand, we find no evidence for more (beneficial) product differentiation in diverse teams, nor for categorization (or similarity/attraction) processes based on surface level differences. Apparently, since the net effect is positive, the evidence suggests that the benefits from complementarities and learning are dominating the possible costs associated with conflicts.

6 Discussion and conclusion

This paper shows that the causal impact of cultural diversity in teams on their business performance is insignificant for a large part of the distribution. The relationship between team performance and cultural diversity in our field experiment turns out to be flat or slightly declining for the subsample of teams where the majority of team members is from the same (Dutch) cultural background. Currently, this is the team composition of most teams in (business) practice. Indeed, this part of our result is in line with earlier results from analyzing real-world teams. However, we find a different result for teams in which the majority of their members have different cultural backgrounds. In case of a culturally diverse team, more cultural diversity in the team turns out to have a significantly positive marginal effect on business performance. Hence, it seems that only if cultural diversity is sufficiently substantial the net effect of cultural diversity on team performance is positive.

The case of truly culturally diverse teams has not been studied before, but will probably become very realistic in the near future. Our findings provide suggestive evidence consistent with a threshold or tipping point. Malcolm Gladwell (2000, p.29): "The Tipping Point is that magic moment when an idea, trend or social behavior crosses a threshold, tips, and spreads like wildfire. At what point does it become obvious that something has reached a boiling point and is about to tip?" Within truly culturally diverse teams, something starts spreading like wildfire. In line with the information-processing perspective (Pelled et al., 1999; Van Knippenberg and Schippers, 2007), our data suggest that 'this something' truly culturally diverse teams benefit from is a more diverse pool of relevant knowledge facilitating (mutual) learning.

Our study is motivated by the fact that many decisions in organizations are nowadays assigned to teams (Hamilton et al., 2003) that become increasingly diverse due to the changing composition of Western populations (e.g., Ozgen et al., 2011). The diversity of teams may be an interesting instrument to optimize organizational performance. However, studies analyzing the causal effect of cultural diversity on team performance in the longer run have been scarce.

More in general, studies that have separately studied various dimensions of demographic diversity (in relation to business performance) such as gender and cultural background are rare. Yet, a better understanding of the causal and possibly subtle impact of the various dimensions of a team's diversity on organizational performance can only be obtained in this manner. We have tried to fill this gap in the empirical literature by conducting a 'double experiment', where randomization took place both in the gender dimension (see Hoogendoorn et al., 2013) and, in the current study, in the dimension of cultural background. A comparison between the results of both studies suggests that the diversity-performance relationships follow different patterns for gender and cultural diversity in teams. For gender, a marginal increase in diversity seems to be beneficial throughout the whole distribution. On the contrary, for cultural diversity, the marginal effect turns positive only after diversity is substantial already. The sizes of both effects are rather similar. For example, profits increase by about one-third of a standard deviation for every 10 percentage point increase in both the share of women (below a maximum of roughly 0.50) and the share of minorities (above a threshold of approximately 0.55).²⁶

Several limitations pertain to this study. There are discrepancies between the business

²⁶The interaction between gender and cultural diversity does not seem to generate an additional surplus in terms of the business performance of teams.

teams in our study and teams in business practice. Individuals in our teams are relatively young, lack serious labor market experience and some of the teams have unprecedented high degrees of cultural diversity. These characteristics may, to some extent, limit the external validity of our study. Moreover, although advantageous for the internal validity of our study, the random composition of teams is probably not representative of common practice in business. There are few startups where team members are randomly assigned. In this sense, the randomization causes a loss of external validity. For example, Boudreau and Lakhani (2011) show that when individuals have the choice to determine the institutional regime that works for them, they are likely to exert more effort. Finally, like most other studies our experimental design does not allow for a causal interpretation of mediating mechanisms such as (mutual) learning that may lead to higher performance of culturally diverse teams.

