

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 cultural background. We measure the impact of cultural diversity on the performance of business teams using a field experiment. Companies are set up by teams of undergraduate students in business studies in realistic though similar circumstances. We vary the cultural composition of otherwise randomly composed teams in a multi-cultural student population. Our data indicate that a moderate level of cultural diversity has no effect on team performance in terms of business outcomes (sales, profits and profits per share). However, if at least the majority of team members is culturally diverse then more cultural 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 culturally diverse teams.

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

Keywords: cultural diversity, ethnic diversity, team performance, field experiment, entrepreneurship, (mutual) learning

*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), University of Aalborg, Barcelona (DRUID), Berlin (CASE Distinguished Lectures, Humboldt University), Bonn (EALE), Boston (AOM), Oslo (National Science Foundation), Potsdam (IZA Workshop) and Utrecht (School of Economics) 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).

Recent developments have strengthened the relevance of diversity for organizations: (i) the share of women in higher education and the labor force has increased, (ii) local populations have become more culturally diverse in a globalized world and the share of minorities in Western populations is rising sharply, and (iii) performance increasingly depends on human capital and the accumulation of knowledge in information-based economies.

Teams have become increasingly important as decision making bodies in many sorts of public and private organizations (Hamilton et al., 2003). Therefore, team diversity can be employed as an organizational tool to create value. Consequently, the effective composition or diversity of teams has become an interesting topic of research. Yet, the precise impact of diversity in teams on their performance may be subtle. And so far, most studies have measured several dimensions of a team's or organization's diversity simultaneously (for example Hansen et al., 2006), hindering the measurement of the possibly subtle impact of each diversity dimension in isolation.

We have performed a 'double' randomized field experiment where randomization took place based on two orthogonal dimensions, i.e., gender and cultural background.¹ This enables us to measure the separate effects on performance of the two most important demographic dimensions of diversity in business teams. The gender experiment has been discussed in Hoogendoorn et al. (2013). This is the companion paper about cultural diversity. In the current study, we measure the effect of team composition in terms of cultural diversity on the performance of business teams using the field experiment. The choice for this topic and approach are easily motivated. One of the most salient dimensions of team heterogeneity is cultural background (Alesina and La Ferrara, 2005). Cultural diversity implies heterogeneity in (mother) languages, religions, races and cultures (Alesina

¹See footnote 7 in Hoogendoorn et al., 2013.

and La Ferrara, 2005). It is commonly measured based on country of birth, of the individual or of his/her parents. Cultural diversity also coincides with a variety of norms, information sets, knowledge and ability levels (Lazear, 1999; Morgan and Vardy, 2009). This variety affects the formation and performance of teams. Cultural diversity would benefit team performance due to a more diverse pool of skills and knowledge that leads to complementarities and (mutual) learning. For example, due to complementarities and learning opportunities, culturally diverse teams are associated with more creativity and innovation (Alesina and La Ferrara, 2005; Lee and Nathan, 2011; Ozgen et al., 2011b). On the other hand, the costs associated with more cultural diversity would be related to more difficult communication and coordination (Lazear, 1999; Morgan and Vardy, 2009).² All in all, cultural diversity is a potentially influential source of heterogeneity.

Cultural diversity is highly relevant in an increasingly globalized world. Multinational firms often staff teams internationally and local populations - especially in big cities - become more mixed and multicultural. Cultural diversity is a current fact of life and the share of minorities in Western populations is increasing sharply (Alesina and La Ferrara, 2005; Lazear, 1999; Ozgen et al., 2011b). 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). As a consequence, it is likely that any team will become more and more diverse in terms of cultural background, even if the optimal team formation would indicate otherwise.

The relevance and potential impact of cultural diversity in teams motivate our choice for the topic of this study. Our approach of a field experiment among business teams that start up in identical circumstances is motivated as follows. Team formation is obviously driven by prospective productive consequences. If the situation carries a higher likelihood that an culturally diverse team is beneficial, the team composition will be more mixed (Boisjoly et al., 2006). Hence, the measured effects of cultural diversity on performance in real-world teams are likely to be biased due to endogenous team composition. Examples of studies conducted on the cultural diversity of teams in real organizations include Carter et al. (2010), Hamilton et al. (2012), Kahane et al. (2013), Leonard et al. (2010) and

²Cultural diversity may also affect group formation and performance through its influence on the group culture (Earley and Mosakowski, 2000; Richard et al., 2004) and the strategic behavior of group members (Alesina and La Ferrara, 2005).

