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

Elizabeth Lyons†

April, 2015

Abstract

Teams of workers are increasingly diverse in their nationality and skill sets. This paper studies the effect of diversity on team performance 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 different skills. Further investigation of the data suggests that teamwork benefits nationally homogeneous teams by increasing the returns to effort. In contrast, teamwork in diverse teams reduces the returns to effort.

JEL Classification: D40, D83, F23, J24, M54

*I thank Ajay Agrawal, Avi Goldfarb, Nicola Lacetera, and András Tilcsik for their valuable input and guidance. I also benefited from the comments of Chris Forman, Alberto Galasso, Joshua Gans, Josh Graff-Zivin, Mitch Hoffman, John Horton, Sarah Kaplan, Bill Kerr, Bill McEvily, Anita McGahan, Amanda Pallais, Mike Ryall, Brad Staats, as well as participants at the 2013 CCC Doctoral Colloquium and 2013 REER, and seminar participants at Cornell, Harvard Business School, Johns Hopkins, London Business School, Toronto, USC, UC San Diego, and Washington University in St. Louis. I gratefully acknowledge funding support from the Centre for Innovation and Entrepreneurship at the Rotman School of Management, the Martin Prosperity Institute, the Rotman School of Management doctoral program, and the Social Sciences and Humanities Research Council of Canada grants #493140, #494787, and #486944. This research was approved by the University of Toronto Social Sciences and Humanities Research Ethics Board, protocol reference #28221.

†IR/PS, University of California, San Diego, 9500 Gilman Drive, MC 0519 La Jolla, CA 92093-0519. lizlyons@ucsd.edu.

1 Introduction

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 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, 2009; 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, 2013; Lazear, 1999). 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 team performance, the direction and extent of the effect of national diversity on team performance remains ambiguous (e.g. 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.

A framework developed in section 2 of this paper develops several propositions that predict what the effect of teamwork will be as a function of communication difficulties and gains from complementarities. Specifically, the framework predicts that, without any differential benefits to complementarities across types of teams, (1) total group effort invested in production and output will be *higher* with teams than independent work when communication requires a small proportion of effort and vice versa when communication requires a large proportion of effort, and (2) the costs of even very difficult communication can be offset by sufficiently high benefits from complementarities. If communication is more difficult in nationally diverse teams than in nation-

[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. International migration motivated by labor force opportunities is so common that, when combined, the population of these migrants is at least equal to the population of Brazil (International Labor Organization, 2012b).

[2] MNE global investment increased by more than four times during the 1990s, and 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 the research on organizational diversity, see Harrison and Klein (2007).

ally homogeneous teams then, without differential benefits to complementarities, the framework predicts that effort invested in production and output will be higher in nationally homogeneous teams than in nationally diverse teams and that effort invested in production and output are more likely to be higher with teamwork than with independent work when groups are nationally homogeneous. With differential benefits to complementarities between worker effort across team types, for instance because workers from different countries have access to different types of information or skills, the difference between nationally homogeneous and nationally diverse pairs becomes ambiguous.

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, to test the role of national differences on work team success. I identify the causal impact of organizing workers into teams by comparing 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 3, 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.[5] 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.[6] 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 5, 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

[5] In the experiment, I hold language constant. 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 in an attempt to reduce the effect of language diversity on my results.

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

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 perhaps because they give workers a shared context through which they are able to understand each other or because they allow workers to avoid costly communication altogether. Further investigation of the data suggests that communication is more difficult for nationally diverse teams than it is for nationally homogeneous teams, and that national diversity reduces the returns to investing effort in the task.

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. Moreover, my findings suggest that coordination problems that arise due to difficulties with communication may be difficult to overcome if they are not addressed ex ante, because they lead to lower effort investments in the team task.

This paper proceeds as follows. Section 2 develops a motivating framework; section 3 describes the experiment design; section 4 describes the data collected from the experiment and my estimation strategy; section 5 presents the empirical results, and section 6 summarizes and concludes the paper.

2 Conceptual Framework

The intention of this framework is to give some formal structure to what the existing literature predicts about the benefits and costs of teamwork and how this varies with team diversity. It is a simplified version of some of the key findings in prior research that are relevant for the experiment reported in this paper. In particular, the framework allows for performance benefits to teamwork, and for differential communication costs and benefits to complementarities for different types of teams. The production function reflects the type of task developed for the experiment presented here but is intended to be general enough to be useful in other settings as well. The proofs for the propositions below are presented in Appendix B.

2.1 Model Set-up

2.1.1 Production

A team is composed of two workers, $i = 1, 2$. Teams can be made up of two workers from the same country or two workers from different countries. Workers are assigned to complete a task within a finite amount of time and can choose their own effort levels, e_i. Production is a function of both workers' efforts and allows for complementarities between teammate efforts to reflect the idea that, for instance, they can share knowledge. The production function is $y = \lambda e_1 + \lambda e_2 + \delta(\lambda e_1 \lambda e_2)$, where $\lambda \varepsilon [0, 1]$ is the fraction of effort that goes towards production, $1 - \lambda$ is the fraction of effort that goes towards communication, and $\delta \geq 1$ if a pair of workers works as a team and 0 if they work independently. When no effort goes into communication, λ is equal to 1. Given the prior evidence that communication can be more difficult when team members are diverse, (Bengtsson and Hsu, 2013; Hegde and Tumlinson, 2014; Lazear, 1999), λ may be smaller for nationally diverse teams. Similarly, δ may be larger for nationally diverse teams because the gains from complementarities may be larger for diverse teams, for instance because they have access to different types of knowledge (Guzzo and Dickson, 1996; Lazear, 1999; Stahl et al., 2009; Watson et al., 1993).