The contribution of this field experiment is based on a couple of its appealing characteristics. It was conducted in the context of a worldwide leading entrepreneurship education program in a college in Amsterdam with a truly multi-cultural population. As outside researchers, we were allowed to exogenously vary the cultural composition of otherwise randomly composed teams. Student teams found real companies. Teams' substantial and genuinely joint task with strong incentives to maximize performance of a real company in which team members have time to establish roles and observations of other members closely resembles the functioning of teams in business practice. Since the program requires students to deliver annual reports, we could measure their performance in a uniform way. All in all, this resulted in the opportunity to measure the longer term effect of cultural diversity on team performance in realistic though controlled circumstances for teams conducting activities that are likely to be similar to the activities of management and entrepreneurial teams.

Given the upcoming increase of the share of minorities in the labor force it is likely that any team will become more and more diverse in terms of cultural background. Our study might provide a realistic preview of the impact that a high degree of cultural diversity may have on the performance of teams.

References

- Alesina, A. and La Ferrara, E. (2005). Ethnic diversity and economic performance. *Journal of Economic Literature*, 43(3):762–800.
- Bell, S. (2007). Deep-level composition variables as predictors of team performance: A meta-analysis. *Journal of Applied Psychology*, 92(3):595–614.
- Bellini, E., Ottaviano, G. I., Pinelli, D., and Prarolo, G. (2013). Cultural diversity and economic performance: evidence from european regions. *Geography, Institutions and Regional Economic Performance*, pages 121–141.
- Bendersky, C. and Hays, N. (2012). Status conflict in groups. *Organization Science*, 23(2):323–340.
- Bernstein, R. and Edwards, T. (2008). An older and more diverse nation by midcentury. Technical report, US Census Bureau.
- Bisin, A. and Verdier, T. (2000). Beyond the melting pot: Cultural transmission, marriage, and the evolution of ethnic and religious traits. *Quarterly Journal of Economics*, pages 955–988.
- Blau, P. (1977). *Inequality and Heterogeneity: A Primitive Theory of Social Structure*. New York: Free Press.
- Boisjoly, J., Duncan, G., Kremer, M., Levy, D., and Eccles, J. (2006). Empathy or antipathy? The impact of diversity. *American Economic Review*, 96(5):1890–1905.
- Boudreau, K. and Lakhani, K. (2011). The confederacy of heterogeneous software organizations and heterogeneous developers: Field experimental evidence on sorting and worker effort. In Lerner, J. and Stern, S., editors, *The Rate and Direction of Inventive Activity*, volume 50. National Bureau of Economic Research.
- Card, D., Mas, A., and Rothstein, J. (2008). Tipping and the dynamics of segregation. *Quarterly Journal of Economics*, 123(1):177–218.
- Carter, D., D’Souza, F., Simkins, B., and Simpson, G. (2010). The gender and ethnic diversity of US boards and board committees and firm financial performance. *Corporate Governance: An International Review*, 18(5):396–414.

- Constant, A. and Zimmermann, K. (2008). Measuring ethnic identity and its impact on economic behavior. *Journal of the European Economic Association*, 6(2-3):424–433.
- Cox, T., Lobel, S., and McLeod, P. (1991). Effects of ethnic group cultural differences on cooperative and competitive behavior on a group task. *Academic Management Journal*, 34(4):827–847.
- Crant, J. (1996). The proactive personality scale as a predictor of entrepreneurial intentions. *Journal of Small Business Management*, 34(3):42–49.
- Earley, P. and Mosakowski, E. (2000). Creating hybrid team cultures: An empirical test of transnational team functioning. *Academy of Management Journal*, 43(1):26–49.
- Ely, R. and Thomas, D. (2001). Cultural diversity at work: the effects of diversity perspectives on work group processes and outcomes. *Administrative Science Quarterly*, 46(2):229–273.
- Gladwell, M. (2000). *The Tipping Point: How Little Things Can Make a Big Difference*. New York: Little, Brown & Company.
- Goldberg, L. (1990). An alternative description of personality: The big five factor structure. *Journal of Personality and Social Psychology*, 59(6):1216–1229.
- Guiso, L., Sapienza, P., and Zingales, L. (2008). Social capital as culture. *Journal of the European Economic Association*, 6(2-3):295–320.
- Guiso, L., Sapienza, P., and Zingales, L. (2009). Cultural biases in economic exchange? *Quarterly Journal of Economics*, 124(3):1095–1131.
- Hambrick, D., Cho, T., and Chen, M. (1996). The influence of top management team heterogeneity on firms’ competitive move. *Administrative Science Quarterly*, 41(4):659–684.
- Hambrick, D. and Mason, P. (1984). Upper echelons: The organization as a reflection of its top managers. *Academy of Management Review*, 9(2):193–206.
- Hamilton, B., Nickerson, J., and Owan, H. (2003). Team incentives and worker heterogeneity: An empirical analysis of the impact of teams on productivity and participation. *Journal of Political Economy*, 111(3):465–497.