Parrotta et al. (2010), and their results are ambiguous (Alesina and La Ferrara, 2005).³

Experiments in the lab have established results that are largely consistent with the theory proposed by Lazear (1999): an optimal degree of heterogeneity is determined by the trade-off between the benefits of more cultural diversity and the associated increased costs of communication and coordination (Alesina and La Ferrara, 2005). However, laboratory studies have a limited resemblance to real-world situations, typically measuring short-term effects. As Hoogendoorn et al., 2013 argued, the consequences of a team's diversity in terms of, for example, coordination and communication or complementarities and learning are not likely to become evident instantaneously (Boisjoly et al., 2006). It is thus useful to study the effects of team composition in the longer run and preferably in more realistic circumstances (in which people develop real interactions).

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. Hansen et al. (2006) measure the impact of demographic diversity (age, gender and cultural background) in student groups of four to five students on the team's academic performance and find no effect.⁴ Boisjoly et al. (2006) find that attitudes and behaviors change when people of different cultural background are randomly assigned to live together at the start of their first year of college. White students assigned to African-American roommates show to be significantly more empathetic to these groups.

We conclude that randomized controlled trials measuring the causal impact of a team's cultural diversity on its performance are scarce, or actually, non-existent. Cultural background is often considered a sensitive (and sometimes not directly observable) dimension of team diversity. Studying longer term team dynamics and outcomes in realistic business situations for a wide variety of culturally (more and less) diverse teams outside a laboratory can be a fundamental contribution. Especially if the experiment takes place among business teams that have arguably the same kind of tasks and objectives as venture and

³Other studies examine the effects of cultural diversity on productivity at the country (Guiso et al., 2009; Montalvo and Reynal-Querol, 2005), region (Ozgen et al., 2011a) or city level (Lee and Nathan, 2011; Ottaviano and Peri, 2006). Related are also studies measuring the impact of cultural composition of schools or neighborhoods on educational outcomes (e.g., Angrist and Lang, 2004; Aslund et al., 2011; Card and Rothstein, 2007; Hanushek et al., 2009; Hoxby, 2000).

⁴Group composition is random and no exogenous stratification is imposed. Teams are required to select one of three contract forms that determine the authority of grading. The drawback of this design, when interested in the bare effect of cultural diversity on team performance, is that the effect of interest may be confounded by the contract choice (that may be related to cultural diversity).

management teams that put, obviously, a very heavy stamp on value creation in general.

This kind of measurement is the objective of our study. To this end, we conducted a field experiment in the context of a compulsory entrepreneurship program for undergraduate students in international business at the Amsterdam University of Applied Sciences. We follow 550 students who set up and manage 45 real companies as part of their curriculum in the academic year of 2008-2009. We exploit the fact that the - otherwise homogeneous - population consists of 55% students with a non-Dutch cultural background from 53 different countries of origin and that we, as outside researchers, were allowed to exogenously vary the cultural composition of otherwise randomly composed teams. The resulting percentage of students with a non-Dutch cultural background per team varies from 20% to 90%. We use a rather common definition of cultural background, i.e., parents' country of birth.

Our field experiment implies a level of controlled circumstances comparable to the laboratory: teams start up companies simultaneously and under equal circumstances and all members come from a pool of students, minority or Dutch, of the same caliber. The substantial variation (20-90%) enables us to capture possible non-linearities in the relationship between cultural diversity and team performance. Moreover, this creates the novel opportunity to study teams in which the majority of its members is diverse. Currently, this circumstance is not yet realistic. However, this is rapidly changing.

At the same time, the entrepreneurship program and the truly joint task with strong incentives to maximize sales, profits and shareholder value of a company in a real market resemble realistic circumstances that are not likely to be mimicable in a laboratory setting. In particular, these circumstances give rise to the formation of a real team in which people have time to establish roles and observations of other team members.