In this set up, team work is exogenously assigned, and workers cannot decide whether or not to work with each other. In reality, the decision of whether or not to work with a colleague may be made with the costs of communication and benefits to collaboration taken into account. The exogeneity of the team work assignment reflects the empirical set up in this paper. First, workers are randomly assigned to teams or independent work so they do not get to select into team work or independent work. Second, although workers in the team treatment have the option to ignore each other, the uncertainty with respect to how inputs should be coordinated may make eliminating communication altogether difficult. This necessary communication due to uncertainty is not unique to this paper. In particular, managers face a trade-off when deciding how to coordinate employees working on a joint goal. As Dessein and Santos (2006) point out, in environments with little uncertainty about how coordination should occur, it might be optimal to specify which actions workers should take and how they should be coordinated ex ante. In less certain environments, it may be optimal to allow workers the flexbility to respond to their local information and decide how to coordinate their inputs ex post through communication. In situations where managers have not specified how worker inputs should be coordinated, workers may not be able to avoid communication if they are going to complete the job. In my experiment, workers face this type of uncertainty because the instructions they received did not specify how the coordination within the

team should occur.

2.1.2 Contractor Utility

Each worker incurs a private cost of effort equal to $c(e_i)$ which is not dependent on λ, with $c'(e_i) > 0$ and $c''(e_i) > 0$. For simplicity, assume that $c(e_i) = \frac{1}{2}e_i^2$. The convexity of cost reflects the increasing disutility of effort. Workers receive a payment p for their work, where p is equal to a fraction π of total production. This payment can also be interpreted as a reputational boost to reflect the incentives that workers face in the experiment.[7] For simplicity, π equals to 1. This assumption should not change the predictions of the model as long as utility is a monotonic function. Therefore, worker i's utility is $U_i = \pi(\lambda e_i + \lambda e_j + \delta\lambda e_i \lambda e_j) - c(e_i) = \lambda e_i + \lambda e_j + \delta\lambda e_i \lambda e_j - \frac{1}{2}e_i^2$.

2.2 Example with Predictions

2.2.1 Independent Work

If working independently, worker i does not have to invest in communicating with worker j. Therefore, worker i chooses effort to maximize $U_i = e_i + e_j - c(e_i)$. Thus, with the first order condition $1 - c'(e_i) = 0$, independent worker optimal effort is set such that $c'(e_i^{IW}) = 1 = e_i^{IW}$.

2.2.2 Team Work

If working in teams, worker i chooses effort that solves the first order condition $\lambda + \delta\lambda^2 e_j - c'(e_i) = 0$. Because of the symmetry of worker utility functions, at the optimal solution $e_i = e_j$. Therefore, $e_i = \frac{\lambda}{1-\delta\lambda^2}$, $e_j = \frac{\lambda}{1-\delta\lambda^2}$, and optimal effort for workers in teams is $e^{TW,ext} = \frac{\lambda}{1-\delta\lambda^2}$.

Therefore:

Proposition 1 *The effort of workers in teams is increasing in λ.*

Proposition 2 *Conditional on a sufficiently large marginal cost of effort, for a given δ there exists a value λ^{TW} such that for all $\lambda > \lambda^{TW}$, effort is higher with teamwork than with independent work. For $\lambda < \lambda^{TW}$, effort is lower with teamwork than with independent work.*

Production is increasing in effort, so Propositions 1 and 2 predict that output will be higher on average for teams made up of workers who do not need to invest a significant amount of effort into communication than those made up of workers who do. In addition, output will be higher

[7] On oDesk and online contract labor markets more generally, contractor reputations are important for success in the market (Horton et al., 2011; Pallais, 2012).

when work is done in teams than when it is done independently if workers do not have to invest a large amount of effort into communication. However, if the the share of effort that goes towards communication is large enough, output will be lower with teamwork than with independent work.

Proposition 3 *The effort of workers in a team is increasing in* δ.

This suggests that in team work, returns to effort are higher when complementarities are higher.

Proposition 4 *Output and effort losses due to communication that requires a high share of effort can be offset by high complementarities.*

2.3 Discussion

This model predicts that teamwork is most beneficial when the majority of worker effort can be invested in production rather than communication and when there are strong complementarities between teammates' effort.[8] Given prior findings that diversity in teams can increase costs of communication (Bengtsson and Hsu, 2013; Hegde and Tumlinson, 2014; Lazear, 1999), it is likely that communication requires more effort in nationally diverse teams than in nationally homogeneous teams. If this is the case, this leads to the prediction that teamwork is more likely to be beneficial when teammates are from the same country and that independent work is more likely to lead to higher output when teammates are from different countries. Finally, given prior findings that diverse teams are better at problem solving than homogeneous teams are because they approach problems from more than one angle and have access to different types of information (Guzzo and Dickson, 1996; Lazear, 1999; Stahl et al., 2009; Watson et al., 1993), it is possible that nationally diverse teams benefit more from complementarities between teammate efforts. This leads to the prediction that for high enough complementarities, teamwork can lead to higher output than independent work even when teammates are from different countries and face difficult communication. In terms of optimal organizational design, these predictions suggest that teamwork is preferred to independent work when workers are from the same country and communication is quite easy, and that independent work is preferred to teamwork when workers are from different countries and communication is quite hard unless the knowledge complementarities between workers can compensate for the communication difficulties.