- Hamilton, B., Nickerson, J., and Owan, H. (2012). *Diversity and Productivity in Production Teams*, chapter in: A. Bryson (ed.), *Advances in the Economic Analyses of Participatory & Labor-Managed Firms*, volume 13, pages 99–138. Bingley: Emerald Group Publishing Limited.
- Hansen, Z., Owan, H., and Pan, J. (2006). The impact of group diversity on performance and knowledge spillover: An experiment in a college classroom. NBER Working Paper.
- Harrison, D., Price, K., and Bell, M. (1998). Beyond relational demography: time and the effects of surface- and deep-level diversity on work group cohesion. *Academic Management Journal*, 41(1):96–107.
- Harrison, D., Price, K., Gavin, J., and Florey, A. (2002). Time, teams, and task performance: changing effects of surface- and deep-level diversity on group functioning. *Academic Management Journal*, 45(5):1029–1045.
- Hartog, J. (2011). *Ethnic Minorities in the Netherlands*, chapter in: M. Kahanec and K. Zimmerman (eds.), *Ethnic Diversity in European Labor Markets: Challenges and Solutions*, pages 198–221. Cheltenham: Edward Elgar Publishing.
- HBO-raad (2010). Feiten en cijfers: Afgestudeerden en uitvallers in het hoger beroepsonderwijs. Technical report, HBO-raad.
- Hofstede, G. (1980). Culture and organizations. *International Studies of Management & Organization*, 10(4):15–41.
- Hofstede, G. (1983). The cultural relativity of organizational practices and theories. *Journal of International Business Studies*, 14(2):75–89.
- Hofstede, G. (1984). Cultural dimensions in management and planning. *Asia Pacific Journal of Management*, 1(2):81–99.
- Hofstede, G. (1985). The interaction between national and organizational value systems. *Journal of Management Studies*, 22(4):347–357.
- Hofstede, G. (1990). A reply and comment on joginder p. singh: ‘managerial culture and work-related values in india. *Organization Studies*, 11(1):103–106.
- Hofstede, G. (2001). *Culture’s consequences: Comparing values, behaviors, institutions and organizations across nations*. Thousand Oaks, second edition.

- Hoogendoorn, S., Oosterbeek, H., and Van Praag, M. (2013). The impact of gender diversity on the performance of business teams: Evidence from a field experiment. *Management Science*, 59(7):1514–1528.
- Horwitz, S. and Horwitz, I. (2007). The effects of team diversity on team outcomes: A meta-analytic review of team demography. *Journal of Management*, 33(6):987–1015.
- Jackson, S., Jackson, J., Sessa, V., Cooper, D., Julin, J., and Peyronnin, K. (1991). Some differences make a difference: Interpersonal dissimilarity and group heterogeneity as correlates of recruitment, promotion, and turnover. *Journal of Applied Psychology*, 76(6):675–689.
- Jehn, K. (1995). A multi-method examination of the benefits and detriments of intragroup conflict. *Administrative Science Quarterly*, 40(2):256–282.
- Jehn, K., Chadwick, C., and Thatcher, S. (1997). To agree or not to agree: the effects of value congruence, member diversity, and conflict on workgroup outcomes. *International Journal of Conflict Management*, 8(4):287–305.
- Jehn, K., Northcraft, G., and Neale, M. (1999). Why differences make a difference: a field study of diversity, conflict, and performance in workgroups. *Administrative Science Quarterly*, 44(4):741–763.
- Joshi, A. and Roh, H. (2009). The role of context in work team diversity research: a meta-analytic review. *The Academy of Management Journal (AMJ)*, 52(3):599–627.
- Kahane, L., Longley, N., and Simmons, R. (2013). The effects of coworker heterogeneity on firm-level output: Assessing the impacts of cultural and language diversity in the national hockey league. *Review of Economics and Statistics*, 95(1):302–314.
- Karlan, D. and Valdivia, M. (2011). Teaching entrepreneurship: Impact of business training on microfinance clients and institutions. *Review of Economics and Statistics*, 93(2):510–527.
- Kilduff, M., Angelmar, R., and Mehra, A. (2000). Top management-team diversity and firm performance: Examining the role of cognitions. *Organization Science*, 11(1):21–34.
- Lane, H., Maznevski, M., DiStefano, J., and Dietz, J. (2009). *International Management Behaviour: Leading with a Global Mindset*. Wiley, sixth edition.

