Team Production in International Labor Markets: Experimental Evidence from the Field

Elizabeth Lyons[*]

July 1, 2016

Abstract

Co-workers are increasingly diverse in their nationality and skill sets. This paper studies the effect of diversity on how workers are organized using data from a field experiment conducted in an environment where diversity is pervasive. Findings show that team organization improves outcomes when workers are from the same country. The opposite is true when workers are nationally diverse. These results are more pronounced for teams of workers with specialized skills. Further investigation of the data suggests that diverse teams have difficulty communicating. I find no evidence that preferences for or expectations about working with someone from the same country is affecting team performance.

The globalization of labor markets is an important phenomenon. The number of immigrants employed in the U.S. grew from 11.7 million in 1994 to 19.3 million in 2003, and this pattern is not restricted to the U.S. (International Labor Organization, 2012a).[1] Multinational enterprises

[1]Brookings reports that the share of immigrants in the U.S. labor force grew from 4.9% in 1970 to 16.4% in 2010. The population of international migrants motivated by labor force opportunities is at least equal to the population of Brazil (International Labor Organization, 2012b).

are contributing to the international nature of labor markets as well.[2] In addition to increases in labor market diversity, team-based production in knowledge work is also becoming more common (Jones, 2009a; Wuchty et al., 2007), in part because of improvements in communication technologies (e.g. Dessein and Santos, 2006). Given the extent of labor market globalization, hiring a multinational labor force is becoming inevitable in many countries and industries. Combined with the growth in team-based knowledge production, organizational success will likely become increasingly dependent on the success of nationally diverse teams of workers.[3]

Because of the importance of communication for success in teams (e.g. Hinds and Mortensen, 2005), national differences in teams may be costly if they impede communication, for instance, by increasing conflict and decreasing knowledge sharing (Hamilton et al., 2012; Hjort, 2014). However, national differences also may have benefits, including increases in creativity, problem-solving abilities, and innovativeness (Haas, 2010; Lazear, 1999).[4] Although the existing literature recognizes cultural and ethnic heterogeneity in teams as an important determinant of their performance, evidence on the direction and extent of the effect of national diversity on organizational performance is sparse and ambiguous (e.g. Bertrand and Duflo, 2016; Horwitz and Horwitz, 2007; Reagans et al., 2004). By testing the causal effects of national diversity on team performance relative to nationally homogeneous teams and groups of independent workers, this paper addresses this gap. Moreover, while the existing empirical literature provides evidence on the performance difference between diverse and homogeneous teams of workers, it does not address the performance differences between organizing diverse workers into non-collaborative groups and organizing them into teams. Understanding how to organize conditional on diversity is particularly important as diverse work forces become harder to avoid, and increasingly expected of organizations.[5]

Existing theoretical and empirical literature on team work and diversity suggests potential gains from and costs from coordination that may be intensified when workers are more different from one another. In particular, team work introduces potential problem solving and learning benefits due to complementary knowledge and ideas (e.g. Cooper and Kagel, 2005; Lazear, 1999). Beneficial complementarities may be higher in diverse than homogeneous teams because, for instance, they have access to different information sources (Hamilton et al., 2012). However, team work also introduces coordination costs that increase as communication becomes more difficult (e.g. Dessein and Santos, 2006). Communication costs may be higher in diverse teams than in homogeneous teams because, for instance, workers with different background have different understandings or interpretations of similar statements and events even when they speak the same language (Lang, 1986). Given this evidence, it is possible that diverse teams perform better or worse than ho-

---

[2]The size of total FDI stock as a percentage of world GDP in 2000 was 20%. Moreover, MNEs were responsible for at least 3.4% of the world's employment by 2000 (Kim, 2006).

[3]Cross-country partnerships in MNEs are ubiquitous (e.g., Hays, 1974; Manev and Stevenson, 2001), and the management of cultural differences among workers from different countries remains a major challenge (e.g., Gong, 2003; Kirkman and Shapiro, 2005; Makela et al., 2012).

[4]For a more thorough review of research on organizational diversity, see Harrison and Klein (2007).

[5]Organizations' reputations are increasingly affected by how diverse their work forces are. Diversity may contribute to their ability to retain and attract workers and consumers. For instance, Fortune annually ranks companies on the percentage of minorities on their work forces (Fortune, 2015).

mogenous teams. Moreover, if communication costs from diversity are sufficiently high and gains from complementarities do not compensate for these costs, diverse teams may under-perform non-collaborative groups even on tasks that would otherwise benefit from team work.

In order to address some of the uncertainty about worker organization and, in particular, to identify the causal impact of national diversity on team outcomes, I designed and conducted a natural field experiment on the online market for contract labor, oDesk[6]. Workers on oDesk have selected into an international labor market and thus may be more willing to engage with a diverse set of employers and co-workers than workers in a more local labor market. My experiment tests the impact of organizing nationally diverse workers into teams by comparing diverse and homogeneous teams to sets of independent workers. This design permits me to identify the differential impact of national diversity on the value of teamwork.

The experiment I ran, described in detail in Section 2, is a two-by-two design in which I randomly assigned contractors into groups of two to complete a web programming task that requires a Javascript and a PHP programmer. Groups were randomly assigned to either work as a team or not and were either made up of workers from different countries or not. All workers in the experiment are able to communicate in English.[7] To reduce the potential for free riding problems within groups, I assigned each group member to a specific portion of the task and required that the task be completed in a setting where individual team member monitoring is possible.[8] Field participants were unaware that they were being studied and subject to conditions consistent with those that are typical in their lines of work.

The findings from the empirical analysis, reported in Section 4, show that organizing contractors into teams improves outcomes for contractors in nationally homogeneous teams but worsens outcomes for contractors in nationally diverse teams. Specifically, relative to the sample means, teamwork leads to a 30% increase in output and a more than 50% increase in productivity for nationally homogeneous teams and a 33% decrease in output and a 75% decrease in productivity in nationally diverse teams. Consistent with the idea that team members are more likely to rely on collaboration when they have different skill sets, teamwork has a more positive impact on outcomes for same-country teams when team members have skill differences, but the opposite is true for cross-country teams. In particular, nationally diverse teams with overlapping skill sets perform significantly better than those with specialized skill sets, however, they still perform worse than nationally homogeneous teams and independent workers. These results are consistent with teamwork being costlier for diverse teams than homogeneous teams and also suggest that overlapping skills may compensate for some of the communication costs associated with a lack of shared nationality. Further investigation of the data suggests that communication is more difficult for nationally diverse teams than it is for nationally homogeneous teams and that contractors in nationally diverse

---

[6]The platform has since been renamed "Upwork".

[7]Given that language diversity may impact teamwork through different mechanisms than diversity due to cross-country differences in social and other work-related values (e.g. Hofstede, 1983; Landes, 1999), I ensure a minimum level of English language ability among participants.

[8]With this task assignment design, it is relatively easy to identify if a given group member cheats.

teams are unaware of this relatively more difficult communication. I do not find any evidence that contractors prefer to work with a country-mate, or that they have ex-ante expectations that diverse teams perform worse than homogeneous ones.

These results provide robust evidence of the importance of communication costs in obstructing team coordination. They imply that when communication costs, such as those present between workers from different countries, cannot be managed, independent work may outperform team work even on tasks that would otherwise benefit from collaboration.

This paper proceeds as follows. Section 1 reviews related literature; section 2 describes the experiment design; section 3 describes the data collected from the experiment and my estimation strategy; section 4 presents the empirical results, and section 5 summarizes and concludes the paper.

1   Labor Market Diversity Literature

This paper contributes to a large and vibrant literature on diversity in the labor market and, in particular, how this diversity relates to organizational performance. A number of dimensions of diversity have been considered in this literature including gender (Adams and Ferreira, 2009), functional (Boone and Hendriks, 2009), and linguistic (Lazear, 1999).[9] Existing theories of the performance implications of diversity highlight the trade-offs between coordination costs and returns from access to a more disparate idea and knowledge sets across all types of worker diversity (e.g. Dessein and Santos, 2006). However, the extent of these costs and benefits, and their overall impact on performance vary depending on the source of diversity. Given that the experiment I present in this paper considers nationally diverse groups of workers hired for a task that requires two types of skills, in this section I focus on literature related most closely to national and skill diversity.

1.1   Performance Implications of National Diversity

Empirical literature on immigrant worker hiring provides some of the most relevant evidence on national diversity in organizations for this paper. In general, this literature finds that immigrants are less likely to be hired in their new countries than seemingly equally qualified natives. For instance, in an audit study performed in Canada, Oreopoulos (2011) finds that applicants with non-English names are significantly less likely to receive callbacks than those with English names. Employers surveyed for the study self-report that this is due to concerns about language abilities although the experimental evidence presented in the paper does not support this explanation. Using data from a national survey, Leslie and Lindley (2003) find some support for these employers'

---

[9]Refer to Bertrand and Duflo (2016) for a more thorough review of diversity in the labor market.

claims. In particular, they find that some of the employment and earnings differences between whites and non-whites in Britain can be explained by language disadvantages. Kaas and Manger (2012) also provide evidence of applicant discrimination from an audit study for the German labor market with additional evidence that differential callback rates for immigrants disappear when applications include reference letters on applicants' personalities. In a related study undertaken in Belgium, Baert et al. (2015) find that applicants with foreign names are more likely to receive lower callback rates for jobs with easier to fill vacancies. McGinnity and Lunn (2011) find similar evidence among immigrants and natives in Ireland, however, they find no evidence that this discrimination depends on occupation. In an analysis of the UK's General Household Surveys, Bell (1997) demonstrates increases in labor market performance among black immigrants as they accumulate more local experience. The author attributes this to an assimilation effect. Agrawal et al. (2013b) show similar patterns among employers hiring through online labor markets. In particular, developed country employers disproportionately hire developed country workers despite the large wage premiums they command. This hiring differential falls as workers accumulate more experience on the platform. These studies demonstrate that organizations prefer to hire native workers suggesting a preference for less national diversity. In addition, they provide some suggestive evidence that discrimination against foreign workers may be due to doubts about how well foreign hires' will communicate and integrate with existing work forces.

Despite the apparent preference among hiring managers for nationally homogeneous work forces, whether they outperform nationally diverse work forces is unclear. To my knowledge, empirical studies of the firm-level performance implications of nationally diverse work forces are limited to several studies of cross-subsidiary team performance in MNEs and of cross-country virtual teams.[10] Combined, these studies find that the performance of these teams depends on the way they are managed (e.g. DiStefano and Maznevski, 2000; Gong, 2003; Jarvenpaa et al., 1998). These studies do not analyze whether nationally homogeneous teams would perform differently than nationally diverse ones under the same circumstances. They do, however, point to the role of management in reducing costs and capturing benefits in these teams.[11]

While the literature of immigrant labor market performance and of cross-country team management provide important evidence on how firms deal with national diversity, they do not provide evidence on whether nationally diverse groups of workers perform better or worse than nationally homogeneous groups, or how national diversity affects the returns to team work. This paper will contribute to these literatures by testing these questions, in addition to providing suggestive evidence that manager concerns about communication difficulties among nationally diverse workers may be warranted.

---

[10]There have been several studies on how population diversity and immigration affect regional economic performance. These studies generally find positive impacts of population diversity and economic performance (e.g. Alesina et al., 2013; Ottaviano and Peri, 2006). As these papers highlight, the impact of national diversity on team performance does not capture the many other ways that national diversity may affect an economy.

[11]This is similar to evidence presented in Kochan et al. (2003) which shows that the performance implications of racial and gender diversity in firms depend on how they are managed. Given this evidence, the lack of managerial support in my experiment is an important boundary condition.

Beyond the literature on national diversity, a number of papers consider one or more element of diversity that may also vary between countries. Among the most relevant of these elements for my analysis is ethnic diversity.[12][13] Hjort (2014) experimentally tests the performance effects of ethnic diversity among teams at a flower production plant in Kenya. He finds worse performance outcomes for teams made up of workers from rival tribes than those made of workers from the same tribe, and provides evidence that this is due to taste-based discrimination between workers. Marx et al. (2015) also experimentally test the performance implications of ethnic diversity in teams in Kenya, and find that ethnic diversity between workers of the same rank has negative performance effects but that diversity between managers and subordinates has positive performance implications. Furthermore, the authors provide survey evidence that homogeneous co-workers made better production decisions potentially because they were more willing to collaborate. Both of these studies demonstrate that ethnic diversity can be harmful when workers have preferences for working with others from their own ethnicity. In the Netherlands, a setting with (arguably) less ethnic tension than Kenya, (Hoogendoorn and Praag, 2014) experimentally test the effect of ethnic diversity on the performance of start-ups among students. In contrast to Hjort (2014) and Marx et al. (2015), they find that ethnically diverse teams outperform ethnically homogeneous teams. The authors provide suggestive evidence that this is due to returns from increased knowledge diversity in diverse teams.

These papers provide useful evidence for thinking about the performance of nationally diverse teams. However, ethnic diversity does not always vary across countries (Alesina et al., 2013) and, in some cases, may vary more within a country than across countries.[14] As a result, ethnic diversity may have different effects on performance than national diversity. Moreover, the mechanisms through which they affect performance may differ. For example, labor market culture appears to differ more at the national level than the regional level (Hofstede, 1984). In contrast within-country differences in average educational attainment across ethnicities differs more in some instances than average educational attainment across countries (e.g. Barro and Lee, 2013; Kao and Thompson, 2003). It is plausible, therefore, that ethnic diversity may matter more or less than national diversity for the performance of teams and that the direction and causes of their impact on performance differ.

---

[12]In my setting, ethnicities differ both within and across countries. I find no evidence, however, that within country ethnic diversity affects the performance of teams in my study.

[13]Other worker characteristics that may vary at the national level include race, and colonial ties. Given that my experiment is restricted to workers from India, Pakistan, and Bangladesh, these are unlikely to vary much in my setting.

[14]For instance, Alesina et al. (2011) demonstrate that national borders have frequently been drawn in such a way that ethnic groups are divided into more than one country. In my empirical setting, some ethnic groups are divided by national borders. For instance, a number of enclaves, defined as "...a portion of one state completely surrounded by the territory of another state", exist on both sides of the India-Bangladesh border (Jones, 2009b).