[8]For this not to hold, marginal cost has to be quite close to zero which is unlikely to be the case for workers with outside labor options and outside leisure options.

2.4 Further Considerations

Aside from national diversity, there may be other reasons communication costs and complementarities differ across teams. Lazear (1999) points out that differences in skill sets or knowledge between team members increases the benefits of working in teams. 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. These possibilities are tested empirically in section 5.1.2.

In addition to not addressing the potential effects of skill specialization on team production, the model presented here does not address the mechanisms driving different effort levels due to communication costs. Communication costs may be higher in nationally diverse teams for instance because of language differences (Lazear, 1999), lack of trust which can lead to conflict (De Dreu and Weingart, 2003), 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 by workers from different countries (Dahlin et al., 2005). In the experiment presented in this study, workers are all able to communicate in English, so the first explanation may not be relevant here although even when speaking the same language, linguistic differences between nationally diverse workers may persist (Kachru, 1992). All of these possibilities, as well as the possibility of other mechanisms, may be driving communication costs in nationally diverse teams. Testing for which is present in the empirical setting studied in this paper is beyond the scope of this work, but an important area for future research.

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, 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 even when communication is easy, workers prefer to invest less effort in a task that requires them to work with someone from a different country. Workers choose 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. This explanation is tested empirically in section 5.2.2. Suggestive evidence indicates animosity is not driving differences in performance across teams and individual workers in this setting.

3 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.

3.1 Research Setting

I generate data for this analysis through a natural field experiment conducted using oDesk, an international online market for contract labor. oDesk is the largest contract labor platform in the world in terms of earnings, and the fastest-growing.[9] 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 operates as a mediator and matches employers with contract laborers who 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

[9]The total amount spent on the site in the six years after the company was founded in 2004 was less than $25 million. By 2012, the amount spent per *quarter* was almost $100 million.

profile pages when they join the site. These pages can include information on their education, prior work histories,[10] and scores from tests administered by oDesk. Profile pages also provide information on where contractors live.[11]

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. Employers on oDesk frequently require that multiple hires work in the same team room.[12] Second, the team room facilitates contractor monitoring by employers through frequent screen shots, memos, work diaries, and activity meters. 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.

3.2 The Task

To test my research question, I needed contractors to be assigned a task for which collaboration is not unnatural. In addition, because I am specifically 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.[13] Moreover, an objective evaluation of the work contractors did on the task 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

[10] 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.

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

[12] During the past eight years 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 post more jobs and more knowledge-intensive jobs online, the prevalence of work teams in the international online market for contract labor will likely continue to grow.

[13] 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., Harrison and List, 2004; Klebe Trevino, 1992; Levitt and List, 2007).

another to complete the PHP portion.[14] 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. 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 C. Nothing in the instructions or the task design specifically 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 forces working against a single contractor working 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.

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.[15]

3.3 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 cross-country teams in order to identify the

[14] 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.

[15] Contractors had between midnight UTC one day until midnight EST the next. In the first round of this experiment's pilot, I required contractors to work all their hours within the same eight hour period as the other contractor in their pairing. Because of other commitments, this was very difficult for many contractors, and many of them chose to end their contracts before working on the task. In addition, contractors did not seem to think it was a legitimate requirement, which may have resulted in some of them becoming suspicious of the purpose of the task. In contrast, 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 my time zone was EST (this information was posted on my employer profile), information gathered in the second round of the pilot suggested that they seemed to expect the deadline to be midnight EST rather than midnight UTC. To minimize confusion, I decided to keep the deadline at midnight EST.

effect of multinational teamwork on performance.

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, and Bangladesh in the experiment. Contractors from these countries frequently applied to my job postings and bid an amount that was within the hourly wage criteria. 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.[16] These differences are discussed in more detail in section 5.

3.4 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. 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[17] and the maximum bid that would be accepted ($4.00 USD).[18] 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 C. Applicants chose which job to apply for, either Javascript or PHP, however, some applicants are knowledgeable in both languages. It 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 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,

[16] 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. Specifically, for the power distance cultural dimension, Bangladesh has a score of 80, Pakistan 55, and India 77. In the individualism dimension, Bangladesh has a score of 20, Pakistan 14, and India 48. In uncertainty avoidance, Bangladesh scores 60, Pakistan 70, and India 40.

[17] 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. In contrast, fixed price contracts do not guarantee payment so there is less incentive to perform work while logged into the team room.

[18] 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 (on a $4.44 bid, $4.00/hour would go to the contractor and $0.44 would go to oDesk).

their English language abilities, and the countries they have lived in. The interview text sent to Javascript and PHP job interviewees is in Appendix C. 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.[19] Some contractors applied for both jobs, but were hired for the task in which they met the hiring requirements.

3.5 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.[20]

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 questions 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 the survey, their contracts ended. I provided feedback based on contractors' output to all paid contractors.