- Lane, H., Maznevski, M., Mendenhall, M., and McNett, J. (2004). *Blackwell handbook of global management: A guide to managing complexity*. Blackwell Publishers.
- Lau, D. and Murnighan, J. (1998). Demographic diversity and faultlines: the compositional dynamics of organizational groups. *Academic Management Review*, 23(2):325–340.
- Lau, D. and Murnighan, J. (2005). Interactions within groups and subgroups: The dynamic effects of demographic faultlines. *Academy of Management Journal*, 48(4):645–659.
- Lazear, E. (1999). Globalisation and the market for team-mates. *Economic Journal*, 109(454):15–40.
- Lee, N. and Nathan, M. (2011). Does cultural diversity help innovation in cities: Evidence from London firms. SERC Discussion Paper.
- Leonard, J., Levine, D., and Giuliano, L. (2010). Customer discrimination. *Review of Economics and Statistics*, 92(3):670–678.
- Mateos, P. (2007). A review of name-based ethnicity classification methods and their potential in population studies. *Population, Space and Place*, 13(4):243–263.
- McLeod, P., Lobel, S., and Cox, T. (1996). Ethnic diversity and creativity in small groups. *Small Group Research*, 27(2):246–264.
- Minniti, M. and Bygrave, W. (2001). A dynamic model of entrepreneurial learning. *Entrepreneurship Theory & Practice*, 25(3):5–16.
- Nathan, M. and Lee, N. (2011). Does cultural diversity help innovation in cities: evidence from london firms. *SERC Discussion Papers*.
- Niebuhr, A. (2010). Migration and innovation: does cultural diversity matter for regional r&d activity? *Regional Science*, 89(3):563–585.
- Nielsen, B. and Nielsen, S. (2013). Top management team nationality diversity and firm performance: a multilevel study. *Strategic Management Journal*, 34(3):373–382.
- Nielsen, S. (2010). Top management team diversity: a review of theories and methodologies. *International Journal of Management Reviews*, 12(3):301–316.

- Oosterbeek, H., Van Praag, M., and IJsselstein, A. (2010). The impact of entrepreneurship education on entrepreneurship skills and motivation. *European Economic Review*, 54(3):442–454.
- Ottaviano, G. and Peri, G. (2006). The economic value of cultural diversity: Evidence from us cities. *Journal of Economic Geography*, 6(1):9–44.
- Ozgen, C., Nijkamp, P., and Poot, J. (2011). The impact of cultural diversity on innovation: Evidence from dutch firm-level data. IZA Discussion Paper.
- Parker, S. (2009). *The Economics of Entrepreneurship*. Cambridge: Cambridge University Press.
- Parrotta, P., Pozzoli, D., and Pytlikova, M. (2010). Does labor diversity affect firm productivity? Norface Migration Discussion Paper.
- Pelled, L., Eisenhardt, K., and Xin, K. (1999). Exploring the black box: An analysis of work group diversity, conflict, and performance. *Administrative Science Quarterly*, 44(1):1–3.
- Richard, O. (2000). Racial diversity, business strategy, and firm performance: A resource-based view. *Academy of Management Journal*, 43(2):164–177.
- Richard, O., Barnett, T., Dwyer, S., and Chadwick, K. (2004). Cultural diversity in management, firm performance, and the moderating role of entrepreneurial orientation dimensions. *Academy of Management Journal*, 47(2):255–266.
- Richard, O., Kochan, T., and McMillan-Capehart, A. (2002). The impact of visible diversity on organizational effectiveness: Disclosing the contents in pandora’s black box. *Journal of Business and Management*, 8(3):265–291.
- Stahl, G., Maznevski, M., Voigt, A., and Jonsen, K. (2010). Unraveling the effects of cultural diversity in teams: A meta-analysis of research on multicultural work groups. *Journal of International Business Studies*, 41(4):690–709.
- Østergaard, C. R., Timmermans, B., and Kristinsson, K. (2011). Does a different view create something new? the effect of employee diversity on innovation. *Research Policy*, 40(3):500–509.