1.1.1 Explanations for Performance Effects of National Diversity

As the studies on ethnic diversity summarized above point out, diversity may affect performance for a number of reasons. For instance, nationally diverse work forces may perform worse on average than nationally homogeneous ones because they communicate less effectively as predicted by Lang (1986) and Lazear (1999). Communication costs may be higher in nationally diverse teams for instance because of language differences (Lazear, 1999), lack of trust (De Dreu and Weingart, 2003; Guiso et al., 2009), differences in social and work norms which may lead to different in interpretations and understandings of concepts and goals (Hofstede, 1984), and differences in cognitive frameworks which may also lead to differences in how the same information is processed (Dahlin et al., 2005).[15] All of these possibilities, as well as the possibility of other mechanisms, may be driving communication costs in nationally diverse teams.[16]

An alternative explanation to that of higher communication costs for relatively poor performance of nationally diverse teams is animosity for workers from other countries (Alesina and La Ferrara, 2005). In particular, as demonstrated by (Hjort, 2014; Marx et al., 2015), individuals may get positive utility from working with others who are like themselves, and negative utility from working with those unlike themselves. This is conceptually distinct from communication costs because workers prefer to avoid working with those who are different from themselves based on preferences, as opposed to being unable to effectively work with them due to communication barriers.

An additional possible explanation for worse performance among nationally diverse teams is that workers expect that they will not be able to work well with someone from another country even if it this belief is not ex-ante accurate, and even if workers have no animosity towards those from other countries. This explanation is consistent with evidence that stereotype threats can decrease performance (e.g. Schmader et al., 2008). Alternatively, this could occur because of the common belief that ones own type has superior knowledge and abilities than those who are different (Brewer, 1979) which could lead workers to not listen to or take advice from teammates from other countries.

As in Hoogendoorn and Praag (2014), national diversity can also increase team performance by giving teams access to a more diverse knowledge or idea set. In my setting, as discussed in sections 2.3 and 4, nationally diverse teams may well have larger knowledge sets than nationally homogeneous teams do. Moreover, given that tasks requiring creativity and problem solving benefit more from knowledge diversity (Audretsch and Feldman, 2004), the task they are assigned (described in detail in section 2.2) likely benefits from more knowledge and ideas. Due to measurement limitations, I do not test directly for whether nationally diverse teams approach completing the

---

[15]In the experiment presented in this paper, workers are all able to communicate in English, and anecdotal evidence from chat room screen shots suggests they do communicate with each other in English. Language differences, therefore, may not be a relevant explanation for performance differences in this context. However, even when speaking the same language, linguistic differences between nationally diverse workers may persist (Kachru, 1992).

[16]Testing for which are present in my empirical setting is beyond the scope of this paper, but an important area for future research.

assigned task in more creative ways than nationally homogeneous teams, or whether nationally diverse teams have less overlap between team members in how they approach the same problem.[17]

Causally establishing the reason for the performance differences between nationally diverse and homogeneous teams is beyond the scope of this paper. However, I test for and suggestively rule out some of the explanations outlined above. In particular, I find no evidence that workers in diverse teams dislike their teammates with more frequency than those in homogeneous teams nor do I find any evidence that workers in diverse teams have difference ex-ante expectations about their performance. Moreover, anecdotal evidence from chat room screen shots demonstrates that all team types appear to communicate in English. I do find evidence indicative of communication difficulties in diverse teams. As a result this paper contributes to our understanding of the mechanisms that drive the performance effects of diversity in teams by demonstrating that even in the absence of taste-based discrimination, language differences, and poor expectations about team performance, national diversity can lead to large enough difficulties that gains from teamwork are eliminated.

1.2   Performance Implications of Skill Diversity

Related to knowledge diversity is skill diversity. Consistent with Ingram and Neumann (2006), skills in this context are classifications for specific abilities needed for the completion a given task. It is thus easier to characterize and measure skill diversity than knowledge diversity which may occur even between workers with the same skill set. Like national diversity, skill diversity may introduce additional communication costs and complementarities in teams. Lazear (1999) shows that teams with less overlap between their skill sets lead to more gains from complementarities when the skills are relevant for the job. However, skill diversity may also lead to higher communication costs because workers are less able to evaluate or understand ideas or input on a skill they are not familiar with (Ferreira and Sah, 2012; Galbraith, 1977), or because workers are less able to know what types of knowledge to ask for from their teammates (Rulke and Galaskiewicz, 2000).

Empirical evidence on the impact of skill diversity on organizational performance is relatively thin. For instance, Hamilton et al. (2003) and Hamilton et al. (2012) find that increases in skill diversity withing teams leads to increases team performance. However, the authors define skill diversity as differences in average productivities or ability levels as opposed to occupation categories. Mas and Moretti (2009) also find that combining low and high productivity workers has a positive impact on performance relative to having workers with similar productivities work together. The mechanisms through which productivity diversity affect performance are likely different than those through which skill diversity as defined in this paper affect performance.

Importantly, the costs and benefits of skill diversity may enhance or dampen those introduced by national diversity. For instance, at the regional level, Alesina et al. (2013) find that skill diversity

---

[17]Given that nationally diverse teams perform significantly worse than all other organizational forms in my study, if there are benefits to diverse knowledge sets in my setting, they do not outweigh the costs of diversity.

among immigrants increases the positive economic impact from immigrant birth country diversity. In contrast, Dessein and Santos (2006) argue that one way of reducing communication costs within teams is to decrease team member specialization so that there is more overlap in skill sets across team members. Combined, these papers suggest that skill specialization in teams may increase the benefits to complementarities between team members, but it may also exacerbate communication problems.[18] By testing how skill diversity affects national diversity in teams this paper provides, to my knowledge, the first empirical evidence of how the effects of skill specialization vary across teams with different communication costs and complementarities.

2 Experimental Design

Identifying the causal impact of national diversity on the value of teamwork is problematic because teams are selected and not randomly assigned. This endogenous team formation makes it difficult to disentangle both the value of teamwork as compared to independent work and the differential impact of teamwork for nationally diverse teams. For example, it is plausible that employers and workers choose which jobs to complete as a team and that they choose who should work together in teams according to their performance expectations. To address these concerns, I test the effects of national diversity on work team success by conducting a field experiment on the world's largest online contract labor platform, oDesk.

To perform the study, I set up an employer account on the oDesk website and posted jobs to attract contractor bids. I hired contractors who met the requirements for participation in the experiment and randomly put them into groups of two to complete a programming task. I define teams in this study as pairs of contractors working towards a common goal. Below I describe the setting of the experiment, the task, team compositions, and the hiring and work process used for the experiment.

2.1 Research Setting

I generate data for this analysis through a natural field experiment conducted using oDesk, an international online market for contract labor. At the time of the experiment, oDesk was the largest contract labor platform in the world in terms of earnings, and the fastest-growing. By offering more flexible conditions and competition than traditional labor markets as well as the potential for more efficient matching and investments in human capital (Autor, 2001; Grossman and Helpman, 2002; Horton, 2010; Oster and Millett, 2010), online labor markets such as oDesk may be able to absorb a meaningful portion of jobs that are now completed offline and locally.

oDesk (now "Upwork") operates as a mediator and matches employers with contract laborers who

---

[18]I empirically test whether skill diversity affects nationally diverse and homogeneous teams differently in section 4.1.2.

work on jobs remotely. To hire on oDesk, employers create an account on the site and post jobs for which they are looking to hire. Job postings typically include a brief description of the job, and the type of contract being offered (fixed price or hourly wage payment). Contractors interested in being hired for jobs can apply to them by submitting a bid to indicate the amount they are willing to accept to work on the job and a cover letter describing why they are suited for the job. Contractors also advertise their skills and abilities to employers through their profile pages. Contractors set up their profile pages when they join the site. These pages can include information on their education, prior work histories,[19] and scores from tests administered by oDesk. Profile pages also provide information on where contractors live. Contractors can be actively employed by multiple jobs at once.[20]

There are several reasons I chose to run this experiment on oDesk. First, oDesk provides a team room application which allows employers to put hired contractors into teams so that they can work together online. In particular, contractors working in the same team room can monitor what the others are doing and chat with each other through instant messaging. Contractors can only be logged into one team room at a time. Employers on oDesk frequently require that multiple hires work in the same team room.[21] Second, the team room facilitates contractor monitoring by employers through frequent screen shots, memos, work diaries, and activity meters. As a result, the majority of employers on oDesk require that their hires complete their tasks while logged into their team room. Another important feature of oDesk for this research is the option to interview. oDesk encourages employers to interview contractors they are interested in hiring before making any offers. These interviews can take place through messages exchanged on the site. Lastly, oDesk is an international labor market which contractor participants have selected into knowing they would almost surely be working with employers and contractors from other countries. Therefore, we might expect to see less distaste for working with a nationally diverse organization than in more local labor markets.

2.2   The Task

To test my research question, I needed contractors to be assigned a task for which collaboration is not unnatural. At the same time, because I am comparing team work to independent work, contractors working on their own should be capable of completing at least some portion of the

---

[19]Platform-specific work histories include a list of all jobs performed on oDesk with a brief description of each job, the hiring employer, and feedback and a rating if the employer has chosen to provide them.

[20]For further information on platform features, refer to Agrawal et al. (2013a).

[21]During the eight years prior to 2014 on oDesk, teams of two to three contractors have grown from 10 per month to more than 5,000 per month, and in mid-2012, there were almost 20 teams of at least 64 on the site (Ipeirotis, 2012). Work teams can enhance productivity in these markets, for instance by allowing workers to further specialize and share ideas (e.g., Hamilton et al., 2003; Jones, 2011). As employers increasingly outsource knowledge-intensive jobs through digital platforms, the prevalence of international work teams will likely continue to grow.

task.[22] In addition, because I am interested in testing the coordination and communication costs associated with national diversity in teams, free-riding on the assigned task should be difficult. The task also needed to be consistent with the types of tasks posted on oDesk so that contractors did not become suspicious about the purpose of the job.[23] Moreover, an objective evaluation of contractor performance was necessary. Finally, it must have been possible to complete the task remotely. Given these constraints, I developed a web development task that required both back-end (PHP) and front-end programming (Javascript) to assign contractors hired for the experiment. Hired contractors were asked to add a list of features in these languages to existing code. This is similar to a web development task for which a group of workers are responsible for both the design and functionality of a web page.

To reinforce the collaboration aspect of this job and to reduce free-riding in teams (Holmstrom, 1982), I hired one contractor in each pairing to complete the Javascript portion of the task and another to complete the PHP portion.[24] I assigned each pair three features to add to the code, one that required only Javascript programming, one that required only PHP programming, and one that required both.[25] Including both individual tasks and a team task accomplishes two objectives. First, it allows for both independent and for team work so that neither of the organizational form treatments lead to contractor frustration.[26] Second, it allows me to disentangle difficulties associated with collaboration from difficulties associated with the job in general. All hired contractors received the same list of features to be added, and I instructed them to work on the features that corresponded to the language for which they were hired. I display the instructions provided to contractors in Appendix B. Nothing in the instructions or the task design explicitly prevents contractors from working on all three features of the task, including the parts that are not in the language they were hired for. The reasons a single contractor may not work on both languages are first, that some contractors are only familiar in coding one of the two languages, and second, that they have been asked to focus on one of the two languages. Importantly, the feature that requires both programming languages to be completed requires that the codes from each language work together. A contractor who only knows one of the two languages would not be able to complete this feature on his or her own. Therefore, contractors in the independent work treatment would not be able to complete all three features unless they are familiar with both PHP and Javascript.

---

[22]During a pilot test of the experiment, contractors were giving a job that only had collaborative tasks (i.e. tasks that required both JavaScript and PHP to be completed). Those in the independent work treatment complained quite a bit about the task being unrealistic for them to work on without being able to communicate with their pairmates.

[23]This was to minimize bias caused by contractors performing differently than they would in real labor market situations. For instance, if contractors believed they were being observed for the purposes of a research study, they may have tried harder than they normally would have to complete the task or to behave in socially desirable ways (e.g., Levitt and List, 2007).

[24]The team rooms further reduce the potential for free-riding because contractors are aware that their individual performance can be monitoring by employers through the team room application.

[25]An aspect of the task that I had not taken into account prior to running the experiment is the script all three features were to be added to was PHP-based. As a result, it turned out that some PHP knowledge was also important for adding the Javascript feature to the script.

[26]For instance, if the job only required a joint task, the contractors in the independent work treatment may be confused as to why the job was designed the way it was. Similarly, if the job only included independent tasks, contractors in the team treatment may wonder why they had been asked to work in a team.

In order for the contractors in the teamwork condition to be able to work with each other, I gave them a total of eight hours to complete the task; they had to work these hours within the same period as the other contractors in their groups.[27]

2.3   Contractor Pair Composition

I randomly assigned each hired contractor to one of four types of groups of two, and all groups were made up of one contractor hired to work on the Javascript code and one to work on the PHP code. As Figure 1 demonstrates, the experiment had two treatments: pairs were either from different countries or not and were either permitted to work in teams or not. Pairs in the team treatment worked in the same team room as each other, and those in the independent work treatment worked in separate team rooms and were unaware of who their pair mate was so that team work was not an option. The purpose of this experimental design is to compare the benefits of teamwork in homogeneous teams with the benefits of teamwork in nationally diverse teams. This allows me to identify how organizing workers into teams compares with organizing them as independent workers, and how this is affected by national diversity.

Figure 1: Experiment Design

		Nationality	
		Same	**Different**
Team Room — Separate		Independent work by contractors of same nationality (N=40 pairs)	Independent work by contractors of different nationalities (N=41 pairs)
Team Room — Shared		Team work by contractors of same nationality (N=40 pairs)	Team work by contractors of different nationalities (N=41 pairs)

Notes: This matrix reports the four experimental conditions.

To control for country-specific differences in contractor quality, I limited the number of countries from which I hired. Based on information about the set of applicants my job postings attracted during the pilot phase of my study, I included English speaking contractors from India, Pakistan,

---

[27]Contractors had between midnight UTC one day until midnight EST the next. During a pilot of this project, it became clear that restricting contractors to work within a single 8 hour period was unrealistic. Giving contractors about a day to finish the job is consistent with many of the jobs on oDesk and still gave contractors the chance to work with their teammate if they wanted to, particularly given that there was very little variation in time zones across contractors in my sample. oDesk contracts automatically begin at midnight UTC on the start day specified in the contract, and because contractors were aware that the employer time zone was EST (this information was posted on my employer profile), to minimize confusion the deadline was midnight EST.

and Bangladesh in the experiment.[28] Contractors from these countries frequently applied to my job postings and bid an amount that was within the hourly wage criteria. Moreover, workers in these countries have very similar time zones making it easier to allow for remote team work among them. While these three countries are similar in many respects - until 1947, all three were part of India - there are also important differences between them that may result in, for example, different knowledge sets and communication styles.[29] These differences are discussed in more detail in section 4.