After the experiment had been completed, I sent all participants a debriefing message to inform them that the job they had been hired for was for research purposes and to describe the goal of the research.[21]

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

[20] I provide the four different versions of the instructions in Appendix C.

[21] This message also explained that their identities would be protected and that if they did not feel comfortable being included in the analysis, they would be dropped as an observation. None of the contractors in my sample asked to be removed from the study.

4 Data and Estimation Methodology

4.1 Overview of the Data

I collected data between January 2013 and June 2013. In total, I hired 324 contractors, who made up 162 pairs.[22] There are 80 contractors and 40 pairs in the nationally homogeneous teamwork and independent work treatment groups,[23] and 82 contractors and 41 pairs in the nationally diverse teamwork and independent work treatment groups.[24][25]

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 interviews 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. I also recorded the total number of non-overlapping errors added by each team; so, for instance, if both teammates added the same feature, this was only counted as once for the team level measure of added features.[26] 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 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

[22] In 12 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).

[23] As discussed in Section 3, 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.

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

[25] I based the number of observations on a sample size analysis done with data collected during the pilot phase of this study, which predicted that I would need 40 observations per group to obtain statistically significant differences between the groups at the 5% level with a power of 80%.

[26] This overlap could occur if, for instance, a contractor hired for the PHP portion of the task is also versed in Javascript and decides to complete more than just the PHP portion of the task.

had been hired once or twice on the platform),[27] 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% 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 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.[28]

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 and, within the teamwork treatment, nationally diverse teams and non-nationally diverse teams. These averages show that teamwork significantly improves outcomes when team members are from the same country and significantly reduces output when team members 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.

[27] It is not very surprising that the contractors who participated in the experiment were inexperienced on the platform given the relatively low wage I offered for a job that requires a non-negligible amount of skill. Contractors with more experience on the site are likely able to capture a higher wage for this type of task (Agrawal et al., 2013b).

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

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-value 0.01) relative to nationally diverse independent contractor pairs.

These comparisons are consistent with the propositions made in Section 2. Specifically, they suggest that when a team is nationally homogeneous, communication is easy enough that teamwork outperforms independent work and that communication 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.

4.2 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,[29] 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

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

within a given pair of countries with the average performance of teams made up of contractors in the same pair of countries.

5 Results

I begin by presenting baseline 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 then present results on possible explanations for these main findings. In particular, I explore the possibility that effort decisions are affected by team work and national diversity, and that workers differentially dislike teammates who are from other countries.

5.1 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 country pair and week 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.

The coefficient estimates in Column 3 show that teamwork is beneficial for nationally homogeneous teams and harmful for nationally diverse teams. Specifically, 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. Specifically, 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.

5.1.1 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.[30] I estimate Equation 1 with individual output and individual productivity as the dependent variables and present the results in Table A6. 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 A6 relative to the sample mean performance are comparable 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.

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. These results are reported in Table A7. 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.

5.1.2 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. Column 3 presents the estimated coefficient of the effect of teamwork for nationally homogeneous and nationally diverse pairs of

[30] 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.

contractors on the likelihood of implementing the feature that requires both Javascript and PHP. As with the implementation of the Javascript feature, 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. Given that the features were to be added to a PHP-based script, these results suggest that teamwork has a larger effect on performance when two types of knowledge are required and, therefore, when coordination is more important. The combined feature required both PHP and Javascript to be implemented, and, while the writing of the Javascript feature only required Javascript knowledge, contractor pairs may have benefited from knowledge of PHP when determining where in the script to implement the feature.

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 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. 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 12% level.[31] Similar findings using interactions between the skill differences, national diversity, and team work variables are presented in Table A8.

5.2 Interpretation

In this subsection, I consider two possible explanations for my findings. Consistent with the predictions of the model presented in section 2, I explore the data to determine whether differences in effort due to communication costs can explain the output and productivity differences. I also test for evidence of dislike for national diversity as a mechanism for my results. I find suggestive evidence in support of the former explanation, and no evidence in support of the latter.

[31] 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).

5.2.1 Communication & Effort

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, 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).[32]

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.[33]

To further investigate whether communication costs differ for nationally homogeneous and nationally diverse teams, I perform two additional tests. First, 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.[34] 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. Second, I estimate Equation 1 with an indicator for whether or not contractors in the same team worked on implementing the same feature. This is a proxy for whether or not contractors were able to divide up the work. Both of

[32] 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 [Bengladeshi political culture] with its emphasis on extreme individualism" a result of "...frontier conditions [of Bengal households] under the British...compounded by 1971 guerilla war." 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 empirical evidence that these differences show up in labor markets. For instance, India and Pakistan differ significantly in their human resources management practices - Pakistan is more similar to China than to India on most dimensions of HRM (Aycan et al., 2000). Measures of labor market cultural distance also show large differences between Bangladesh, India, and Pakistan on several dimensions (Hofstede, 1983).

[33] 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.