- Tajfel, H. (1982). Social psychology of intergroup relations. *Annual Review of Psychology*, 33(1):1–39.
- Thomas, D. (1999). Cultural diversity and work group effectiveness an experimental study. *Journal of Cross-Cultural Psychology*, 30(2):242–263.
- Tsui, A., Egan, T., and O'Reilly, C. (1992). Being different: relational demography and organizational attachment. *Administrative Science Quarterly*, 37(4):549–579.
- Van Dijk, H., Van Engen, M., and Van Knippenberg, D. (2012). Defying conventional wisdom: A meta-analytical examination of the differences between demographic and job-related diversity relationships with performance. *Behavior and Human Decision Processes*, 119(1):38–53.
- Van Knippenberg, D. and Schippers, M. (2007). Work group diversity. *Annual Review of Psychology*, 58:1008–1022.
- Van Knippenberg, D., Van Knippen, C., and Homan, A. (2004). Work group diversity and group performance: an integrative model and research agenda. *Journal of Applied Psychology*, 89(6):1008–1022.
- Vogel, R., Puhan, T., Shehu, E., Kliger, D., and Beese, H. (2014). Funding decisions and entrepreneurial team diversity: A field study. *Journal of Economic Behavior & Organization*, 107:595–613.
- Watson, W., Kumar, K., and Michaelson, L. (1993). Cultural diversity's impact on interaction process and performance: comparing homogeneous and diverse task groups. *Academic Management Journal*, 36(3):590–602.
- Wernerfelt, B. (1984). A resource?based view of the firm. *Strategic Management Journal*, 5(2):171–180.
- Wiersema, M. and Bantel, K. (1992). Top management team demography and corporate strategic change. *Academy of Management Journal*, 35(1):91–121.
- Wiersema, M. and Bird, A. (1993). Organizational demography in japanese firms: Group heterogeneity, individual dissimilarity, and top management team turnover. *Academy of Management Journal*, 36(5):996–1025.

- Williams, K. and O'Reilly, C. (1998). Demography and diversity in organizations: A review of 40 years of research. *Research in organizational behavior*, 20(20):77–140.
- Zellmer-Bruhn, M., Maloney, M., Bhappu, A., and Salvador, R. (2008). When and how do differences matter? an exploration of perceived similarity in teams. *Organizational behavior and human decision processes*, 107(1):41–59.
- Zenger, T. and Lawrence, B. (1989). Organizational demography: The differential effects of age and tenure distributions on technical communicatio. *Academy of Management Journal*, 32(2):353–376.
- Zhao, H. and Seibert, S. (2006). The big five personality dimensions and entrepreneurial status: A meta-analytical review. *Journal of Applied Psychology*, 91(2):259–271.