2.4   Hiring Process

Using an employer account on oDesk, I attracted contractor bids by posting two job types. One type of job posting asked for bids from contractors able to code in PHP and the other for bids from contractors able to code in Javascript. For each pair in my sample, one PHP job and one Javascript job was advertised. Contractors therefore selected into the programming language they preferred. To minimize attrition among hired contractors, the job postings described the type of work the task required, the date the work needed to be done on, and the number of hours contractors had to complete the task. The job postings also specified that an hourly wage contract was being offered[30] and the maximum bid that would be accepted ($4.00 USD).[31] To avoid bias due to contractor selection into job applications, the job postings did not mention that the work would be completed in teams nor did they mention anything about country-specific requirements. I provide screenshots of the Javascript and PHP job postings used for the experiment and further details about the postings in Appendix B. Applicants chose which job to apply for, either Javascript or PHP, however, some applicants are knowledgeable in both languages. It therefore seems likely that applicants applied to the job that required the language that they are most comfortable coding in.

The first applicant to each job posting who bid at most $4.00 and who was from one of the countries

---

[28]In the independent work, no national diversity treatment, 34 workers are from Bangladesh, 38 are from India, and 8 are from Pakistan. In the independent work, national diversity treatment, 22 are from Bangladesh, 34 are from India, and 26 are from Pakistan. In the team work, no national diversity treatment, 26 workers are from Bangladesh, 52 are from India, and 2 are from Pakistan. In the team work, national diversity treatment, 30 workers are from Bangladesh, 28 are from India, and 24 are from Pakistan.

[29]The education system in each country is overseen by their respective governments, and the curricula in the countries differ (National Curriculum & Textbook Board, Bangladesh, 2013; Ministry of Human Resource Development Government of India, 2013; Islamabad, 2013). Moreover, Hofstede (2013)'s measures of labor market culture suggests noticeable differences across the countries, particular in terms of their uncertainty avoidance, individualism, and power distance measures.

[30]This is to ensure contractors perform their work while logged into the team room. Hourly wage contracts on oDesk guarantee payment to contractors for all hours worked as long as the hours are performed in the team room. Fixed price contracts do not guarantee payment so there is less incentive to perform work while logged into the team room.

[31]oDesk charges 10% of each transaction on the site so a $4.00 bid from a contractor's perspective shows up as a $4.44 bid from an employer's perspective.

included in my experiment received a list of interview questions. The purpose of the questions was to get an idea of the level of Javascript and PHP knowledge each contractor had, their English language abilities, and the countries they have lived in. The interview text sent to Javascript and PHP job interviewees is in Appendix B. I gave each interviewee two hours to reply to the interview questions before I interviewed the next applicant who met the bid and country requirements. I hired the first interviewee to reply with the exception of interviewees who did not provide pertinent answers due to their lack of English language ability.[32] Some contractors applied for both jobs, but were hired for the task in which they met the hiring requirements.

2.5 Job Completion

Once hired, contractors were sent the appropriate instructions and the file with the code to be edited. There were four versions of the job instructions to reflect whether the contractor was hired for the Javascript or the PHP portion of the job and whether pairs were permitted to work as a team or not. Instructions do not vary across the national diversity treatment. I provide the four different versions of the instructions in Appendix B.

In addition to describing the features to be added to the code, the instructions noted the timeline for the task and the country that contractors' co-worker lived in (i.e. the country of residence of the other contractor in the pairing). I included contractors' pair-mate's country of residence in the task instructions because it is not possible to restrict access to this information in the team work treatment, so I controlled for this knowledge across treatment groups to ensure that it was not driving my findings. The instructions asked contractors to note what they were working on while in the team room and to include information on what they had completed when they turned in their work.

I asked contractors who had not turned in their work by the deadline to submit any work they had completed. Once contractors had turned in their work or after the deadline had passed, I asked them whether they would be willing to answer a few questions about the job for a $0.50 bonus. The purpose of the survey is to obtain a measure of how well contractors think they did on the job and their perceptions on teamwork. Once contractors had completed the survey or indicated that they did not want to complete it, their contracts ended. I provided feedback and ratings based on contractors' output to all paid contractors.[33]

---

[32]For instance, some interviewees replied with text unrelated to the questions and others provided answers to the questions that were not comprehensible.

[33]Because ratings are weighted by the amount a job pays on oDesk, contractors who earn $0 on a job cannot be rated by employers.

14

3 Data and Estimation Methodology

3.1 Overview of the Data

I collected data between January 2013 and June 2013. In total, I hired 324 contractors, who made up 162 pairs.[34] There are 80 contractors and 40 pairs in the nationally homogeneous teamwork and independent work treatment groups,[35] and 82 contractors and 41 pairs in the nationally diverse teamwork and independent work treatment groups.[36]

For each hired contractor, I collected information provided on their profile pages including their work histories on and off oDesk, education, oDesk test scores, and advertised wages. I also recorded the number of examples of Javascript and PHP work they provided in their interview responses as well as the amounts they bid on my jobs. For each job posting, I recorded the number of applicants and the amount of money I had spent as an employer on the site up to that point.

After contractors completed their work or once the deadline had passed, I recorded information on the number of features they added to the code, both with and without error, the amount of time they spent working on the task, the difference between the number of features they reported having added and the number they actually had added, which features they attempted to add, and the total amount they were paid for the hours worked. In addition to these performance outcomes, I collected information on contractor effort. Specifically, I recorded whether or not contractors attempted to implement each of the three features in the task, and how their effort levels (measured by oDesk using contractor mouse and keyboard activity) evolved over the course of the job. Table 1 describes these outcome variables, the key independent variables and the contractor and job descriptive variables.

I present summary statistics of hired contractor-pair characteristics in Table 2. I measure each variable as the average between the two contractors in each pair. It is worth noting that hired contractors were, on average, relatively inexperienced on oDesk given that about 60% of them had never received a rating on the platform prior to being hired (though many of those without a rating had been hired once or twice on the platform),[37] that the average bid on the job was lower than the average wage contractors indicated they would work for on their profiles, that more than 85%

---

[34]In 9 of the pairs, at least one of the contractors did not contact me again after receiving the instructions. I have coded these contractors as not having completed any of the work. However, given that there might be reasons for disappearing other than their ability to do the work, I check whether the results are robust to excluding these pairs, and to assigning them the sample average performance (see Appendix A).

[35]As discussed in Section 2, contractors assigned to pairs in the independent work treatment work in separate team rooms from each other. These contractors were not aware of who their pair-mates were, and were not able to communicate with them.

[36]The difference in the number of observations across groups is a result of some groups taking longer to meet the observations requirement than others.

[37]The inexperience of contractors in this study is liklely due to the relatively low wage I offered for a job that requires some specialized skills (Agrawal et al., 2013b).

of hired contractors were male, and that the average hired contractor had a Bachelor's degree. I present summary statistics of pair outcomes in Table 3. Here it is worth noting that the average pair did not work the full 16 hours given (eight hours per contractor). This could be because contractors are incentivized by both the financial returns to the job and the potential reputational returns. In particular, contractors on oDesk receive feedback and a ratings score once they have completed a job, and they may believe that working fewer hours will result in better feedback. Alternatively, it could be because contractors on oDesk can only work on one hourly wage job at a time, and they may have decided that their time was more valuable spent on another task. It is also

Table 1: Variable Definitions

Variable	Description
Dependent Variables:	
Joint Output	Number of non-overlapping features added to code by the Pair
Joint Productivity	Number of non-overlapping features added to code by the Pair divided by number of hours worked
Individual Output	Number of features added to code by the Contractor
Individual Productivity	Number of features added to code by the Contractor divided by number of hours worked
Difference between Reported and Completed	Absolute difference between what contractors reported as completed and what was actually completed by the Pair
Attempt Javascript Feature	Equal to one if at least one Contractor in the Pair attempts to add the feature that requires Javascript to the code, zero otherwise
Attempt PHP Feature	Equal to one if at least one Contractor in the Pair attempts to add the feature that requires PHP to the code, zero otherwise
Attempt Combined Feature	Equal to one if at least one Contractor in the Pair attempts to add the feature that requires both Javascript and PHP to the code, zero otherwise
Reported Doing More Work than Teammate	Contractor response to survey question "Did you or your teammate do more work on this project?" Response range: 1 (My teammate did almost all of the work) to 7 (I did almost all of the work)
Willingness to Work with Teammate Again	Contractor response to survey question "Would you prefer working with your teammate from this task or someone else on another project?" Response range: 1 (I would prefer working with someone else) to 7 (I would prefer working with my teammate from this task)
Independent Variables:	
National Diversity	Equal to one if Contractors in the Pair are from different countries, zero otherwise
Team Work	Equal to one if Contractors in the Pair are permitted to work as a team, zero otherwise
Skill Differences	Equal to one if at least one Contractor in the Pair has knowledge of a relevant skill (i.e., Javascript or PHP) that their co-worker does not have, zero otherwise
Indicator for Pairs with Different Javascript Knowledge	Equal to one if only one Contractor in the pair has knowledge of Javascript, zero otherwise
Indicator for Pairs with Different PHP Knowledge	Equal to one if only one Contractor in the pair has knowledge of PHP, zero otherwise
Outcome Descriptives:	
Total Hours Worked	Total hours worked by Pair on job
Amount Paid for Hours Worked	Total amount paid to Pair for job
Contractor and Job Descriptives:	
Number of Job Applications	Number of Contractors who applied to a job posting
oDesk Rating	Job size weighted feedback score out of five provided by prior oDesk Employers
No oDesk Rating	Equal to one if Contractor has no oDesk rating at time of hiring, zero otherwise
oDesk Experience	Number of contracts Contractors have been hired to complete on oDesk at time of hiring

Profile Picture	Equal to one if Contractor has a profile picture, zero otherwise
Education	Measurement of highest level of education. Equal to one for College Diploma, two for Bachelor's Degree, three for Master's Degree, four for Doctorate, zero otherwise
Number of Non-oDesk Jobs Reported	Number of non-oDesk jobs reported on profile
Average oDesk Test Score	Average score out of five on oDesk tests taken
Number of oDesk Tests	Number of tests taken on oDesk
Wage Bid	Amount Contractor bid on job posting
Advertised Wage on profile	Wage Contractors post on profile pages as amount they are willing to work for
Female	Equal to one if Contractor is female, zero otherwise
Employment Agency Member	Equal to one if Contractor belongs to an oDesk employment agency, zero otherwise
Number of portfolio Items	Number of items in Contractor's profile page portfolio

Table 2: Pair Characteristics Summary Statistics

Pair Averages	Mean	(Std. Dev.)
Number of Job Posting Applications	8.218	(3.18)
oDesk Rating Prior to Hire	4.577	(0.66)
No Rating Prior to Hire	0.611	(0.37)
Number of oDesk Contracts Prior to Hire	4.213	(7.466)
Indicator for having a Profile Picture	0.843	(0.27)
Level of Education	1.846	(0.688)
Number of Offline Jobs Listed on Profile	1.191	(0.643)
Average Score on oDesk Tests	3.428	(0.422)
Number of oDesk Tests Taken	2.509	(2.283)
Wage Bid on the Job	3.759	(0.607)
Wage Posted on Profile	6.888	(3.543)
Indicator for Female Contractor	0.145	(0.253)
Indicator for Agency Membership	0.253	(0.312)
Number of Items in Portfolio	3.87	(4.806)
Indicator for Pairs with Different Javascript Knowledge	0.329	(0.471)
Indicator for Pairs with Different PHP Knowledge	0.360	(0.481)
N		162

worth noting that the majority of pairs attempted to implement the Javascipt and PHP features but that only 33% of pairs attempted the combined feature. This suggests contractors perceived this feature as more complex, perhaps because it requires both programming languages to be implemented and, therefore, relies more heavily on teamwork.[38]

To verify the randomness of the treatments across pairs, Table 4 presents hired contractor-pair characteristics across the four treatment groups. A multivariate analysis of variance test of equality of the four group means indicates that there are no statistically significant differences across the four groups. Table A1 compares hired contractor-pair characteristics across the teamwork treatment (in Panel A) and across the national diversity treatment (in Panel B). Given that I randomly assigned pairs to treatment groups, characteristics should be statistically the same across the groups and, with one exception, that is the case here. The exception is the number of oDesk tests taken across the national diversity treatment. This measure is statistically different at the 10% level between the two groups. However, the medians between the two groups are identical.

Figures 2 and 3 compare pair outcomes across treatment groups. Figure 2 compares the number of features implemented by pairs that worked independently and those that worked in teams. The figure also separates the teamwork treatment by the national diversity treatment in order to compare the treatments of interest. These averages show that teamwork significantly improves outcomes when team members are from the same country and significantly reduces output when team mem-

---

[38]The three features in the task were designed to be equally difficult for programmers who are experts in both PHP and Javascript.

Table 3: Pair Outcome Summary Statistics

Pair Averages	Mean	(Std. Dev.)
Team Total Amount Paid	35.555	(20.336)
Team Total Number of Hours Worked	9.35	(5.123)
Number of Features Implemented	1.068	(0.773)
Number of Features Implemented Divided by Hours Worked	0.131	(0.164)
Difference between Actual and Reported Features Added	0.327	(0.555)
Attempt Javascript Feature	0.920	(0.273)
Attempt PHP Feature	0.821	(0.385)
Attempt Combined Feature	0.327	(0.471)
No Attempt	0.056	(0.230)
Increasing Effort	0.364	(0.483)
Teammates Worked on the Same Feature	0.179	(0.385)
N		162

bers are from different countries. These effects are economically significant as well. In particular, teamwork improves output for nationally homogeneous pairs by 38% (p-value 0.04) relative to nationally homogeneous independent contractor pairs and reduces output for nationally diverse pairs by 30% (p-value 0.03) relative to nationally diverse independent contractor pairs.