[34] 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 that there is differential 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 because if they took advice from their teammate and believed it solved a problem that it didn't, they may end up misreporting what they were able to complete.

these tests are done on the sample of teams.[35] I present the results of these estimations in Table 8. The results in Column 1 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. Column 2 does not suggest any significant differences across team types in whether contractors in a team worked on implementing different features. However, there may be benefits to working together on a feature.[36]

Taken together, findings on task completion, skill specialization, and misreporting are consistent with communication being more difficult for cross-country teams. The conceptual framework presented in Section 2 suggests that difficult communication harms performance by reducing the value of effort because a lower proportion of effort will be spent on production when more of it has to be spent on communication. To test whether this prediction is consistent with contractor behavior in the experiment, I estimate Equation 1 with indicators for whether or not a pair attempted to implement each feature and whether a pair invested effort in any of the features as measures of effort. I present these estimates in Table 9.

The results show that whether or not effort was invested in a task did not change significantly for work on the feature that required only Javascript knowledge or for the feature that required only PHP knowledge. However, effort does appear to have differed across the treatment groups for the feature that required both PHP and Javascript knowledge. This is the part of the task that likely requires the most collaboration to be completed so pairs that could communicate easily were more likely to benefit from teamwork when trying to complete it and pairs that had a hard time communicating were less likely to benefit from teamwork when trying to complete it. To make the latter point more concrete, when working independently, contractors can invest all their effort into production.[37] When working in teams, pairs with difficulty communicating may realize that a large proportion of the effort exerted on the joint language feature (i.e., the third feature) will go towards coordinating with their teammates rather than into production. As a result, these pairs may not be willing to invest any effort in trying to complete it if the gains from joint production do not outweigh the costs. Consistent with this interpretation, effort on the joint language feature appears to increase for nationally homogeneous pairs when they work in teams relative to when they work independently, and the reverse appears to be true for nationally diverse teams. In addition

[35] The sample is restricted to teams because independent contractors cannot communicate with each other.

[36] Ideally, I would be able to test for communication differences directly by analyzing text data collected through the screenshots taken while contractors work on oDesk. Unfortunately, screenshots are only taken every 10 minutes, and contractors receive a warning before a screenshot is taken so they can decide whether or not to minimize their conversations before the screenshot is taken. As a result of this, and due to the fact that conversations seem to be concentrated within specific time periods, communication data from screenshots is sparse and complete conversations are rarely observable.

[37] Contractors in the independent treatment who only know how to code in one of the two languages may not see a benefit of investing effort in the third feature, however, those who know both languages will have been able to complete the feature.

to being statistically significant, the results are economically significant. Teamwork in nationally homogeneous teams increases the likelihood that a pair of contractors exerts effort on the third feature by more than 17% and decreases effort on the third feature for nationally diverse teams by more than 26%.

One alternative explanation for the drop in effort on the combined feature for nationally diverse teams is that these pairs were less productive (as demonstrated in Table 5), and if pairs attempted the combined feature last, nationally diverse teams were less likely to have time to work on this feature even if they intended to. The coefficient estimates in Column 4 suggest that this productivity difference is not the only factor contributing to the fall in effort for nationally diverse teams. The estimates suggest that nationally diverse teams of contractors are more likely to invest no effort into any part of the task than nationally diverse independent pairs of contractors or than nationally homogeneous teams or independent pairs. This finding holds when any of the country pairs are dropped, so it is not being driven by a particular pair of countries.

As an additional test of the prediction that the returns to effort are higher in teams when communication is sufficiently easy, I investigate how effort changes over time across my treatment groups. To do this, for each team, I code the average effort level for every ten minute period across the total amount of time the 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 for independent contractors, contractors in nationally homogeneous teams, and those in nationally diverse teams. The figure demonstrates three patterns worth noting. First, pairs who work independently exert more effort on average throughout the time they spend on the task.[38] This is consistent with theory that coordination 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.[39] Second, 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. Third, effort in nationally diverse teams falls over time and becomes significantly lower than average effort in nationally homogeneous teams once about 50% of the total time spent on the job has been completed. Consistent with the prediction that difficult communication reduces the returns to effort, as contractors in cross-country teams learn about the returns to their efforts, they invest

[38]There is no difference across treatments in the total amount of time spent on the task.

[39]Coordination likely requires contractors to wait for their teammates to show up, and to reply to messages sent through the team room.

less. The same is not true in nationally homogeneous teams who are significantly more likely than independent contractors and nationally diverse teams to increase their effort over the course of the job (see Table A5).

5.2.2 Animosity in Nationally Diverse Teams

Based on the structure of the task and of contractor incentives on oDesk, the theoretical framework presented in Section 2 assumes 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' payoffs change based on preferences for teammate nationality unrelated to communication difficulties and complementarities because of negative feelings towards people from other countries (Alesina and La Ferrara, 2005). The average effort levels across treatments presented in Figure 4 suggest nationally diverse teams begin the task with aligned incentives,[40] 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.[41] 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. Table 10 presents results from regressions of the cross-country team indicator on these survey questions.[42] 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.

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.[43] To test whether this dispute is driving

[40] Contractors are told which country their pair-mate is from before they begin the task. As Figure 4 demonstrates, this knowledge alone does not appear to have an effect on effort levels.

[41] The survey response rate is 83%.

[42] I only gave these survey questions to contractors in the teamwork treatment, so the results are restricted to that sample.

[43] Hjort (2013) 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 11. The results in Table 11 are consistent with those in Table 5,[44] and are not consistent with the negative effect of teamwork on nationally diverse pair performance being driven by dislike for national diversity in teams.[45]

6 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 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, and that the returns to investing effort in the task are lower when nationally diverse groups of workers collaborate. I find no evidence that animosity between workers from different countries affects their performance.