Appendix Table A1. Cultural diversity and team performance (spline functions with three segments)

	Sales			Profits			Pos. profits			Profits per share		
	OLS (1)	Median (2)	Robust (3)	OLS (4)	Median (5)	Robust (6)	OLS (7)	OLS (8)	Median (9)	Robust (10)		
C1: Spline (0.50 and 0.60)												
1st segment (N=13)	-72.3 (3383.4)	-903.9 (2832.8)	-1475.2 (2542.9)	-25.5 (1305.0)	19.5 (1091.9)	-298.3 (1186.7)	-0.566 (1.553)	5.519 (20.586)	-1.568 (21.003)	-1.451 (21.464)		
2nd segment (N=12)	-3855.7 (5694.8)	-1702.8 (4160.3)	379.6 (3194.1)	-1665.6 (1924.9)	-2380.6 (1993.9)	-1827.1 (1789.0)	-3.438 (2.537)	-40.931 (34.724)	-44.992 (41.217)	-45.332 (38.860)		
3rd segment (N=18)	1121.1 (1511.8)	1863.9 (1730.1)	832.5 (1324.0)	1486.8** (750.4)	1533.8* (836.2)	1272.7* (743.1)	3.431*** (1.329)	33.194** (13.609)	30.027* (17.104)	33.510** (15.465)		
R^2	0.05	0.05	0.07	0.09	0.09	0.16	0.20	0.09	0.09	0.14		
C2: Spline (0.45 and 0.65)												
1st segment (N=10)	-1982.7 (3386.2)	-4112.3 (3331.4)	-3445.7 (3215.0)	-70.7 (1601.9)	-729.5 (1746.8)	-546.6 (1919.0)	-0.833 (2.146)	9.147 (28.721)	-14.834 (31.870)	-3.369 (37.745)		
2nd segment (N=21)	-441.6 (2170.7)	-52.1 (1717.5)	660.5 (1429.9)	-362.8 (782.4)	-376.4 (971.8)	-463.8 (727.6)	-0.804 (1.413)	-13.556 (17.582)	-6.730 (22.359)	-14.604 (18.522)		
3rd segment (N=12)	593.0 (1832.0)	1953.7 (1960.7)	869.7 (1948.5)	1233.6 (816.3)	1519.5 (1036.6)	1091.6 (952.3)	3.130* (1.753)	32.469* (16.937)	25.126 (24.939)	32.904 (21.982)		
R^2	0.04	0.06	0.14	0.05	0.06	0.08	0.13	0.06	0.05	0.08		
C3: Spline (0.40 and 0.60)												
1st segment (N=6)	-3825.4 (4174.3)	-8217.8 (4985.2)	-5895.6 (5886.8)	542.9 (2788.9)	1449.7 (3150.5)	-107.4 (4286.7)	1.138 (3.301)	26.214 (48.628)	34.250 (53.268)	9.394 (77.093)		
2nd segment (N=19)	179.6 (1455.4)	749.1 (1535.9)	662.8 (1291.8)	-1070.0 (828.7)	-1424.8* (807.7)	-1119.5 (680.6)	-2.768** (1.213)	-26.479 (17.121)	-32.143* (16.135)	-28.274 (18.670)		
3rd segment (N=18)	-91.0 (1172.0)	786.2 (1482.4)	487.1 (1219.8)	1426.1** (685.2)	1340.6* (693.3)	1187.0** (584.7)	3.475*** (1.169)	32.448*** (12.559)	27.426** (13.511)	31.785** (13.311)		
R^2	0.05	0.08	0.22	0.10	0.10	0.15	0.22	0.11	0.10	0.14		

Note: Based on information from 43 teams. All specifications include a team level dummy for the field of study Trade management Asia. OLS, Median and Robust refer to the estimation method. Median and robust specifications for positive profits are excluded since this variable is dichotomous. Bootstrapped standard errors in parentheses (1000 replications). ***/**/* denotes significance at the 1%/5%/10%-level.