Figure 2: Output by Treatment
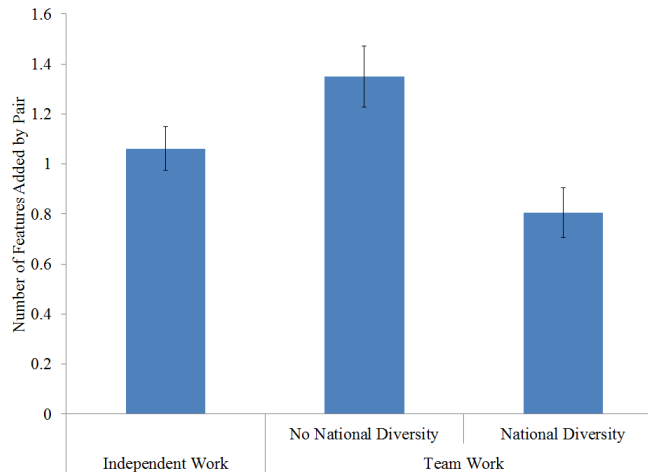
Figure 3 compares the number of features implemented divided by the total number of hours worked by pairs and shows the same pattern, that teamwork increases productivity when teammates are from the same country and decreases it when they are from different countries. Again, these effects are economically significant. Teamwork increases the productivity of nationally homogeneous pairs by almost 90% (p-value 0.04) relative to nationally homogeneous independent contractor pairs and decreases the productivity of nationally diverse pairs by more than 50% (p-

Table 4: Pair Characteristic Summary Statistics Across Treatment Groups

Variable	No National Diversity		National Diversity		p-value[+]
	No Team Team Work	Team Work	No Team Work	Team Work	
Number of Job Posting	7.900	8.400	8.805	7.768	0.435
Applications	(2.565)	(3.086)	(3.239)	(3.660)	
oDesk Rating Prior to	4.423	4.657	4.631	4.589	0.645
Hire	(0.919)	(0.520)	(0.448)	(0.650)	
No Rating Prior to	0.625	0.613	0.646	0.561	0.760
Hire	(0.386)	(0.381)	(0.373)	(0.337)	
Number of oDesk Contracts	4.775	2.475	5.402	4.171	0.331
Prior to Hire	(8.068)	(3.699)	(10.387)	(5.672)	
Indicator for having a	0.850	0.875	0.780	0.866	0.382
Profile Picture	(0.280)	(0.245)	(0.315)	(0.223)	
Level of Education	1.925	1.825	1.817	1.817	0.873
	(0.622)	(0.784)	(0.752)	(0.575)	
Number of Offline Jobs	1.200	1.200	1.122	1.244	0.860
Listed on Profile	(0.810)	(0.513)	(0.519)	(0.686)	
Average Score on oDesk Tests	3.500	3.427	3.340	3.452	0.449
	(0.379)	(0.536)	(0.348)	(0.405)	
Number of oDesk Tests Taken	2.238	2.113	2.817	2.854	0.329
	(1.862)	(1.847)	(2.359)	(2.829)	
Wage Bid on the Job	3.734	3.842	3.724	3.738	0.803
	(0.685)	(0.513)	(0.578)	(0.636)	
Wage Posted on Profile	7.460	5.762	7.385	6.933	0.116
	(3.511)	(1.877)	(4.439)	(3.594)	
Indicator for Female Contractor	0.100	0.163	0.159	0.158	0.644
	(0.257)	(0.261)	(0.259)	(0.234)	
Indicator for Agency Membership	0.200	0.263	0.256	0.293	0.605
	(0.314)	(0.318)	(0.297)	(0.314)	
Number of Items in Portfolio	4.950	3.588	2.914	4.049	0.282
	(6.486)	(3.506)	(2.737)	(5.398)	
Indicator for Pairs with	0.293	0.350	0.366	0.310	0.888
Different Javascript Knowledge	(0.072)	(0.076)	(0.076)	(0.072)	
Indicator for Pairs with	0.341	0.350	0.415	0.333	0.867
Different PHP Knowledge	(0.075)	(0.076)	(0.078)	(0.074)	
Number of Observations	40	40	41	41	

Notes: Standard deviations are in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%. [+] Test for equality of four group means using multivariate analysis of variance

value 0.01) relative to nationally diverse independent contractor pairs.
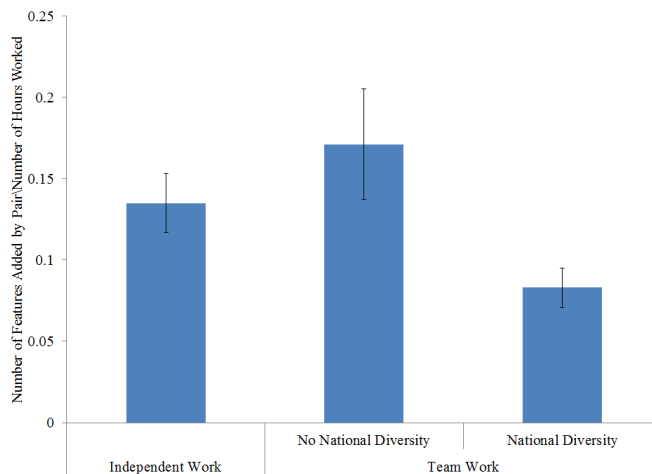
Figure 3: Productivity by Treatment

These comparisons suggest that when a team is nationally homogeneous, coordination is easy enough that teamwork outperforms independent work and that coordination when teammates are from different countries is sufficiently difficult that independent work leads to higher output and productivity. Moreover, the findings suggest that complementaries are not high enough to compensate for the costs of diversity in this setting.

I collected the data from a randomized experiment; therefore the simple correlations reported here can be interpreted as causal relationships. However, to further allay concerns about omitted variables, and to potentially increase precision, in the next section I estimate the effect of cross-country teamwork on performance in a multivariate regression framework.

3.2 Estimation Strategy

The regression estimates derive from versions of the following linear model:

$$Y_i = \alpha + \beta_1 TeamWork_i + \beta_2 NationallyDiverse_i * TeamWork_i +$$
$$\theta X_i + \delta CountryPair_i + \psi Week_i + \varepsilon_i \quad (1)$$

where Y_i is a measure of pair i's success, measured both as joint output (the total number of features implemented by the pair) and joint productivity (the number of features added divided by the total number of hours worked). $NationallyDiverse_i$ is an indicator for whether or not pair i is made up of contractors from two different countries, $TeamWork_i$ in an indicator for whether or not pair i is working as a team, X_i is a vector of controls for the average characteristics between contractors

in pair i, *CountryPair$_i$* is a fixed effect for pair i's pair of countries,[39] and *Week$_i$* is a vector of binary indicators for the week during which pair i worked. With the inclusion of country-pair fixed effects, I estimate β_1 and β_2 by comparing the average performance of independent contractors within a given pair of countries with the average performance of teams made up of contractors in the same pair of countries.

4 Results

I begin by presenting estimations of the effects of teamwork, pair diversity, and the interaction between the two on performance, followed by estimates of how these coefficients differ depending on how important teamwork is for a given pair. Specifically, I consider whether or not the task being worked on requires both PHP and Javascript skills, and whether or not pair mates have specialized skills. I provide several robustness tests of my findings, including evidence that national diversity is not simply a proxy for ethnic or language diversity, or for geographic distance. I then present results on possible explanations for these main findings. In particular, I explore the possibility that communication is harder in diverse teams, that workers have more negative ex-ante expectations about the performance of nationally diverse teams, and that workers differentially dislike teammates who are from other countries.

4.1 Main Results

The results from estimating Equation 1, presented in Table 5, show findings that are consistent with those presented in Figures 2 and 3. Columns 1-3 estimate the effects of teamwork and national diversity on joint output, and Columns 4-6 estimate these effects on joint productivity. Columns 1 and 4 estimate the effects of teamwork and national diversity on performance without controls or fixed effects, Columns 2 and 5 add country pair and week fixed effects, and Columns 3 and 6 add controls. Including the fixed effects and controls does not affect the estimated coefficients on teamwork. I present the full set of estimated control coefficients in Appendix A. For the remainder of the paper, I focus exclusively on coefficients estimated with the full set of fixed effects and controls.

---

[39]I include 3 contractor countries in the experiment, so there are six country pairings.

Table 5: Effect of Team Work & National Diversity on Output and Productivity

	(1)	(2)	(3)	(4)	(5)	(6)
	Joint Output			Joint Productivity		
Team Work	0.375**	0.322*	0.318*	0.079**	0.070*	0.072**
	(0.183)	(0.180)	(0.164)	(0.037)	(0.036)	(0.033)
National Diversity*Team Work	-0.716***	-0.666***	-0.669***	-0.172***	-0.179***	-0.174***
	(0.237)	(0.242)	(0.236)	(0.050)	(0.053)	(0.051)
National Diversity	0.171			0.084**		
	(0.177)			(0.035)		
Country Pair & Week Fixed Effects	No	Yes	Yes	No	Yes	Yes
Controls	No	No	Yes	No	No	Yes
H_o: Team Work+National Diversity*Team Work=0		0.014**	0.008***		0.003***	0.001***
Observations	162	162	162	162	162	162
Mean Dependant Variable Independent Work		1.062			0.135	
R-squared	0.069	0.200	0.265	0.070	0.228	0.278

Notes: An observation is a pair of workers. Robust standard errors in parentheses. Controls are team averages for member education, platform experience, non-platform work experience, number of platform tests, presence of a profile page, gender, wage bid, and agency membership. Joint Output is the total number of features added by an observation. Joint productivity is the total number of features added by an observation divided by the number of hours worked on the task by the pair. The H_o: Team Work+National Diversity*Team Work=0 row reports the p-values of this test.
* significant at 10%; ** significant at 5%; *** significant at 1%

The coefficient estimates in Column 3 show that teamwork is beneficial for nationally homogeneous teams and harmful for nationally diverse teams; teamwork increases output by 30% relative to the independent contractor sample mean of 1.06 for nationally homogeneous teams and decreases it by 33% for nationally diverse teams. The coefficient estimates in Column 6 suggest even larger effects for productivity. In particular, relative to 0.14 features per hour (the mean productivity of independent contractors) teamwork increases productivity by about 50% for nationally homogeneous teams and decreases productivity by about 75% for nationally diverse teams. For a clearer interpretation, I estimate the effects of teamwork national diversity on joint output using an ordered logit regression. The results of this estimation are in Table A4. The estimates show that team work among nationally homogeneous pairs significantly decreases the likelihood of adding no features, and significantly increase the likelihood of adding 2 and 3 features. The reverse is true for nationally diverse pairs of contractors. It is worth noting that team work does not significantly affect the number of hours worked on the job (see Table A5). The differences in productivity across treatment groups is driven almost completely by the differences in output, and, therefore, in the effectiveness with which different treatments groups used the same amount of time.

As the p-values for the test of whether the sum of the team work and interaction coefficients is equal to zero demonstrate, team work leads to worse outcomes than independent work for nationally diverse teams. This suggests that nationally diverse teams could have performed better than they did simply by ignoring each other and working as independent workers. One explanation for why this is occurring is that contractors dislike people from other countries and gain more utility from harming a teammate from another country than they do from the pay-offs of performing well on the task.[40] However, as I show in section 4.2.3, this does not appear to be the case. An alternative explanation is that contractors in nationally diverse teams lose efficiency because of coordination problems and are not aware that this is occurring. I provide evidence in section 4.2.2 consistent with this latter hypothesis.

4.1.1   Robustness of Baseline Estimates

In addition to team-based measures of performance, I also collected data on individual contractor performance. Individual contractor performance is measured as the number of features a contractor added and the number of features per hour worked the contractor added.[41] I estimate Equation 1 with individual output and individual productivity as the dependent variables and present the results in Table A9. The results show that teamwork significantly improves individual output for contractors in nationally homogeneous pairs but does not significantly impact their productivity. Team work significantly reduces both output and productivity for contractors in nationally diverse pairs. The size of the coefficients in Table A9 relative to the sample mean performance are com-

---

[40]In this case, the primary pay-off from good performance is the reputational boost contractors receive through ratings and feedback. They may also anticipate that there is a possibility of being re-hired for a subsequent job.

[41]This is distinct from team performance because if a given feature is added by both contractors in a pair, it counts once towards team output.

parable to the size of the coefficients in Table 5. Given that the only difference between individual and team level measures of performance is whether or not duplicate feature completions within pairs get counted once or twice, this finding suggests that teamwork does not have a large effect on whether pair members complete overlapping features or not.

I also test the effect of national diversity on team work by restricting the sample to those pairs in the team treatment. The results are presented in Table A10. The estimates show a sizable negative effect of national diversity on team performance that is consistent with the estimates on the full sample of pairs.

To ensure that national diversity is not acting solely as a measure of geographic distance, I test whether including the geographic distance in kilometers between pair-member cities eliminates the effect of national diversity on performance. This analysis is restricted to contractors for whom I was able to collect precise city-location data for.[42] These results are reported in Table A11. As these results show, geographic distance does not have a significant effect on joint output or productivity once national diversity is included in the regressions. In addition, though some statistical significance is lost (perhaps due at least in part to the smaller sample size), including geographic distance does not eliminate the effects of national diversity.

National diversity may also be acting as a proxy for ethnic diversity which also varies within country. To test whether this is the case, or whether it is capturing something that varies specifically at the national level I perform several tests. First, given that India is the most ethnically diverse country in my sample (e.g. Fearon, 2003), I test whether team organization among Indian contractors performs worse than among Pakistani or Bangladeshi contractors. Second, I test whether, conditional on having two contractors from India in a pair, teams with one contractor from Southern India and one from Northern India perform worse.[43] These results are presented in Tables A6 and A7. Consistent with the notion that what is driving performance differences across teams is at the national level, I do not find that Indian teams performance worse than Pakistani or Bangladeshi teams, nor do I find that teams of contractors from North and South India perform worse than those from the same region of India.

To further test whether shared or diverse ethnicities drive my main findings, I test whether teams with one contractor from the Punjab State in India and one from Pakistan and teams with one contractor from the West Bengal State of India and one from Bangladesh suffer less from their national diversity than other nationally diverse teams.[44] I find no evidence that these teams perform differently than other nationally diverse teams (see Table A8). However, only 5% of the team sample is made up of these similar pairs.[45]

---

[42]Some city names do not identify a unique city in contractor countries. Contractors from these cities are not included in this analysis.

[43]Northern and Southern India are arguably culturally quite distinct (e.g. Dyson and Moore, 1983).

[44]These regions are more similar to one another than other cross-country regions in my sample. For instance, Punjabi is spoken in both the Punjab State of India and in Pakistan (CIA, 2016b; Government of Punjab, 2016) and that Bangla is spoken in both West Bengal and in Bangladesh (CIA, 2016a; UNICEF, 2016).

[45]An additional test would be whether religious differences in teams affect my results. However, in my sample,

Lastly, although I hired contractors who speak English and communicated with them exclusively in English, contractors in homogeneous teams may benefit because they are able to speak in their native language. Anecdotal evidence from team room screen shots suggests homogeneous teams did not primarily communicate in their native languages. Of the eight observed conversations within nationally homogeneous teams, six were conducted completely in English and two were conducted in a mix of English and another language.[46] Moreover, the evidence that teams of contractors from North and South India perform no worse than teams of contractors within similar regions in India and that teams of contractors from similar regions across countries perform no better than those from more ethnically distinct regions (see Tables A7 and A8) also suggest language difficulties are not driving my results.

4.1.2 Team Collaboration

Table 6 shows the coefficients from estimating Equation 1 with an indicator for whether or not each feature included in the task was implemented. Columns 1 and 2 present the estimated coefficients of the effect of teamwork for nationally homogeneous and nationally diverse pairs of contractors on the likelihood of implementing the Javascript and PHP features, respectively. These results show that teamwork had more of an effect on the implementation of the Javascript feature than the PHP feature for which teamwork had no significant effect. The findings reported in Column 1 are not what I expected. In particular, the Javascript feature was designed to only require Javascript knowledge to be implemented, and therefore, I did not anticipate that it would depend on teamwork. The reason I believe that it does appear to be affected by teamwork is because of an unanticipated attribute of the task. In particular, although the writing of the Javascript feature only required Javascript knowledge, the script that the feature was to be added to is PHP-based. As a result, contractor pairs may have benefited from knowledge of PHP when determining where in the script to implement the feature.