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. Workers in longer term contracts who do not have as many outside options as oDesk contractors do may also have more of an incentive to continue

[44] The statistical significance of the coefficients in Table 11 is lower, but the sample size is also significantly lower, so that is not surprising.

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

investing high levels of effort throughout the course of the job. 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.[46] Finally, communication encompasses a number of different elements including interpretation, trust, and knowledge (Kintsch, 1988; Rode, 2010). Cultural differences may affect many aspects of communication, and 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 evidence in support of the theoretical explanation that my findings are driven by changes in the returns to effort and the costs of communication. 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). 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; Leslie and Lindley, 2003; Oreopoulos, 2011). 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.[47] 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

[46] 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.

[47] 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).

providing cultural training programs for international market participants.

7 References

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.

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.

Bengtsson, O. and D. H. Hsu, "Ethnic Matching in the U.S. Venture Capital Market," *Management Department Working Paper*, 2013.

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

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.

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.

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.

Guzzo, R. A. and M. W. Dickson, "Teams in Organizations: Recent Research on Performance and Effectiveness," *Annual Review of Psychology*, 1996, *47*, 307–338.

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.

Harrison, G. W. and J. A List, "Field Experiments," *Journal of Economic Literature*, 2004, *42*, 1009–1055.

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

Hegde, Deepak and Justin Tumlinson, "Does Social Proximity Enhance Business Partnerships? Theory and Evidence from Ethnicity's Role in US Venture Capital," *Management Science*, 2014, *60* (9), 2355–2380.

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, J., "Ethnic Divisions and Production in Firms," Working Paper 2013.

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.

Horton, J. J., D. G. Rand, and R. J. Zeckhauser, "The online laboratory: Conducting experiments in a real labor market," *Experimental Economics*, 2011, *14* (3), 399–425.

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

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

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.

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.

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

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.

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

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.

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.

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.

Pallais, Amanda, "Inefficient Hiring in Entry-Level Labor Markets," Working Paper 2012.

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.

Stahl, G. K, M. L. Maznevski, A. Voigt, and K. Jonsen, "Unraveling the Effects of Cultural Diversity in Teams: A Meta-Analysis of Research on Multicultural Work Groups," *Journal of International Business Studies*, 2009, *41*, 690–709.

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.

Trevino, L. Klebe, "Experimental Approaches to Studying Ethical-Unethical Behavior in Organizations," *Business Ethics Quarterly*, 1992, *2*, 121–136.

Watson, W. E., K. Kumar, and L. K. Michaelsen, "Cultural Diversity's Impact on Interaction Process and Performance: Comparing Homogeneous and Diverse Task Groups," *Academy of Management Journal*, 1993, *36*, 590–602.