What kind of results might we expect? We combine Lazear's argument (1999) that there is a trade-off between the costs and benefits of an culturally more diverse team with recent insights from Earley and Mosakowski (2000) and Richard et al. (2004) based on Blau's theory of heterogeneity (1977). The non-formal models in these studies refine Lazear's argument by allowing the costs of communication and coordination to be a non-linear function of cultural diversity. Moderately heterogeneous teams would incur higher costs of communication and coordination than teams that are homogeneous or het-

erogeneous in terms of cultural background. In these moderately heterogeneous teams subgrouping along cultural lines may have negative effects such as distrust, conflicts or stereotyping between distinct subgroups.

The resulting non-linear relationship between communication costs and cultural diversity could perhaps be illustrated by considering three types of team compositions: (i) with a low percentage of minorities, (ii) with a substantial subgroup of minorities, and (iii) with a majority group of minorities. In the first situation, the communication costs are probably low. The few minorities present will perhaps not participate in the team process and be left aside. This is costly in the sense that their productivity is lost, but communication can be low cost too and based on habits, language and norms of the majority group (i.e., with a single identity). In the second situation, the group that may not take part in the productive process of the majority is larger. Segregated subgroups may be formed by the majority and the others with distinct manners of work. Hence, the costs of diversity are higher at this stage. In the third situation, in a truly culturally diverse team, the costs of communication have probably become lower due to the absence of subgroups and the desirability for every team member to participate in a hybrid team culture with a diverse identity (Earley and Mosakowski, 2000; Richard et al., 2004). We thus expect that the costs of cultural diversity follow an inverse u-shaped pattern, whereas the benefits of complementarities and (mutual) learning may be an(y) increasing function of cultural diversity (Lazear, 1999).

Due to the fact that we do not have any conjectures about the specific forms of the cost and benefit functions relative to each other, the net effect of cultural diversity on team performance remains a question that needs to be answered empirically. If anything, the net benefits are expected to start increasing after a certain threshold. Before that threshold has been reached, minorities, as a small part of the team, tend to be 'marginalized' or form a different group, but such marginalization tends to disappear when the share of minorities is sufficiently large.

Our empirical analysis shows a flat or even slightly negative impact of a team's cultural diversity on its performance as measured in terms of business outcomes but this seems to turn positive when the team starts to be truly heterogeneous, i.e., when the majority of the team is culturally diverse. The flat or slightly negative effect that we find for the

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 and this novelty is possibly due to the fact that we have been the first to study the very right handside of the distribution of cultural diversity.

Altogether, these results are consistent with the expectations based on narrative theories, discussed before. They suggest 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). With respect to underlying mechanisms, our data are consistent with the idea that more culturally diverse teams tend to have a more diverse pool of relevant knowledge facilitating more (mutual) learning that may lead to better business results. However, like other studies, our tests of underlying mechanisms that we explore, such as learning, cannot provide causal evidence.

In what follows, Section 2 provides information on the context and design of the field experiment that are similar to the field experiment described in Hoogendoorn et al. (2013). Therefore parts of these descriptions repeat theirs or are omitted with reference to this companion paper. Section 3 describes the data and randomization checks. Section 4 shows the empirical findings. Section 5 provides a discussion and conclusion.

2 Context and design

2.1 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. Students thus frequently interact, build up relationships, and create routines and processes to achieve their common goal.⁵ Table 1 lists team characteristics such as their cultural composition, size, business performance and product or service.

Cultural background

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). The share of students in our sample with a non-Dutch cultural background is 0.55 and the percentage of students with a non-Dutch cultural background varies per team from 20% to 90%.⁶ We measure cultural diversity dichotomously. This is consistent with previous empirical studies (Boisjoly et al., 2006; Carter et al., 2010; Hamilton et al., 2012; Hansen et al., 2006) as well as with our assignment of students to teams.⁷ We do not use any other more refined measures of cultural background by distinguishing *degrees* or *kinds* of minorities such as country of birth, degree of integration, language or religion. Students with a non-Dutch cultural background are born abroad in 38% of the cases, with 82% of the fathers born abroad and 87% of the mothers. In 69% of the cases both the father and the mother are born abroad.