Column 3 presents the estimated coefficients of the effect of teamwork for nationally homogeneous and nationally diverse pairs of contractors on the likelihood of implementing the feature that requires both Javascript and PHP. These results show that teamwork significantly increases the likelihood that the combined feature will be implemented when pairs of contractors live in the same country and significantly decreases the likelihood of implementation when contractors in a pair are from different countries. Conditional on my reasoning for why teamwork matters for the Javascript feature being true, these results suggest that suggest that teamwork has a larger effect on performance when two types of knowledge are required and, therefore, when coordination is more important.

The findings in Table 6 suggest that teamwork has a large impact when collaboration is more important. It follows then that teamwork will be most important when contractors in a pair specialize

---

there is not enough variation in religion measured by contractor last names across countries to perform this test.

[46] All observed nationally diverse team communication was in English.

Table 6: Effect of Team Work & National Diversity by Task Feature

	(1) Javascript Feature	(2) PHP Feature	(3) Combined Feature
Team Work	0.272**	-0.130	0.165***
	(0.110)	(0.104)	(0.059)
National Diversity*Team Work	-0.349**	-0.054	-0.274***
	(0.161)	(0.148)	(0.090)
H_o: Team Work+National Diversity*Team Work=0	0.014**	0.747	0.002***
Observations	162	162	162
Mean Dependent Variable Independent Work	0.568	0.444	0.074
R-squared	0.225	0.314	0.240

Notes: An observation is a pair of workers. Robust standard errors in parentheses. Controls and fixed effects included in these regressions are equivalent to those included in Table 5. The dependent variable is equal to one if the feature specified in the column headers was successfully implemented, and zero otherwise. The H_o: Team Work+National Diversity*Team Work=0 row reports the p-values of this test. * significant at 10%; ** significant at 5%; *** significant at 1%

in one of the two skills required for the task (e.g., when one member knew Javascript but not PHP and the other knew PHP but not Javascript). Testing whether this is the case helps clarify whether collaboration is driving the main results. In addition, it provides a test of whether skill diversity in teams changes the effects of national diversity which contributes to our understanding of whether skill specialization exacerbates or improves communication difficulties. I do this test by estimating Equation 1 separately for pairs with skill differences and those without any. I present the results of this estimation in Table 7. The estimates in Table 7 confirm that the costs of teamwork for nationally diverse pairs are highest when contractors have different skill sets. The coefficients on the interaction term in Columns 1 and 2 are significantly different at the 5% level and the interaction term estimates in Columns 3 and 4 significantly different at the 15% level.[47] Similar findings using interactions between the skill differences, national diversity, and team work variables are presented in Table A12.

4.2 Interpretation

In this subsection, I consider three explanations for my findings. In particular, I explore the data to determine whether differences in communication costs can explain the output and productivity

---

[47]The coefficients on the estimated effect of teamwork for nationally homogeneous pairs are higher with skill differences but not significantly so (p-value=0.18 when output is the dependent variable, p-value=0.35 when productivity is the dependent variable).

Table 7: Effect of Team Work & National Diversity on Performance by Pair Skill Differences

VARIABLES	(1)	(2)	(3)	(4)
		Joint Output		Joint Productivity
	No Skill Difference	Skill Difference	No Skill Difference	Skill Difference
Team Work	0.059	0.436	0.045	0.092*
	(0.234)	(0.262)	(0.036)	(0.053)
National Diversity*Team Work	-0.277	-1.062***	-0.109*	-0.228***
	(0.372)	(0.328)	(0.057)	(0.085)
H_o: Team Work+National Diversity* Team Work=0	0.547	0.008***	0.083*	0.010***
Observations	75	87	75	87
Mean Dependent Variable Independent Work	1.083	1.044	0.132	0.137
R-squared	0.517	0.437	0.550	0.310

Notes: An observation is a pair of workers. Robust standard errors in parentheses. Controls and fixed effects included in these regressions are equivalent to those included in Table 5. A pair has a skill difference if the individuals in the pair have different knowledge of Javascript and/or PHP. The H_o: Team Work+National Diversity*Team Work=0 row reports the p-values of this test. * significant at 10%; ** significant at 5%; *** significant at 1%

differences. I also test for evidence of dislike for national diversity, and for negative ex-ante expectations about the performance of nationally diverse teams as a mechanism for my results. I find suggestive evidence in support of the former explanation, and no evidence in support of the latter two.

4.2.1 Communication

One possibility for differences in performance across the treatment groups is that the costs of communication differ for nationally homogeneous and nationally diverse teams. Although the countries included in my experiment may be more similar to each other than many other countries and all contractors speak English, communication barriers likely remain. For example, Pakistan, India, and Bangladesh were officially a single country until 1947, but citizens began calling for separation well before then (Khan, 2007; Thursby, 1975). Importantly, while the drawing of borders between the three countries was arguably arbitrary, a significant amount of migration based on employment opportunities, and cultural and political preferences occurred between the regions following separation (Khan, 1974; Robinson, 1989).[48]

---

[48]Historians maintain that important societal and cultural differences exist between the three countries, and that these differences have been reinforced through periods of discord. For example, Robinson (1989) suggests the existence of a "...winner-take-all atmosphere of the post-independence period in [Bangladeshi political culture] with its emphasis on extreme individualism." Thursby (1975) writes that "toward the end of the nineteenth century in northern India the lines of religion, language, and script were becoming established between Hindus and Muslims." There is also

That communication difficulties are causing poor performance in cross-country teams is consistent with the findings on feature completion and skill specialization reported in the previous subsection. In particular, these results demonstrate that cross-country team performance suffered significantly more when collaboration, and therefore communication, was more important.[49]

To further investigate whether communication costs differ for nationally homogeneous and nationally diverse teams, I estimate Equation 1 with the difference between what contractor pairs reported as having implemented and what was actually implemented as the outcome variable. At the completion of the task, contractors are required to report which features are completed in the script they turn in.[50] This measure is a proxy for how successfully contractors communicated with their teammates by measuring whether contractors were aware of what their teammate was able to complete. This test is done on the sample of teams[51] and the results are presented in Table 8. The estimates suggest that nationally homogeneous teams are better able to communicate than nationally diverse teams because the error in what contractors report as completed is significantly higher in diverse teams than in homogeneous teams.

Although I do not observe a representative or full sample of team communication[52], the 16 non-random conversations that I do observe support communication being more difficult in nationally diverse teams. Specifically, although I observe the same number of teams communicating in both the nationally diverse and homogeneous treatments, the average number of observed words in nationally diverse team chat logs is 111 and the average in nationally homogeneous chat logs is 63. This suggests that nationally diverse teams spent more time trying to coordinate than nationally homogeneous teams.

Taken together, the evidence from task completion, skill specialization, misreporting, and chat logs suggests that communication is more difficult for nationally diverse teams. More difficult communication requires more time, which can harm performance by shifting time invested in the

---

empirical evidence that these differences show up in labor markets. For instance, India and Pakistan differ significantly in their human resources management practices (Aycan et al., 2000). Measures of labor market cultural distance also show large differences between Bangladesh, India, and Pakistan on several dimensions (Hofstede, 1983).

[49]An alternative but related explanation for the finding that skill specialization worsened the performance of cross-country teams is that knowledge of a particular skill set includes knowledge of a type of language. When contractors share a skill, they also share knowledge of how to communicate about the skill and this may substitute for other communication barriers. However, findings that skill differences *benefit* nationally homogeneous teams suggest collaboration within teams with specialized skills is likely more intensive relative to collaboration between contractors with a shared skill.

[50]It is not clear whether contractors understood the request to report what they had completed as requesting their individual addition or the joint addition. However, given the difference in misreporting across the team and independent work treatments, it appears that contractors in the two treatments did interpret the request differently. Moreover, even if contractors in the team treatments decided to report their individual additions, this may still provide a measure of miscommunication if they incorporated advice from their teammate into their individual additions.

[51]The sample is restricted to teams because pairs of independent contractors cannot communicate.

[52]Screenshots are only taken every 10 minutes, and contractors receive a warning before a screenshot is taken so they can decide what is displayed in the pictures. In addition, conversations seem to be concentrated within specific time periods. As a result, communication data from screenshots is sparse and complete conversations are almost never observable.

Table 8: Effect of Team Work & National Diversity on Task Reporting Accuracy

	(1) Difference between Reported and Completed
National Diversity	0.349*
	(0.191)
Observations	81
Mean Dependent Variable National Homogeneous Teams	0.241
R-squared	0.385

Notes: An observation is a pair of workers. Robust standard errors in parentheses. Controls and fixed effects included in these regressions are equivalent to those included in Table 5. The sample is restricted to teams. * significant at 10%; ** significant at 5%; *** significant at 1%

task towards communication and away from production. This is consistent with the finding that discussed in section 4.1 that, on average, homoegenous and diverse teams spend the same amount of time on the task however diverse teams use this time less effectively. Moreover, given that nationally diverse teams perform worse than independent pairs of workers, these teams do not appear to realize the productivity loss attempting to coordinate causes. Tests of whether nationally diverse teams have different ex-ante expectations or ex-post beliefs about their performance than homogeneous teams do in the next two sections support this interpretation.

4.2.2 Performance Expectations

An alternative explanation to communication difficulties for the poor performance of nationally diverse teams is that they had different expectations about the performance of their team going into the task. For instance, they may have believed that they would not be able to work well with someone from another country and, as a result, their performance may have reflected this belief (Schmader et al., 2008). Alternatively, they may have believed that someone from another country would not be able to do a good job on the task, and was therefore not worth trying to work with although anecdotal evidence from chat logs does not support this possibility.

To test whether nationally diverse and homogenous teams had different ex-ante expectations about their performance, I test whether the intensity of effort exerted on the task from the outset is different across the two team types. To do this, I code the average effort level for every ten minute period across the total amount of time each team worked on the job. Effort levels are measured on oDesk through activity meters which measure mouse and keyboard use on a scale from one to ten. Figure 4 shows polynomials of average team effort levels over time contractors in nationally homogeneous teams, and those in nationally diverse teams. The figure demonstrates that the national composition

of teams does not appear to affect effort levels at the point when the job is started. This suggests that nationally diverse and nationally homogeneous teams start the job with the same expectations about the returns to their effort despite knowing which country their teammate was from before beginning the task. At about 50% of the way through the job, however, nationally diverse team effort begins to fall below that of nationally homogeneous effort suggesting that communication difficulties may have slowed these teams down as time went on.[53]

In addition to this more formal evidence, anecdotal evidence from my interactions with contractors suggests they did not did not have more negative expectations about the performance of nationally diverse teams. In particular, I did not receive any complaints from workers about who they had been paired with once they had learned which country their teammate was from. I did, however, receive complains and suggestions on other task requirements.

Figure 4: Effort Over Time by Team Treatment
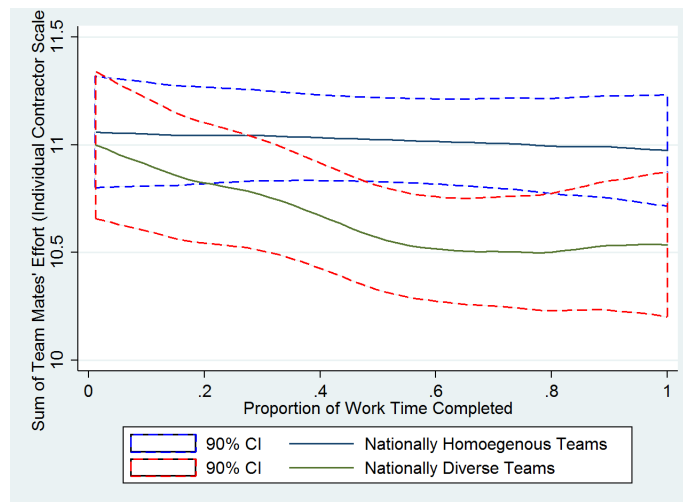
Notes: This chart reports polynomials of average time effort levels over time across the team treatment groups. Dashed lines represent 90% confidence intervals.

4.2.3   Animosity in Nationally Diverse Teams

Based on the structure of the task and of contractor incentives on oDesk, I have thus far assumed that the incentives of the contractors in my study are aligned. One concern, however, is that they may not be in diverse teams, for instance because contractors' pay-offs change based on preferences for teammate nationality unrelated to communication difficulties and complementarities

---

[53]This is consistent with theory that more difficult communication takes effort away from the task (Dessein and Santos, 2006), for instance because workers spend time waiting for each other to show up or to complete a portion of the job, or because time is spent communicating as opposed to writing code. Mouse and keyboard activity is more likely to be higher when workers are working on their code than when they are attempting to coordinate with their teammates.

because of negative feelings towards people from other countries (Alesina and La Ferrara, 2005). The average effort levels across teams demonstrated in Figure 4 suggest nationally diverse teams begin the task with aligned incentives. However, I provide additional empirical tests of this explanation for my findings here.

I investigate whether there is any empirical evidence for this in the data by analyzing contractor survey responses.[54] I asked contractors in the teamwork treatment who agreed to answer the survey questions whether they thought they had done more work than their teammates or vice versa (where 1 is "my teammate did almost all the work" and 7 is "I did almost all the work") and whether they would be willing to work with their teammates from this task or would prefer a different teammate for any follow-up tasks (where 1 is "I would prefer another teammate" and 7 is "I would prefer this teammate"). Contractors knew that the employer was going to see these answers and therefore could have used them to damage their teammates' reputations or to make themselves appear better. Table 9 presents results from regressions of the cross-country team indicator on these survey questions.[55] The results give no indication that contractors in cross-country teams had a more negative view of their teammates or that they were less likely to want to work with their teammates again. This also suggests that contractors in nationally diverse teams may not have recognized the relatively poor performance of their teams.

Table 9: Effect of National Diversity on Opinion of Teammate

	(1) Reported Doing More Work than Teammate	(2) Willingness to Work With Teammate Again
National Diversity	-0.067	0.598
	(0.324)	(0.382)
Observations	120	120
Mean Dependent Variable No National Diversity	5.598	4.133
R-squared	0.182	0.375

Notes: An observation is a pair of workers. Robust standard errors in parentheses. Controls and fixed effects included in these regressions are equivalent to those included in Table 5. Independent variables are collected through a survey. * significant at 10%; ** significant at 5%; *** significant at 1%

These results are suggestive that contractors' preferences for teammates based on their nationality is not driving the results. However, the survey may not capture what contractors really thought and the response rate is not 100%, so there may be some selection in who replied. Another way to test for this mechanism is to restrict the sample to countries that are less likely to have feelings of animosity towards each other. Most concerning in this regard are teammates from India and Pakistan, given the current dispute between these countries.[56] To test whether this dispute is driving

---

[54]The survey response rate is 83%.