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

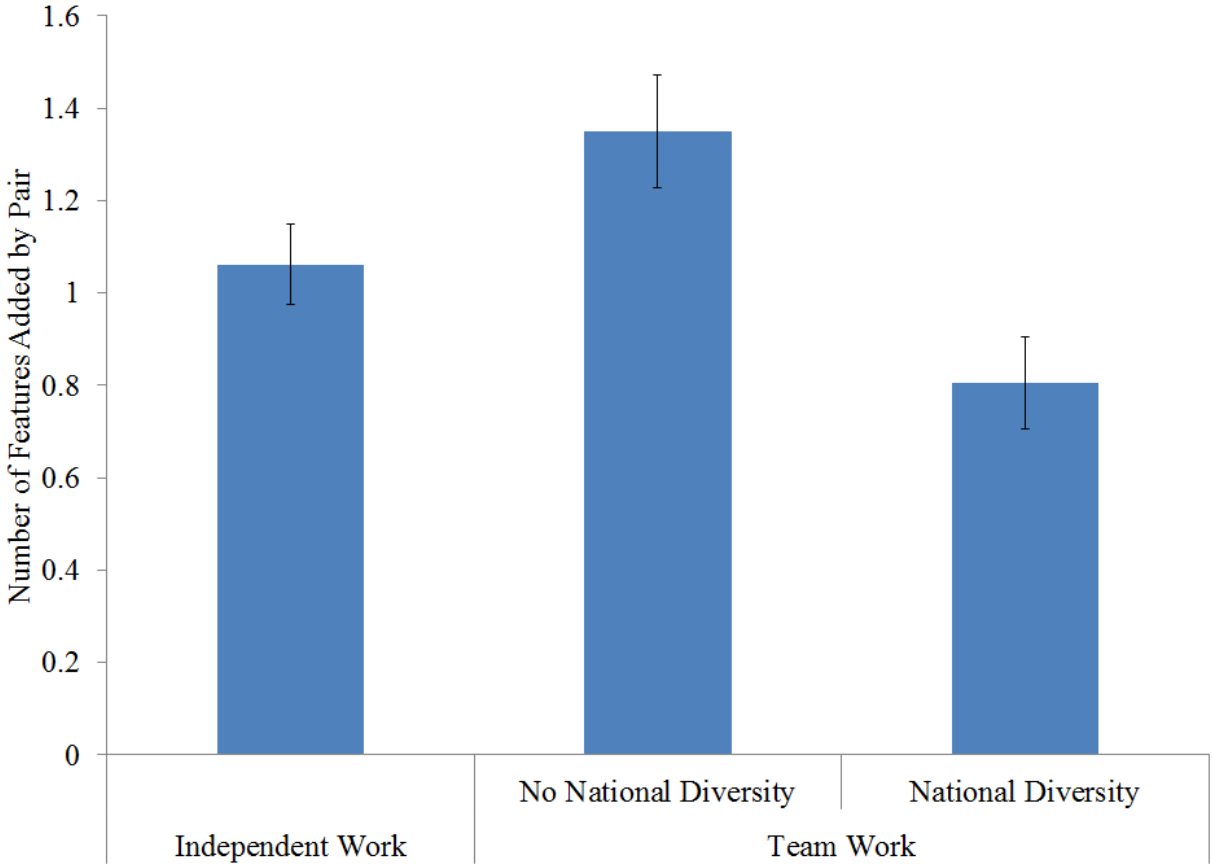
8 Tables and Figures

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)
	Shared	Team work by contractors of same nationality (N=40 pairs)	Team work by contractors of different nationalities (N=41 pairs)

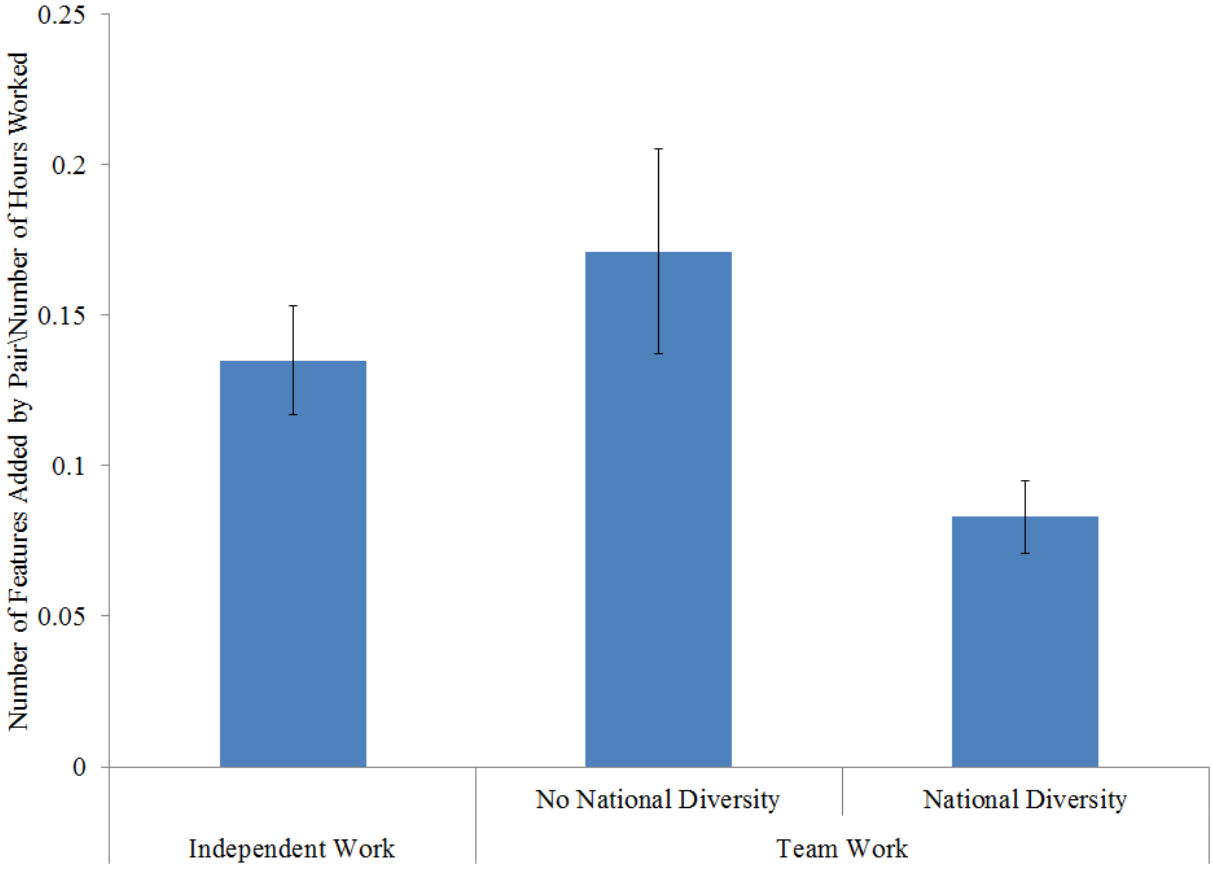
Notes: This matrix reports the four experimental conditions. The performance differences between the average in (1) and (2) and the average in (3) and (4) represent the overall effect of teamwork as compared to independent work. The performance differences between the average in (1) and (3) and the average in (2) and (4) represent the overall effect of having subjects of the same vs. different nationality in each pair. The difference between the performance of pairs in (1) and (2) measures the impact of same vs. different nationality for subject working independently and the difference between (3) and (4) for subjects working in teams.