Table 2 shows that 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 (panel A). Only for the purpose of presentation, the descriptive statistics in panel B are

⁵The circumstances, usual processes (like the election of officers) and (support) procedures experienced by student teams as well as the strong (individual and team) incentives that are in place to align students' interests with the business performance of their company can be found in Hoogendoorn et al. (2013). The relationship between the cultural composition of teams and their allocation of roles across cultural backgrounds is discussed below.

⁶The 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.

⁷An alternative way of measuring cultural diversity takes the country of (parental) origin into account and is expressed in terms of Blau's diversity index (1977), $1 - \sum_{i=1}^N p_i^2$. Here, p is the share of individuals in a (cultural) category i and N the number of (cultural) categories. Disadvantages of Blau's diversity index are that it: (i) implicitly assumes that each category is different from another *to the same extent* and (ii) turns out rather sensitive to the number of (cultural) categories in the sample. The upside, though, may be that it takes into account that diversity is larger when individuals originate from all distinct countries than when they or their parents are all born in the same foreign country. However, our initial assignment of students to teams and, hence, the exogenous variation in teams' diversity is based on the dichotomous measure of cultural background. As a consequence, we can only use that measure convincingly to study the causal effect of cultural diversity on team performance.

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	Flpthat	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).

collapsed into groups (see Parrotta et al., 2010). They indicate that cultural variation among minority students is substantial. Panel C shows that there are virtually no clusters of minorities at the team level. The vast majority of minority students has no team members from the same country of origin (70%), whereas 19% of the minority students has only one team member from the same country of origin (the average number of students with a non-Dutch cultural background per team is 6.8 at baseline). Hence, we can safely assume that the share of students with a non-Dutch cultural background is a true indicator of the team’s cultural diversity.⁸ 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 below.

2.2 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 cultural 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 background 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.

⁸Only two teams form an exception: one accommodates six and the other four members from the same country of origin. Excluding these two teams from the sample does not change the main results. Robustness checks indicate that our findings also remain similar when we control for the number of different countries of origin.

Table 2. Descriptive statistics of cultural diversity

A: Cultural background (team level)		Mean	SD	Min	Max
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 background (collapsed into groups)		Fraction			
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		Fraction			
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		Students	Teams	Share of minorities (per team)	
Business management		240	18	Average	Minimum
Management		60	5	0.52	0.20
Trade management Asia		105	9	0.54	0.31
Business languages		118	11	0.66	0.42
Financial management		27	2	0.50	0.22
Total		550	45	0.66	0.58
				0.55	0.20

Note: 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.

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.⁹ 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 minorities 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 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.¹⁰ Moreover, students of non-Dutch cultural background are not significantly more or less likely to be part of the management team than students of Dutch cultural background 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.

3 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

⁹Dropout 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.

¹⁰These 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. 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
Conflicts	1-5	2.23	0.59	1.00	3.67
Atmosphere	1-5	3.53	0.55	2.33	4.83

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.

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 processes regarding complementarities, (mutual) learning, communication and coordination.¹¹

Table 3 reports descriptive statistics of individual and team characteristics. The pre-treatment 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 and include them in our analyses of complementarities and learning (see Karlan and Valdivia, 2011; Minniti and Bygrave, 2001).¹² We use the individual development of these knowledge areas

¹¹In Hoogendoorn et al. (2013) other processes were analyzed too, for example, those related to monitoring and decision making. In this paper, nonformal theories provide more guidance as to which processes would matter, enabling us to consider a more limited set.

¹²The 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.

during the program as a proxy for learning. The levels reported in Table 3 are knowledge levels at baseline on a scale from 1 to 5. These initial levels also form the basis for our constructed measure of complementarities. As we will discuss more thoroughly in Section 4, self-assessments of team members about the team's atmosphere and the prevalence of conflicts are proxies for communication and coordination in teams.

Business performance metrics are gathered or calculated from the companies' annual reports that we obtained from 43 out of 45 teams. 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 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.¹³

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 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. However, 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).