Appendix Table A2. Complementarities and diversity in relevant knowledge

	Complementarities			Diversity in knowledge of									
	Business/E'ship/Leadership			Business			Entrepreneurship			Leadership			
	OLS (1)	Median (2)	Robust (3)	OLS (4)	Median (5)	Robust (6)	OLS (7)	Median (8)	Robust (9)	OLS (10)	Median (11)	Robust (12)	
A: Linear													
% minorities	0.690 (0.533)	1.397 (0.872)	1.181 (0.920)	0.129* (0.076)	0.186 (0.127)	0.126 (0.094)	0.105 (0.109)	0.079 (0.160)	0.096 (0.132)	-0.059 (0.071)	-0.047 (0.105)	-0.049 (0.084)	
R ²	0.28	0.22	0.41	0.54	0.32	0.46	0.17	0.03	0.14	0.23	0.13	0.19	
B: Polynomial													
% minorities	-3.776 (3.132)	-4.352 (4.273)	-2.763 (4.155)	-0.371 (0.418)	-0.639 (0.662)	-0.402 (0.502)	-0.766 (0.520)	-1.038 (0.809)	-0.786 (0.657)	-0.597 (0.364)	-0.272 (0.580)	-0.599 (0.446)	
% minorities ²	4.065 (2.697)	4.517 (3.784)	3.828 (3.742)	0.457 (0.371)	0.670 (0.574)	0.483 (0.433)	0.800* (0.475)	0.978 (0.733)	0.811 (0.605)	0.493 (0.323)	0.211 (0.518)	0.503 (0.402)	
Minimum	0.46 (0.35)	0.48 (0.24)	0.36 (0.58)	0.41 (0.56)	0.48 (0.34)	0.42 (0.51)	0.48 (0.23)	0.53 (0.10)	0.48 (0.20)	0.61 (0.28)	0.64 (0.15)	0.60 (0.24)	
C1: Spline (0.50)													
1st segment	-1.376 (1.318)	-1.680 (1.862)	-1.321 (1.754)	-0.025 (0.179)	0.040 (0.292)	-0.041 (0.216)	-0.239 (0.208)	-0.356 (0.315)	-0.246 (0.247)	-0.245 (0.164)	-0.156 (0.259)	-0.229 (0.185)	
2nd segment	1.878** (0.812)	2.126* (1.243)	2.144* (1.253)	0.218* (0.124)	0.234 (0.174)	0.224 (0.143)	0.315* (0.170)	0.309 (0.260)	0.311 (0.213)	0.054 (0.104)	0.027 (0.183)	0.061 (0.126)	
R ²	0.37	0.26	0.40	0.55	0.33	0.50	0.23	0.09	0.19	0.27	0.14	0.22	
C2: Spline (0.55)													
1st segment	-0.526 (1.053)	-0.054 (1.437)	-0.392 (1.325)	-0.016 (0.148)	-0.159 (0.232)	-0.033 (0.182)	-0.177 (0.162)	-0.272 (0.244)	-0.183 (0.191)	-0.201 (0.135)	-0.150 (0.185)	-0.189 (0.144)	
2nd segment	1.889* (1.000)	1.960 (1.649)	2.167 (1.451)	0.273* (0.150)	0.330 (0.217)	0.286 (0.188)	0.397** (0.196)	0.464 (0.319)	0.391 (0.258)	0.086 (0.122)	0.031 (0.218)	0.099 (0.159)	
R ²	0.34	0.24	0.37	0.56	0.34	0.51	0.24	0.11	0.19	0.27	0.14	0.22	
C3: Spline (0.60)													
1st segment	0.022 (0.929)	0.783 (1.312)	0.790 (1.246)	-0.012 (0.119)	-0.068 (0.190)	-0.032 (0.158)	-0.134 (0.139)	-0.167 (0.216)	-0.147 (0.173)	-0.172 (0.111)	-0.143 (0.167)	-0.167 (0.124)	
2nd segment	1.831 (1.335)	1.933 (1.932)	2.830 (1.843)	0.370* (0.190)	0.401 (0.279)	0.397* (0.229)	0.532** (0.254)	0.507 (0.395)	0.534 (0.329)	0.141 (0.165)	0.054 (0.271)	0.157 (0.197)	
R ²	0.31	0.22	0.57	0.56	0.35	0.52	0.26	0.11	0.21	0.27	0.14	0.23	

Note: Based on information from 43 teams. All specifications include team level controls for size, field of study Trade management Asia and average knowledge levels at baseline. OLS, Median and Robust refer to the estimation method. Bootstrapped standard errors in parentheses (1000 replications). *** **/* denotes significance at the 1%/5%/10%-level.