[55]The results are restricted to the sample of contractors in the team treatment.

[56]Hjort (2014) shows that when team members are from conflicting ethnicities, productivity suffers.

the negative coefficient on the interaction between cross-country teams and teamwork, I estimate Equation (1) excluding all pairs with at least one contractor from Pakistan. I present the results from this estimation in Table 10. The results in Table 10 are consistent with those in Table 5,[57] and are not consistent with the negative effect of teamwork on nationally diverse pair performance being driven by dislike for national diversity in teams.[58]

Table 10: Effect of Team Work & National Diversity on Performance, Pakistan Excluded

	(1) Joint Output	(2) Joint Productivity
Team Work	0.307*	0.079**
	(0.178)	(0.035)
National Diversity*Team Work	-0.530	-0.176***
	(0.323)	(0.061)
H_o: Team Work+National Diversity*Team Work=0	0.063*	0.004***
Observations	108	108
Mean Dependent Variable Independent Work	1.059	0.121
R-squared	0.325	0.330

Notes: An observation is a pair of workers. Robust standard errors in parentheses. Controls and fixed effects included in these regressions are equivalent to those included in Table 5. Pairs with at least one worker from Pakistan are excluded. The H_o: Team Work+National Diversity*Team Work=0 row reports the p-values of this test. * significant at 10%; ** significant at 5%; *** significant at 1%

5 Conclusion

As technology continues to facilitate international markets, collaboration between market participants from all over the world is becoming more common and multinational work teams are likely to become increasingly necessary. Understanding the trade-offs associated with these teams has important implications for employers, and workers. This paper considers what the value of teamwork is and how it varies with national diversity among contract workers using a field experiment in an international labor market to answer these questions.

Findings from the experiment show that allowing teamwork improves outcomes for contractors in nationally homogeneous pairs but worsens outcomes for contractors in nationally diverse teams. In particular, nationally diverse pairs perform worse when working as teams than they do when working independently. Moreover, teamwork has a more negative impact on outcomes for nationally

---

[57]The statistical significance of coefficients in Table 10 is lower, likely due to the smaller sample size.

[58]The results hold when I drop any one of the six country-pairs of contractors from the sample.

diverse pairs when team members have specialized skills suggesting that overlapping skills either allows workers to avoid coordination, or improves their ability to communicate by providing a shared context on which to understand each other. Further investigation of the data suggests that communication is more difficult for nationally diverse teams than it is for nationally homogeneous teams, but that contractors are not aware of this. I find no evidence that animosity between workers from different countries affects their performance or that nationally diverse teams begin the task expecting to perform worse than nationally homogeneous teams.

It is important to note some boundary conditions of the analysis in this paper. First, teams had eight hours to work on their tasks, and team members had no prior interactions. Therefore, the experiment considers the impact of national diversity on the value of teamwork during the period immediately after team members are first introduced. For longer term labor contracts, as team members spend more time together, contractors may learn how to deal with the communication difficulties that arise due to national differences. Second, the experiment restricts analysis to three countries. The countries included are relatively similar compared to many other country pairs so the results presented here may be a lower bound on the value of teamwork for cross-country pairs. Third, the task assigned in this experiment has both routine and creative elements (e.g., Boudreau et al., 2011), which reflects many common real-world tasks; however, it may not require the type or extent of creativity that prior research has shown diverse teams excel at (e.g., creative generation tasks). To the extent that the assigned task is not a task that benefits from diversity, this study can be thought of as a test of the costs associated with national diversity in teams.[59] Fourth, these teams are not overseen and managed by an employer who helps to facilitate coordination. As prior literature has shown, the success of diverse teams depends on how they are managed (e.g. DiStefano and Maznevski, 2000; Gong, 2003; Jarvenpaa et al., 1998; Kochan et al., 2003). Therefore, the results in this paper provide estimates of the performance of teams in the absence of managerial intervention. Finally, communication encompasses a number of different elements including interpretation, trust, and knowledge (Kintsch, 1988; Rode, 2010). Understanding the causes of communication failures in nationally diverse teams is an important area for future research.

The research I present in this paper contributes to the organizational and labor economics literature by identifying the value of organizing workers in to teams separately for nationally diverse and nationally homogeneous pairs of contractors and by providing suggestive evidence that even in the absence of taste-based discrimination, nationally diverse teams have difficulties coordinating. These results build on theoretical research that considers the costs and benefits of team work when communication costs vary across possible teams (e.g. Dessein and Santos, 2006; Lazear, 1999). It also contributes to literature on diversity in organizations by providing the first experimental test of the performance effects of national diversity in teams. Importantly, my experimental evidence is consistent with national diversity capturing something other than ethnic or geographic diversity. Furthermore, this paper has implications for research on immigrant labor market success. For instance, prior findings in this literature suggest that employers are less willing to hire immigrants than equally capable native workers (Hunt, 2013). Consistent, with survey and observational ev-

---

[59]However, if it is the case that diversity only benefits teams in cases of extreme task creativity, it is unlikely to benefit many real-world teams.

idence (Bell, 1997; Leslie and Lindley, 2003; Oreopoulos, 2011, e.g.), the results presented here suggest that some of this hesitancy may be a result of the costs of assimilating different nationalities in the workplace.

The results presented in this paper also have important practical implications. In particular, there are potentially large gains to cross-country collaboration, including knowledge transfer, market growth, and access to higher paying jobs; however, the findings suggest that managers hiring from an international pool of workers should be cognizant of the costs of a diverse labor force and invest in managing these differences.[60] My findings suggest that one way for managers to reduce the costs associated with national diversity is to ensure teammates have another form of shared knowledge. In addition, participants in international markets may benefit from investing in culture-specific human capital. Education policy may play a role in easing the costs of collaborating across countries by standardizing some of the content of curricula across countries, and by providing cultural training programs for international market participants.

6   References

Adams, Renée B and Daniel Ferreira, "Women in the boardroom and their impact on governance and performance," *Journal of Financial Economics*, 2009, *94* (2), 291–309.

Agrawal, A., J. Horton, N. Lacetera, and E. Lyons, "Digitization and the Contract Labor Market: A Research Agenda," Working Paper 19525, National Bureau of Economic Research 2013.

_ , **N. Lacetera, and E. Lyons**, "Does Information Help or Hinder Job Applicants from Less Developed Countries in Online Markets?," Working Paper 18720, National Bureau of Economic Research 2013.

Alesina, Alberto and Eliana La Ferrara, "Ethnic Diversity and Economic Performance," *Journal of Economic Literature*, 2005, *43*, 762–800.

Alesina, Alberto F, Johann Harnoss, and Hillel Rapoport, "Birthplace Diversity and Economic Prosperity," *NBER Working Paper*, 2013, (w18699).

Alesina, Alberto, William Easterly, and Janina Matuszeski, "Artificial states," *Journal of the European Economic Association*, 2011, *9* (2), 246–277.

Audretsch, David B and Maryann P Feldman, "Knowledge spillovers and the geography of innovation," *Handbook of Regional and Urban Economics*, 2004, *4*, 2713–2739.

---

[60]It is important to note that my experiment does not account for costs of avoiding national diversity. In some instances, it may be necessary to hire from a particular set of countries in order for a task to be completed (for instance, in situations where local knowledge of a set of foreign markets is required).

Autor, David H., "Wiring the Labor Market," *Journal of Economic Perspectives*, 2001, *15* (1), 25–40.

Aycan, Zeynep, Rabindra Kanungo, Manuel Mendonca, Kaicheng Yu, Jürgen Deller, Günter Stahl, and Anwar Kurshid, "Impact of culture on human resource management practices: A 10-country comparison," *Applied Psychology*, 2000, *49* (1), 192–221.

Baert, Stijn, Bart Cockx, Niels Gheyle, and Cora Vandamme, "Is there less discrimination in occupations where recruitment is difficult?," *ILR Review*, 2015, p. 0019793915570873.

Barro, Robert J and Jong Wha Lee, "A new data set of educational attainment in the world, 1950–2010," *Journal of Development Economics*, 2013, *104*, 184–198.

Bell, Brian D, "The performance of immigrants in the United Kingdom: evidence from the GHS," *The Economic Journal*, 1997, *107* (441), 333–344.

Bertrand, Marianne and Esther Duflo, "Field experiments on discrimination," Technical Report, National Bureau of Economic Research 2016.

Boone, Christophe and Walter Hendriks, "Top management team diversity and firm performance: Moderators of functional-background and locus-of-control diversity," *Management Science*, 2009, *55* (2), 165–180.

Boudreau, K. J., N. Lacetera, and K. Lakhani, "Incentives and Problem Uncertainty in Innovation Contests: An Empirical Analysis," *Management Science*, 2011, *57*, 843–863.

Brewer, Marilynn B, "In-group bias in the minimal intergroup situation: A cognitive-motivational analysis.," *Psychological bulletin*, 1979, *86* (2), 307.

CIA, "The World Factbook: Bangladesh," https://www.cia.gov 2016.

_ , "The World Factbook: Pakistan," https://www.cia.gov 2016.

Cooper, David J and John H Kagel, "Are two heads better than one? Team versus individual play in signaling games," *The American Economic Review*, 2005, *95* (3), 477–509.

Dahlin, Kristina B, Laurie R Weingart, and Pamela J Hinds, "Team diversity and information use," *Academy of Management Journal*, 2005, *48* (6), 1107–1123.

Dessein, Wouter and Tano Santos, "Adaptive organizations," *Journal of Political Economy*, 2006, *114* (5), 956–995.

DiStefano, Joseph J and Martha L Maznevski, "Creating value with diverse teams in global management," *Organizational Dynamics*, 2000, *29* (1), 45–63.

Dreu, Carsten KW De and Laurie R Weingart, "Task versus relationship conflict, team performance, and team member satisfaction: a meta-analysis.," *Journal of applied Psychology*, 2003, *88* (4), 741.

Dyson, Tim and Mick Moore, "On kinship structure, female autonomy, and demographic behavior in India," *Population and development review*, 1983, pp. 35–60.

Fearon, James D, "Ethnic and cultural diversity by country," *Journal of Economic Growth*, 2003, *8* (2), 195–222.

Ferreira, Daniel and Raaj K Sah, "Who gets to the top? Generalists versus specialists in managerial organizations," *The RAND Journal of Economics*, 2012, *43* (4), 577–601.

Fortune, "100 Best Companies to Work For," http://fortune.com/best-companies/ 2015.

Galbraith, Jay R, "Organization design: An information processing view," *Organizational Effectiveness Center and School*, 1977, *21*, 21–26.

Gong, Y., "Subsidiary Staffing in Multinational Enterprises: Agency, Resources and Performance," *Academy of Management Journal*, 2003, *46*, 307–338.

Grossman, G. M. and E Helpman, "Integration versus Outsourcing in Equilibrium," *The Quarterly Journal of Economics*, 2002, *117*, 85–120.

Guiso, Luigi, Paola Sapienza, and Luigi Zingales, "Cultural biases in economic exchange," in "Quarterly Journal of Economics" Citeseer 2009.

Haas, M. R., "The double-edged swords of autonomy and external knowledge: Analyzing team effectiveness in a multinational organization," *Academy of Management Journal*, 2010, *53* (5), 989–1008.

Hamilton, B. H., J. A. Nickerson, and H Owan, "Team Incentives and Worker Heterogeneity: An Empirical Analysis of the Impact of Teams on Productivity and Participation," *Journal of Political Economy*, 2003, *111*, 465–497.

_ , _ , **and** _ , "Diversity and Productivity in Production Teams," *Advances in the Economic Analysis of Participatory and Labor-Managed Firms*, 2012, *13*, 99–138.

Harrison, D. A. and K. J. Klein, "What's the Difference? Diversity Constructs as Separation, Variety, or Disparity in Organizations," *The Academy of Management Review*, 2007, *32*, 1199–1228.

Hays, R. D., "Insuring Success and Avoiding Failure," *Journal of International Business Studies*, 1974, *5*, 25–37.

Hinds, P. J. and M. Mortensen, "Understanding Conflict in Geographically Distributed Teams: The Moderating Effects of Shared Identity, Shared Context, and Spontaneous Communication," *Organization Science*, 2005, *16*, 290–307.

Hjort, Jonas, "Ethnic divisions and production in firms," *The Quarterly Journal of Economics*, 2014, *129* (4), 1899–1946.

Hofstede, G., "The Cultural Relativity of Organizational Practices and Theories," *Journal of International Business Studies*, 1983, *14*, 75–89.

_ , "The Hofstede Centre: National Culture," http://geert-hofstede.com/countries.html 2013.

Hofstede, Geert, *Culture's consequences: International differences in work-related values*, Vol. 5, sage, 1984.

Holmstrom, Bengt, "Moral hazard in teams," *The Bell Journal of Economics*, 1982, pp. 324–340.

Hoogendoorn, Sander and Mirjam Praag, "Ethnic Diversity and Team Performance: A Field Experiment," Technical Report, Tinbergen Institute 2014.

Horton, John J., "Online Labor Markets," *Internet and Network Economics*, 2010, *6484*, 515–522.

Horwitz, S. K. and I. B. Horwitz, "The Effects of Team Diversity on Team Outcomes: A Meta-Analytic Review of Team Demography," *Journal of Management*, 2007, *33*, 987–1015.

Hunt, J., "Are Immigrants the Best and Brightest U.S. Engineers?," Working Paper 18696, National Bureau of Economic Research 2013.

Ingram, Beth F and George R Neumann, "The returns to skill," *Labour economics*, 2006, *13* (1), 35–59.

International Labor Organization, *International Labour Migration Statistics* 2012.

_ , *Labor Migration* 2012.

Ipeirotis, Panos, "The Emergence of Teams in Online Work," 2012.

Islamabad, UNESCO, "Teacher Education in Pakistan," http://unesco.org.pk/ 2013.

Jarvenpaa, Sirkka L, Kathleen Knoll, and Dorothy E Leidner, "Is anybody out there? Antecedents of trust in global virtual teams," *Journal of management information systems*, 1998, *14* (4), 29–64.

Jones, B., "The Knowledge Trap: Human Capital and Development Reconsidered," Working Paper 2011.

Jones, Benjamin F, "The burden of knowledge and the "death of the Renaissance man": is innovation getting harder?," *The Review of Economic Studies*, 2009, *76* (1), 283–317.

Jones, Reece, "Sovereignty and statelessness in the border enclaves of India and Bangladesh," *Political Geography*, 2009, *28* (6), 373–381.