¹³Shares 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	High	Low	Moderate	High
		share<0.45 (2)	0.45≥share≤0.65 (3)	share>0.65 (4)	share<0.40 (5)	0.40≥share≤0.60 (6)	share>0.60 (7)
Sales (euros)	Mean	838	881	745	1183	826	735
	SD	707	891	349	622	895	468
	Min	0	0	265	593	130	0
	Max	4209	4209	1267	2226	4209	1953
Profits (euros)	Mean	-69	-146	41	-97	-78	-51
	SD	318	346	164	468	294	305
	Min	-1016	-1016	-247	-1011	-848	-1016
	Max	477	477	294	294	477	294
Positive profits (0/1)	Mean	0.51	0.29	0.75	0.67	0.37	0.61
	SD	0.51	0.46	0.45	0.52	0.50	0.50
	Min	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
Profits per share (euros)	Mean	-0.51	-2.07	1.46	-1.84	-0.58	0.01
	SD	6.42	6.33	5.07	7.05	6.64	6.29
	Min	-15.48	-12.87	-8.04	-15.48	-11.78	-12.87
	Max	15.64	11.36	12.79	3.67	15.64	12.79
N	43	10	21	6	19	18	

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.

Randomization

Before presenting results, 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'. 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. 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' 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 dummy to control for the field of study 'trade management Asia'.

4 Results

4.1 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 in business. 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 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 performance 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 specification in Panel A. Moreover, they do not impose equal curvatures of opposite signs below and above the fraction of minority members at which team performance is lowest, as is the case with a quadratic specification (see Panel B). In particular, using spline functions and focussing on the right handside of the distribution, we may find some indicative evidence of rising net benefits after a certain tipping point as was suggested by

Table 6. Cultural diversity and team performance

	Sales			Profits			Pos. profits			Profits per share		
	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	-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	<i>0.02</i>	<i>0.03</i>	<i>0.03</i>	<i>0.02</i>	<i>0.03</i>	<i>0.00</i>	<i>0.03</i>	<i>0.01</i>	<i>0.01</i>	<i>0.00</i>		
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)		
<i>Minimum</i>	<i>0.65</i>	<i>0.53</i>	<i>0.58</i>	<i>0.51</i>	<i>0.57</i>	<i>0.57</i>	<i>0.53</i>	<i>0.48</i>	<i>0.55</i>	<i>0.54</i>		
R^2	<i>0.04</i>	<i>0.05</i>	<i>0.11</i>	<i>0.05</i>	<i>0.06</i>	<i>0.08</i>	<i>0.11</i>	<i>0.04</i>	<i>0.05</i>	<i>0.07</i>		
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	<i>0.02</i>	<i>0.05</i>	<i>0.08</i>	<i>0.05</i>	<i>0.05</i>	<i>0.08</i>	<i>0.11</i>	<i>0.03</i>	<i>0.05</i>	<i>0.07</i>		
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	<i>0.03</i>	<i>0.05</i>	<i>0.07</i>	<i>0.06</i>	<i>0.08</i>	<i>0.13</i>	<i>0.17</i>	<i>0.05</i>	<i>0.07</i>	<i>0.10</i>		
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	<i>0.03</i>	<i>0.05</i>	<i>0.06</i>	<i>0.08</i>	<i>0.08</i>	<i>0.14</i>	<i>0.18</i>	<i>0.07</i>	<i>0.08</i>	<i>0.12</i>		

Note: Based on information from 43 teams. All specifications include a 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.

nonformal theories. 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 estimate in column (5) of panel C2 implies that profits increase by 100 euros (about one-third of a standard deviation) if the share of minorities is raised from 0.6 to 0.7. 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 the share of minorities in the segment above 60% of minorities.¹⁴

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. 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

¹⁴F-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 (not tabulated).

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.

4.2 Costs and benefits of cultural diversity

In this subsection we explore whether our data, collected through the questionnaires, indeed suggest mechanisms that possibly drive the effect of cultural diversity on team performance. If the variable that we employ as an indicator of a mechanism is unrelated to the cultural diversity of teams then we can conclude that this mechanism will not explain our findings. On the other hand, if we find a relationship between a certain mechanism and cultural diversity then this mechanism possibly contributes to the explanation of our results. However, for the mechanism to actually explain our findings, it also needs to have an association with team performance. And even if both associations are found, we still lack convincing evidence. For example, on the one hand, a positive relationship between learning and team performance may indicate that learning by the team improves performance. On the other hand, team performance may influence team members' individual assessment of how much learning took place. For this reason, the estimated relationship between learning measures and cultural diversity is open to multiple interpretations and certainly not indicative of a causal effect of team composition on learning. Nevertheless, we think it is useful to explore the occurrence of this kind of relationships that may possibly explain our findings. In what follows, we first consider the mechanisms that are associated with the costs of cultural diversity and subsequently the mechanisms that are related to its benefits.

For communication and coordination, we expect that moderately heterogeneous teams incur higher costs of communication and coordination than teams that are homogeneous or heterogeneous in terms of cultural background, in line with the narrative theories introduced in Section 1 (Lazear, 1999; Earley and Mosakowski, 2000; Richard et al., 2004). We measure coordination and communication costs in terms of a team's atmosphere and personal conflicts (see Earley and Mosakowski, 2000; Richard et al., 2004). Teams' atmosphere is administered by asking students to rate the atmosphere within their team on a

5-points scale in the posttreatment survey. Likewise, conflicts in the team are surveyed by asking students to what extent there was conflict or disagreement between team members about personal matters (that are not task-related). More personal conflicts and a worse team atmosphere are expected to coincide with worse communication and coordination due to, for example, subgrouping along cultural lines, distrust or stereotyping. In line with Alesina and La Ferrara (2005), we find that homogeneous and moderately heterogeneous teams tend to experience less conflicts than heterogeneous teams (not tabulated). However, the data also indicate that teams' atmosphere and cultural diversity are not significantly related. In sum, these results do not consistently suggest that the costs of communication and coordination might drive our main findings.

For the benefits, we expect a positive relationship between cultural diversity and business performance due to more complementarities and (mutual) learning in culturally diverse teams. We measure learning in terms of the development in three knowledge areas most relevant for successful entrepreneurship: business, entrepreneurship and leadership (see Karlan and Valdivia, 2011; Minniti and Bygrave, 2001). 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. 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, 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 cultural heterogeneous teams learn more than members of homogeneous and moderately heterogeneous teams. Additional regressions at the individual level, that we run separately for students of Dutch and non-Dutch cultural background, show that the learning benefits of more cultural diversity accrue to similar extents to both groups (not tabulated). Moreover, and probably as

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	<i>0.34</i>	<i>0.24</i>	<i>0.42</i>	<i>0.34</i>	<i>0.18</i>	<i>0.32</i>	<i>0.30</i>	<i>0.16</i>	<i>0.27</i>	
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>	<i>0.50</i>	<i>0.50</i>	<i>0.53</i>	<i>0.56</i>	<i>0.59</i>	<i>0.57</i>	<i>0.51</i>	<i>0.55</i>	<i>0.51</i>	
R^2	<i>0.38</i>	<i>0.27</i>	<i>0.47</i>	<i>0.39</i>	<i>0.25</i>	<i>0.37</i>	<i>0.38</i>	<i>0.24</i>	<i>0.33</i>	
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	<i>0.38</i>	<i>0.26</i>	<i>0.46</i>	<i>0.39</i>	<i>0.26</i>	<i>0.37</i>	<i>0.37</i>	<i>0.26</i>	<i>0.36</i>	
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	<i>0.38</i>	<i>0.28</i>	<i>0.49</i>	<i>0.41</i>	<i>0.27</i>	<i>0.41</i>	<i>0.41</i>	<i>0.28</i>	<i>0.37</i>	
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	<i>0.38</i>	<i>0.28</i>	<i>0.51</i>	<i>0.40</i>	<i>0.25</i>	<i>0.40</i>	<i>0.41</i>	<i>0.25</i>	<i>0.36</i>	

Note: Based on information from 43 teams. All specifications include controls for team size, field of study, trade management Asia¹ and team-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.

Table 8. 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)
<i>Minimum</i>	0.46	0.48	0.36	0.41	0.48	0.42	0.48	0.53	0.48	0.61	0.64	0.60
R^2	0.35	0.24	0.58	0.56	0.34	0.51	0.23	0.10	0.20	0.28	0.15	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 controls for team size, field of study, trade management, Asia, and team-average knowledge levels at baseline. OLS, Median and Robust refer to the estimation method. Bootstrapped standard errors in parentheses. ***/**/* denotes significance at the 1%/5%/10%-level.

an explanatory factor of the higher learning levels in more diverse teams, 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 (see Lazear, 1999). Table 8 indicates that complementarities between team members and the coefficients of variation of business, entrepreneurship and leadership knowledge at baseline tend to be larger in culturally diverse teams.¹⁵

All in all, based on these results we cannot reject the idea that cultural diversity benefits (mutual) learning and heterogeneous knowledge, possibly leading to better team performance. This finding is partly consistent with the theoretical ideas formulated in Section 1. Culturally diverse teams tend to have a more diverse pool of relevant knowledge and (possibly based on this) experience more learning and achieve better results. However, we do not find support for the idea that moderately heterogeneous teams incur higher costs of coordination and communication.