Kaas, Leo and Christian Manger, "Ethnic discrimination in Germany's labour market: a field experiment," *German Economic Review*, 2012, *13* (1), 1–20.

Kachru, Braj B, *The other tongue: English across cultures*, University of Illinois Press, 1992.

Kao, Grace and Jennifer S Thompson, "Racial and ethnic stratification in educational achievement and attainment," *Annual review of sociology*, 2003, pp. 417–442.

Khan, Masihur Rahman, "Pattern of External Migration to and from Bangladesh, 1901-1961," *The Bangladesh Economic Review*, 1974, pp. 599–632.

Khan, Yasmin, *The great partition: The making of India and Pakistan*, Yale University Press, 2007.

Kim, K. B., "Direct Employment in Multinational Enterprises: Trends and Implications," MULTI Working Paper, International Labour Organization 101 2006.

Kintsch, Walter, "The role of knowledge in discourse comprehension: a construction-integration model.," *Psychological review*, 1988, *95* (2), 163.

Kirkman, B. L. and D. L. Shapiro, "The Impact of Cultural Value Diversity on Multicultural Team Performance," *Volume Advances in International Management*, 2005, *18*, 33–67.

Kochan, Thomas, Katerina Bezrukova, Robin Ely, Susan Jackson, Aparna Joshi, Karen Jehn, Jonathan Leonard, David Levine, and David Thomas, "The effects of diversity on business performance: Report of the diversity research network," *Human Resource Management*, 2003, *42* (1), 3–21.

Landes, D. S., *The Wealth and Poverty of Nations: Why Some Are So Rich and Some So Poor* 1999.

Lang, Kevin, "A language theory of discrimination," *The Quarterly Journal of Economics*, 1986, pp. 363–382.

Lazear, Edward P., "Globalisation and the Market for Team-Mates," *The Economic Journal*, 1999, *109* (454), 15–40.

Leslie, D. and J. Lindley, "The Impact of Language Ability on Employment and Earnings of Britain's Ethnic Communities," *Economica*, 2003, *68*, 587–606.

Levitt, S. D. and J. A. List, "What do Laboratory Experiments Measuring Social Preferences Reveal About the Real World?," *Journal of Economic Perspectives*, 2007, *21*, 153–174.

Makela, K., U. Andersson, and T. Seppala, "Interpersonal Similarity and Knowledge Sharing within Multinational Organizations," *International Business Review*, 2012, *21*, 439–451.

Manev, I. M. and W. B. Stevenson, "Nationality, Cultural Distance, and Expatriate Status: Effects on the Managerial Network in a Multinational Enterprise," *Journal of International Business Studies*, 2001, *32*, 285–303.

Marx, Benjamin, Vincent Pons, and Tavneet Suri, "Diversity and Team Performance in a Kenyan Organization," 2015.

Mas, Alexandre and Enrico Moretti, "Peers at Work," *The American Economic Review*, 2009, *99* (1), 112–145.

McGinnity, Frances and Peter D Lunn, "Measuring discrimination facing ethnic minority job applicants: an Irish experiment," *Work, Employment & Society*, 2011, *25* (4), 693–708.

Ministry of Human Resource Development Government of India, "Department of School Education & Literacy," http://mhrd.gov.in/schooleducation 2013.

National Curriculum & Textbook Board, Bangladesh, "Curriculum," http://www.nctb.gov.bd/ 2013.

of Punjab, India Government, http://punjab.gov.in/history 2016.

Oreopoulos, P., "Why Do Skilled Immigrants Struggle in the Labor Market? A Field Experiment with Thirteen Thousand Resumes," *American Economic Journal: Economic Policy*, 2011, *3*, 148–171.

Oster, Emily and M Bryce Millett, "Do call centers promote school enrollment? Evidence from India," Technical Report, National Bureau of Economic Research 2010.

Ottaviano, Gianmarco IP and Giovanni Peri, "The economic value of cultural diversity: evidence from US cities," *Journal of Economic Geography*, 2006, *6* (1), 9–44.

Reagans, R., E. Zuckerman, and B. McEvily, "How to make the team: Social networks vs. demography as criteria for designing effective teams," *Administrative Science Quarterly*, 2004, *49*, 101–133.

Robinson, Francis, *The Cambridge Encyclopedia of India, Pakistan, Bangladesh, Sri Lanka, Nepal, Bhutan and the Maldives*, Cambridge University Press Cambridge, 1989.

Rode, Julian, "Truth and trust in communication: Experiments on the effect of a competitive context," *Games and Economic Behavior*, 2010, *68* (1), 325–338.

Rulke, Diane L and Joseph Galaskiewicz, "Distribution of knowledge, group network structure, and group performance," *Management Science*, 2000, *46* (5), 612–625.

Schmader, Toni, Michael Johns, and Chad Forbes, "An integrated process model of stereotype threat effects on performance.," *Psychological review*, 2008, *115* (2), 336.

Thursby, Gene R, *Hindu-Muslim relations in British India: A study of controversy, conflict, and communal movements in northern India 1923-1928*, Vol. 35, Brill, 1975.

UNICEF, "UNICEF India: West Bengal," http://unicef.in 2016.

Wuchty, Stefan, Benjamin F Jones, and Brian Uzzi, "The increasing dominance of teams in production of knowledge," *Science*, 2007, *316* (5827), 1036–1039.

For Online Publication

Appendix A   Additional Tables

Table A1: Pair Characteristic Summary Statistics By Team Work and National Diversity Treatments

| | Panel A | | |
Variable	No Team Work	Team Work	p-value of difference
Number of Job Posting	8.358	8.08	0.580
Applications	(0.328)	(0.378)	
oDesk Rating Prior	4.527	4.62	0.470
to Hire	(0.11)	(0.083)	
No Rating Prior	0.636	0.586	0.397
to Hire	(0.042)	(0.040)	
Number of oDesk Contracts	5.093	3.333	0.134
Prior to Hire	(1.036)	(0.542)	
Indicator for having a	0.815	0.87	0.191
Profile Picture	(0.033)	(0.026)	
Level of Education	1.87	1.821	0.649
	(0.077)	(0.076)	
Number of Offline Jobs	1.16	1.222	0.543
Listed on Profile	(0.075)	(0.067)	
Average Score on oDesk	3.416	3.44	0.736
Tests	(0.044)	(0.058)	
Number of oDesk Tests	2.531	2.488	0.905
Taken	(0.239)	(0.269)	
Wage Bid on the Job	3.729	3.789	0.527
	(0.070)	(0.065)	
Wage Posted on Profile	7.422	6.355	0.326
	(0.329)	(0.447)	
Indicator for Female	0.13	0.16	0.440
Contractor	(0.029)	(0.028)	
Indicator for Agency	0.228	0.278	0.315
Membership	(0.035)	(0.035)	
Number of Items in	3.92	3.821	0.896
Portfolio	(0.562)	(0.508)	
Number of Observations	80	82	

| | **Panel B** | | |
Variable	No Team Work	Team Work	p-value of difference
Number of Job Posting	8.15	8.287	0.786
Applications	(0.318)	(0.386)	
oDesk Rating Prior	4.543	4.607	0.635
to Hire	(0.112)	(0.080)	
No Rating Prior	0.619	0.604	0.796
to Hire	(0.043)	(0.04)	
Number of oDesk Contracts	3.625	4.787	0.324
Prior to Hire	(0.713)	(0.927)	
Indicator for having a	0.863	0.823	0.356
Profile Picture	(0.029)	(0.031)	
Level of Education	1.875	1.817	0.594
	(0.079)	(0.074)	
Number of Offline Jobs	1.2	1.183	0.866
Listed on Profile	(0.076)	(0.067)	
Average Score on oDesk	3.464	3.394	0.335
Tests	(0.056)	(0.045)	
Number of oDesk Tests	2.175	2.835	0.066*
Taken	(0.207)	(0.288)	
Wage Bid on the Job	3.788	3.731	0.550
	(0.068)	(0.067)	
Wage Posted on Profile	6.611	7.159	0.326
	(0.329)	(0.447)	
Indicator for Female	0.131	0.159	0.495
Contractor	(0.029)	(0.027)	
Indicator for Agency	0.231	0.274	0.38
Membership	(0.036)	(0.034)	
Number of Items in	4.269	3.482	0.299
Portfolio	(0.588)	(0.477)	
Number of Observations	80	82	

Notes: Standard errors are in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%

Table A2: Effect of Team Work & National Diversity on Team Performance, Robustness to Pairs with Contractors who Lost Contact

	(1)	(2)	(3)	(4)
Pairs with Lost Contractors	Average Performance Assigned		Dropped from Sample	
VARIABLES	Joint Ouput	Joint Producitivity	Team Output	Team Productivity
Team Work	0.272	0.062*	0.312*	0.069**
	(0.165)	(0.033)	(0.172)	(0.034)
National Diversity*Team Work	-0.616***	-0.166***	-0.694***	-0.185***
	(0.234)	(0.050)	(0.248)	(0.054)
H_o: Team Work+National Diversity*Team Work=0	0.011**	0.002***	0.006***	0.001***
Observations	162	162	153	153
Mean Dependant Variable Independent Work	1.062		0.135	
R-squared	0.254	0.266	0.274	0.283

Notes: An observation is a pair of workers. Robust standard errors in parentheses. Country pair and week fixed effects are included in all regressions. Controls included in all regressions are team averages for member education, platform experience, non-platform work experience, number of platform tests, presence of a profile page, gender, wage bid, and agency membership. Joint output is the total number of features added by an observation. Joint productivity is the total number of features added by an observation divided by the number of hours worked on the task by the pair. The H_o: Team Work+National Diversity*Team Work=0 row reports the p-values of this test. * significant at 10%; ** significant at 5%; *** significant at 1%

Table A3: Effect of Team Work & National Diversity on Team Performance, Full Set of Controls

	(1) Joint Output	(2) Joint Productivity
Team Work	0.300*	0.070**
	(0.165)	(0.033)
National Diversity*Team Work	-0.630***	-0.168***
	(0.238)	(0.050)
Number of oDesk Contracts Prior to Hire	-0.010	-0.002
	(0.012)	(0.002)
Indicator for having a Profile Picture	0.169	-0.072
	(0.258)	(0.051)
Level of Education	-0.031	-0.002
	(0.098)	(0.016)
Number of Offline Jobs Listed on Profile	0.020	0.022
	(0.125)	(0.024)
Number of oDesk Tests Taken	0.001	0.012
	(0.034)	(0.011)
Wage Bid on the Job	0.197*	-0.003
	(0.106)	(0.019)
Indicator for Agency Membership	-0.545***	-0.072**
	(0.217)	(0.035)
Indicator for Female Contractor	-0.003	-0.031
	(0.293)	(0.037)
Observations	162	162
Mean Dependant Variable Independent Work	1.062	0.135
R-squared	0.255	0.274

Notes: An observation is a pair of workers. Robust standard errors in parentheses. Country pair and week fixed effects are included in all regressions. Joint Output is the total number of features added by an observation. Joint productivity is the total number of features added by an observation divided by the number of hours worked on the task by the pair. * significant at 10%; ** significant at 5%; *** significant at 1%

Table A4: Effect of Team Work & National Diversity on Output, Ordered Logit

	(1) Joint Output	(2)	(3) Predicted Values	(4)	(5)
VARIABLES	Coefficient Estimates	No Features Added	One Feature Added	Two Features Added	Three Features Added
Team Work	0.871*	0.124	0.472	0.351	0.053
	(0.457)				
National Diversity* Team Work	-1.878***	0.353	0.485	0.148	0.014
	(0.666)				
Independent Work		0.178	0.593	0.210	0.019
Observations	162				
Pseudo R-squared	0.1308				
Wald Chi Squared	50.89				

Notes: An observation is a pair of workers. Robust standard errors in parentheses. Country pair and week fixed effects are included in all regressions. Controls included in all regressions are team averages for member education, platform experience, non-platform work experience, number of platform tests, presence of a profile page, gender, wage bid, and agency membership. Joint Output is the total number of features added by an observation. * significant at 10%; ** significant at 5%; *** significant at 1%

Table A5: Effect of Team Work & National Diversity on Number of Hours Worked

	Hours Worked
Team Work	0.075
	(1.191)
National Diversity*Team Work	0.315
	(1.603)
Constant	3.142
	(3.740)
H_o: Team Work+National Diversity*Team Work=0	0.724
Observations	162
Mean Dependant Variable Independent Work	9.140
R-squared	0.242

Notes: An observation is a pair of workers. Robust standard errors in parentheses. Country pair and week fixed effects are included in all regressions. Controls included in all regressions are team averages for member education, platform experience, non-platform work experience, number of platform tests, presence of a profile page, gender, wage bid, and agency membership. Joint Output is the total number of features added by an observation. The H_o: Team Work+National Diversity*Team Work=0 row reports the p-values of this test. * significant at 10%; ** significant at 5%; *** significant at 1%

Table A6: Effect of Team Work & India Pairs on Output and Productivity

	(1)	(2)	(3)	(4)
	Output		Productivity	
Team Work	0.238	0.119	0.025	0.024
	(0.240)	(0.318)	(0.028)	(0.033)
India Pair	0.148	-0.011	0.015	0.016
	(0.275)	(0.265)	(0.031)	(0.062)
India Pair * Team Work	0.171	0.317	0.078	0.105
	(0.352)	(0.430)	(0.062)	(0.069)
H_o: India Pair+India Pair* Team Work=Team Work	0.834	0.776	0.288	0.385
Controls	No	Yes	No	Yes
Observations	80	80	80	80
Mean Dependant Variable Independent Work, Non-Indian Pairs	0.905	0.905	0.085	0.085
R-squared	0.072	0.544	0.090	0.506

Notes: An observation is a pair of workers. Robust standard errors in parentheses. Week fixed effects and controls for team averages for member education, platform experience, non-platform work experience, number of platform tests, presence of a profile page, gender, wage bid, and agency membership are included in columns 2 and 4. Joint Output is the total number of features added by an observation. The H_o: North-South India Pair+North-South India Pair*Team Work=Team Work row reports the p-values of this test. The sample is restricted to pairs of contractors from the same country. * significant at 10%; ** significant at 5%; *** significant at 1%

Table A7: Effect of Team Work & North-South India Pairs on Output and Productivity

VARIABLES	(1) Output	(2) Productivity
Team Work	-0.076	-0.045
	(0.276)	(0.082)
North-South India Pair	-0.539	-0.112
	(0.504)	(0.107)
North-South India Pair*Team Work	0.652	0.325
	(0.567)	(0.269)
H_o: North-South India Pair+	0.710	0.332
North-South India Pair*Team Work=Team Work		
Observations	45	45
Mean Dependant Variable Independent Work, Same Region of India	0.929	0.084
R-squared	0.797	0.632