5 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 turned 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 found a somewhat different result for teams in which the majority of their members have different cultural backgrounds. In case of an 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

¹⁵In Table 8 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.

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 heterogeneous teams, something starts spreading like wildfire. In line with theoretical predictions (Lazear, 1999), our data suggest that 'this something' that 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., 2011b) and the increased participation of females in the labor market. The diversity of teams may be an interesting instrument to optimize organizational performance. One of the most salient and relevant dimensions of team heterogeneity is cultural background (Alesina and La Ferrara, 2005). Until today, however, studies analyzing the causal effect of cultural diversity on team performance in the longer run have been scarce. More 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 very 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' field experiment, where randomization took place both in the gender dimension (see Hoogenboom et al., 2013) and, in the current study, in the cultural background dimension. The field experiment was conducted in the context of a worldwide leading entrepreneurship education program in one of the largest colleges in Amsterdam with a truly multi-cultural population in which student teams found real companies. As outside researchers, we were allowed to exogenously vary the cultural composition of otherwise randomly composed teams. Since the program requires students to deliver annual reports, we could measure their performance in terms of sales, profits and profits per share. All in all, this is a

genuinely interesting opportunity to measure the longer term effect of cultural diversity on team performance in realistic though controlled circumstances in activities that are likely to be similar to the activities of the most influential teams in the business world, i.e., management and entrepreneurial teams.

Several limitations pertain to this study. There are discrepancies between the business 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 might, 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. Finally, our experimental design does not allow for a causal interpretation of mechanisms such as (mutual) learning that lead to higher performance of cultural heterogeneous teams.

Nevertheless, 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. 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.
- Angrist, J. and Lang, K. (2004). Does school integration generate peer effects? Evidence from Boston’s Metco program. *American Economic Review*, 94(5):1613–1634.
- Aslund, O., Edin, P., and Fredriksson, P. (2011). Peers, neighborhoods, and immigrant student achievement: Evidence from a placement policy. *American Economic Journal: Applied Economics*, 3(2):67–95.
- Bernstein, R. and Edwards, T. (2008). An older and more diverse nation by midcentury. Technical report, US Census Bureau.
- 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.
- Card, D., Mas, A., and Rothstein, J. (2008). Tipping and the dynamics of segregation. *Quarterly Journal of Economics*, 123(1):177–218.
- Card, D. and Rothstein, J. (2007). Racial segregation and the black-white test score gap. *Journal of Public Economics*, 91(11-12):2158–2184.
- 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.
- 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.
- 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. (2009). Cultural biases in economic exchange? *Quarterly Journal of Economics*, 124(3):1095–1131.
- 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.
- Hanushek, E., Kain, J., and Rivkin, S. (2009). New evidence about brown v. board of education: The complex effects of school racial composition on achievement. *Journal of Labor Economics*, 27(3):349–383.
- 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.
- 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.
- Hoxby, C. (2000). Peer effects in the classroom: Learning from gender and race variation. NBER Working Paper.

- 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*, forthcoming.
- 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.
- 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.
- Minniti, M. and Bygrave, W. (2001). A dynamic model of entrepreneurial learning. *Entrepreneurship Theory & Practice*, 25(3):5–16.
- Montalvo, J. and Reynal-Querol, M. (2005). Ethnic diversity and economic development. *Journal of Development Economics*, 76(2):293–323.
- Morgan, J. and Vardy, F. (2009). Diversity in the workplace. *American Economic Review*, 99(1):472–485.
- 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. (2011a). Immigration and innovation in European regions. IZA Discussion Paper.

- Ozgen, C., Nijkamp, P., and Poot, J. (2011b). 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.
- 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.
- 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 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.