Notes: An observation is a pair of workers. Robust standard errors in parentheses. Week fixed effects are included in all regressions. Controls included in all regressions are team averages for member education, platform experience, non-platform work experience, number of platform tests, presence of a profile page, gender, wage bid, and agency membership. Joint Output is the total number of features added by an observation. The H_o: North-South India Pair+North-South India Pair*Team Work=Team Work row reports the p-values of this test. The sample is restricted to pairs of contractors from India. * significant at 10%; ** significant at 5%; *** significant at 1%

Table A8: Effect of Team Work & Ethnically Similar Nationally Diverse Pairs on Output and Productivity

	(1)	(2)	(3)	(4)
	Output		Productivity	
Team Work	-0.274	-0.237	-0.103*	-0.100*
	(0.178)	(0.188)	(0.055)	(0.056)
Diverse Pairs in Ethnically Similar Regions	0.017	0.093	-0.090	-0.029
	(0.364)	(0.414)	(0.074)	(0.067)
Diverse Pairs in Ethnically Similar Regions *	-0.444	-0.525	0.099	0.099
Team Work	(0.549)	(0.545)	(0.105)	(0.110)
H_o: Diverse Pairs in Ethnically Similar Regions+ Diverse Pairs in Ethnically Similar Regions* Team Work=Team Work	0.751	0.693	0.245	0.140
Country Pair Fixed Effects	No	Yes	No	Yes
Mean Dependant Variable Independent Work, Non-Similar Pairs	1.156	1.156	0.199	0.199
Observations	83	83	83	83
R-squared	0.288	0.309	0.336	0.387

Notes: An observation is a pair of workers. Robust standard errors in parentheses. Week fixed effects are included in all regressions. Controls included in all regressions are team averages for member education, platform experience, non-platform work experience, number of platform tests, presence of a profile page, gender, wage bid, and agency membership. Joint Output is the total number of features added by an observation. Ethnically similar diverse pairs are pairs made up of one contractor from Punjab in India and one from Pakistan, or one contractor from West Bengal in India and one from Bangladesh. The H_o: Diverse Pairs in Ethnically Similar Regions+Diverse Pairs in Ethnically Similar Regions*Team Work=Team Work row reports the p-values of this test. The sample is restricted to nationally diverse pairs of contractors. * significant at 10%; ** significant at 5%; *** significant at 1%

Table A9: Effect of Team Work & National Diversity on Individual Performance

	(1) Individual Output	(2) Individual Productivity
National Diversity	0.047	0.025
	(0.102)	(0.027)
Team Work	0.200*	0.036
	(0.107)	(0.030)
National Diversity*Team Work	-0.388***	-0.104***
	(0.143)	(0.038)
H_o: National Diversity+Team Work +National Diversity*Team Work=0	0.127	0.049**
Observations	324	324
Mean Dependant Variable,		
No National Diversity	0.675	0.121
R-squared	0.176	0.122

Notes: An observation is an individual worker. Robust standard errors in parentheses. Country pair and week fixed effects included in all regressions. Controls included in all regressions are team averages for member education, platform experience, non-platform work experience, number of platform tests, presence of a profile page, gender, wage bid, and agency membership. Individual Output is the total number of features added by an individual. Individual productivity is the total number of features added by an observation divided by the number of hours worked on the task by the individual. The H_o: National Diversity+Team Work+National Diversity*Team Work=0 row reports the p-values of this test. * significant at 10%; ** significant at 5%; *** significant at 1%

Table A10: Effect of National Diversity on Output and Productivity Among Teams

	(1) Team Output	(2) Team Productivity
National Diversity	-0.807***	-0.168*
	(0.290)	(0.090)
Observations	81	81
Mean Dependent Variable, No National Diversity	1.350	0.171
R-squared	0.470	0.392

Notes: An observation is a pair of workers. Robust standard errors in parentheses. Country pair and week fixed effects are included in all regressions. Controls are team averages for member education, platform experience, non-platform work experience, number of platform tests, presence of a profile page, gender, wage bid, and agency membership. Joint productivity is the total number of features added by an observation divided by the number of hours worked on the task by the pair. Sample is restricted to pairs of workers in the team treatment of the experiment.
* significant at 10%; ** significant at 5%; *** significant at 1%

Table A11: Effect of Geographic Distance & National Diversity on Pair Performance

	(1)	(2)	(3)	(4)
	Joint Output		Joint Productivity	
Team Work	0.392	0.144	0.152**	0.127**
	(0.420)	(0.456)	(0.058)	(0.061)
Log(Geographic Distance Between	0.035	0.009	0.020**	0.011
Pair Member Cities)	(0.043)	(0.055)	(0.009)	(0.009)
Log(Geographic Distance Between	-0.063	0.026	-0.029**	-0.016
Pair Member Cities)*Team Work	(0.065)	(0.091)	(0.012)	(0.013)
National Diversity*Team Work		-0.579		-0.122**
		(0.386)		(0.049)
H_o: Team Work+National Diversity*		0.136		0.002***
Team Work=0				
Observations			119	
Mean Dependant Variable,				
Independent Work		0.983	0.126	
R-squared	0.253	0.291	0.280	0.358

Notes: An observation is a pair of workers. Robust standard errors in parentheses. Country pair and week fixed effects are included in all regressions. Controls are team averages for member education, platform experience, non-platform work experience, number of platform tests, presence of a profile page, gender, wage bid, and agency membership. Sample includes all pairs for which city of residence was available for both contractors. Joint productivity is the total number of features added by an observation divided by the number of hours worked on the task by the pair. The H_o: Team Work+National Diversity*Team Work=0 row reports the p-values of this test. * significant at 10%; ** significant at 5%; *** significant at 1%

Table A12: Effect of Team Work & National Diversity on Performance by Pair Skill Differences, Interaction

	(1) Joint Output	(2) Joint Productivity
Team Work	0.186	0.034
	(0.225)	(0.032)
Different Skill Sets	-0.450*	-0.038
	(0.252)	(0.043)
National Diversity*Team Work	-0.229	-0.104**
	(0.345)	(0.052)
National Diversity * Different Skill Sets	0.543	0.014
	(0.373)	(0.072)
Team Work * Different Skill Sets	0.263	0.068
	(0.330)	(0.064)
National Diversity * Team Work	-0.737	-0.124
* Different Skill Sets	(0.476)	(0.090)
H_o: Team Work+National Diversity*Team Work=0	0.627	0.052*
Observations	162	162
Mean Dependant Variable Independent Work	1.062	0.135
R-squared	0.289	0.292

Notes: An observation is a pair of workers. Robust standard errors in parentheses. Country pair and week fixed effects are included. Controls included in regression are team averages for member education, platform experience, non-platform work experience, number of platform tests, presence of a profile page, gender, wage bid, and agency membership. The H_o: Team Work+National Diversity*Team Work=0 row reports the p-values of this test. * significant at 10%; ** significant at 5%; *** significant at 1%

Appendix B   Data Appendix[61]

Below is the document sent to contractors hired for the job used in this experiment.

I would like some customizations made to DokuWiki, an open source PHP-based wiki engine. DokuWiki uses plan text files so it does not need a database. The site is internal, not available to the public internet so I cannot share the URL with you. For more information on DokuWiki, see http://en.wikipedia.org/wiki/DokuWiki and https://www.dokuwiki.org/features.

The task is as follows: Please add as many of the below *Javascript/PHP* features in the attached code as possible and submit the code with the added features as soon as you have added everything you are able to. One feature is*Javascript/PHP* only and one uses both *Javascript/PHP* and *PHP/Javascript* (you are to work on the *Javascript/PHP* part of this task). You will be paid for eight hours of work on this project and all eight hours of work must be performed on "Day of the week, Month Date". Another contractor from "country teammate is from" has been hired to work on the *PHP/Javascript* features in this code. **You and this contractor will be working on this in the team room at the same time so please communicate with each other to work through this task together/You will work independently of this contractor**. Unfortunately I am the hiring manager and I have little knowledge of the technical aspects of the task. Therefore, I am not available to answer questions so just do the best you can. Please send me your output and let me know which features you were able to add once the eight hours is up. Please also update your memo to let me know what you are working on. Thanks again!

You can login with the username "admin" and the password "asdf".

You'll likely need to fix the permissions on the data/ directory so that the web server can write to them (probably chwon -R www-data data/, where www-data is the user which your webserver runs as).

PHP Task

Login using either username or email address

Currently users must use their username to login. Allow them to use either their username or their email address to login.

For example, the "admin" user has the email address "admin@example.com" and the password "asdf". Allow the "admin" user to login either by entering the username "admin" and the password "asdf" OR by entering the email address "admin@example.com" and the password "asdf".

Javascript Task

Make a popup for the login dialog.

Currently clicking "login" directs the user to a new page. Update this link so that, when clicked, a popup containing the login form is shown (similar to, for example, http://www.meetup.com/).

Combined Javascript & PHP Task

Show when a page is being edited.

Make the "edit page" text red when the current page is being edited.

---

[61]Italicized words indicate content that varies by the type of coding language the contractor was hired to complete. Bolded words indicate content that varies by whether contractors are in the team work treatment or not.

Use AJAX to poll the server every 15 seconds checking to see whether the current page is being edited. If it is, change the color of the "edit page" text to red. Change it back to the original blue when the page is no longer being edited.

For example, if Alice and Bob both have the start page open, then Alice clicks "edit this page", the color of the "edit page" text in Bob's web browser should change to red. Once Alice finished editing the page (either by cancelling the edit or saving the new page) the color of the "edit page" text should return to the original blue.

Appendix B.1   Supplementary Job Posting Information

This section describes additional information about the job posting information beyond what was described in section 2. Other than the title of the job and the job description, the information available to contractors on job postings is standardized by oDesk. In particular, oDesk posts employer information on all job postings to applicants can see how many contracts employers have hired for on the site, how much they have spent and feedback from previous hires. In addition, all job postings have to specify the estimated time the contract will last and the approximate number of hours per week the job will require. Both of these measures must be selected by the employer from a list of pre-specified options. Screenshots of the Javascript and PHP job postings used for the experiment are provided in the appendix of this paper.

An important requirement of oDesk job postings for the purposes of this experiment is that employers must specify at the time of posting which team room contractors hired for the job will work in. Once hired, contractors cannot be removed from this team room. Therefore, to ensure that contractors in the team work treatment were only able to communicate about this task with their teammate and that those in the independent work treatment were not able to communicate with any other contractors about the task, one job posting for each participant was required. This institutional feature of the site made it necessary to determine which jobs would allow team work among pairs of contractors and which would not before contractors were able to apply for them. However, job postings do not indicate which team room the hired contractor will work in, and the job postings do not differ by treatment group so applicants cannot have known in advance of being hired whether they would be working with a teammate or not.

Below are screenshots of the Javascript job posting and the PHP job posting. The employer work history and feedback is blocked out to protect the privacy of contractors on the site.

Figure B1: Javascript Job Posting

Javascript Features

Hourly - Est. Time Less than 1 week, As needed - Less than 10 hrs/week - Posted 2 hours ago

Job Description

I'd like to hire contractors to add some Javascript customizations to DokuWiki, an open source PHP-based wiki engine, within 8 hours. The work will have to be performed Saturday, January 5, please don't apply if you cannot work on the 5th. No PHP knowledge is needed for this job.

Skills Required

javascript

Preferred Qualifications		**Client Activity on this Job**	
Hourly Rate:	$1.00/hr - $4.00/hr	Last Viewed: ⑦ 2 hours ago	
		Applicants:	9
		Interviewing:	0

⊕ **Applicants (9)**

Client's Work History and Feedback (50)

Job Overview

Type:	Hourly
Workload:	As needed - Less than 10 hrs/week
Duration:	Less than 1 week
Posted:	January 4, 2013
Planned Start:	January 5, 2013
Visibility:	Public
Category:	Web Development
Sub-Category:	Web Programming

About the Client

Canada (UTC-06)

Member Since February 27, 2011

Payment Method Verified ⑦

★★★★★ (4.96) 34 reviews

Total Spent:	$984
Hours Billed:	227
Jobs Posted:	158
Paid Contracts:	52
Open Jobs:	2
Active Contracts:	44

Figure B2: PHP Job Posting

PHP Features

Hourly - Est. Time Less than 1 week, As needed - Less than 10 hrs/week - Posted 2 hours ago

Job Description

I'd like to hire contractors to add some PHP customizations to DokuWiki, an open source PHP-based wiki engine, within 8 hours. The work will have to be performed Saturday, January 5; please don't apply if you cannot work on the 5th.

Skills Required

php

Preferred Qualifications	**Client Activity on this Job**	
Hourly Rate: $1.00/hr - $4.00/hr	Last Viewed: ?	2 hours ago
	Applicants:	15
	Interviewing:	1

⊕ **Applicants (15)**

Client's Work History and Feedback (50)

Job Overview

Type:	Hourly
Workload:	As needed - Less than 10 hrs/week
Duration:	Less than 1 week
Posted:	January 4, 2013
Planned Start:	January 5, 2013
Visibility:	Public
Category:	Web Development
Sub-Category:	Web Programming

About the Client

Canada (UTC-06)

Member Since February 27, 2011

Payment Method Verified ?

★★★★★ (4.96) 34 reviews

Total Spent:	$984
Hours Billed:	227
Jobs Posted:	158
Paid Contracts:	52
Open Jobs:	2
Active Contracts:	44

58

Appendix B.2   Supplementary Interview Information

Below is the text provided to interviewees for the Javascript job.

Hello, thanks for applying to my job. I have four interview questions for you to answer. Please answer as honestly as possible as we are looking to hire the person best suited to this job. Please answer all questions in an oDesk message; I'm not available to communicate over Skype.

1) If you have any Javascript experience, please give up to 5 examples of your experience.
2) PHP knowledge is not needed for this job but if you do have any PHP experience, please give up to 5 examples of your experience.
3) Please list all the countries you have lived and/or worked in.
4) In one paragraph, please describe why you think you are well suited for this job.
Also, please confirm whether you are able to work your hours on this job on *(Date job is to be completed on)*.
Below is the text provided to interviewees for the PHP job.

Hello, thanks for applying to my job. I have four interview questions for you to answer. Please answer as honestly as possible as we are looking to hire the person best suited to this job. Please answer all questions in an oDesk message; I'm not available to communicate over Skype.

1) Javascript knowledge is not needed for this job but if you do have any Javascript experience, please give up to 5 examples of your experience.
2) If you have any PHP experience, please give up to 5 examples of your experience.
3) Please list all the countries you have lived and/or worked in.
4) In one paragraph, please describe why you think you are well suited for this job.
Also, please confirm whether you are able to work your hours on this job on *(Date job is to be completed on)*.