Figure 2: Output by Treatment



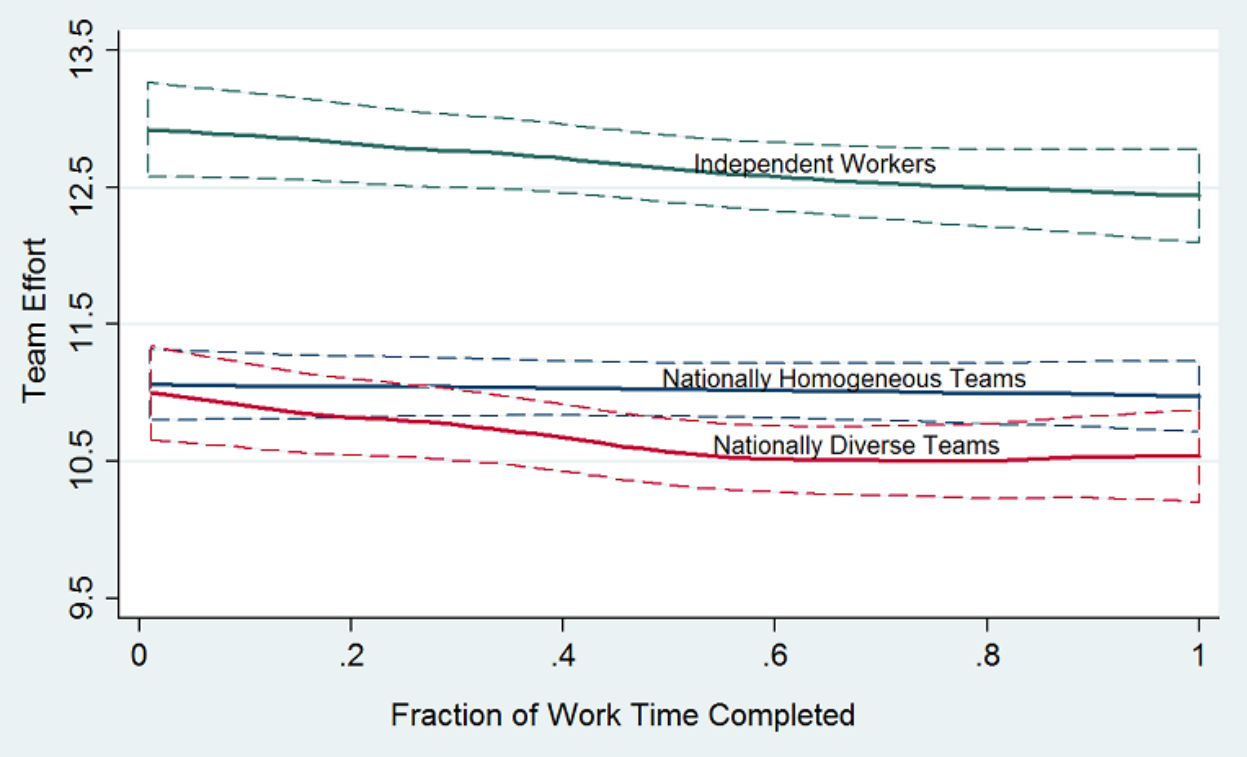
Notes: This chart compares mean output of pairs across treatment groups.

Figure 3: Productivity by Treatment



Notes: This chart compares mean productivity of pairs across treatment groups.

Figure 4: Effort Over Time by Treatment



Notes: This chart reports polynomials of average time effort levels over time across treatment groups. Dashed lines represent 90% confidence intervals.

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
Implement Combined Feature	Equal to one if Pair successfully adds the feature that requires both Javascript and PHP to the code, zero otherwise
Both Attempt Same Feature	Equal to one if both contractors in the pair attempt to add the same feature to the code, zero otherwise
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
No Attempt	Equal to one if both Contractors in Pair do not attempt to add any features 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 "6. 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)
Increasing Effort	Equal to one if effort during second half of job completion is higher than during first half, zero otherwise
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.
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)
N	162	

Table 3: Pair Outcome Summary Statistics

Pair Outcomes	Mean	(Std. Dev.)
Team Total Amount Paid	35.51	(20.408)
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	0.173	(0.379)
N	162	

Table 4: Pair Characteristic Summary Statistics Across Treatment Groups

	No National Diversity		National Diversity		
Variable	**No Team Work**	**Team Work**	**No Team Work**	**Team Work**	p-value^{+}
Number of Job Posting Applications	7.900	8.400	8.805	7.768	0.435
	(2.565)	(3.086)	(3.239)	(3.660)	
oDesk Rating Prior to Hire	4.423	4.657	4.631	4.589	0.645
	(0.919)	(0.520)	(0.448)	(0.650)	
No Rating Prior to Hire	0.625	0.613	0.646	0.561	0.760
	(0.386)	(0.381)	(0.373)	(0.337)	
Number of oDesk Contracts Prior to Hire	4.775	2.475	5.402	4.171	0.331
	(8.068)	(3.699)	(10.387)	(5.672)	
Indicator for having a Profile Picture	0.850	0.875	0.780	0.866	0.382
	(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 Listed on Profile	1.200	1.200	1.122	1.244	0.860
	(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)	
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

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

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
	Joint Output			Joint Productivity		
Team Work	0.350*	0.309*	0.300*	0.0773**	0.0681*	0.0698**
	(0.182)	(0.175)	(0.165)	(0.0373)	(0.0363)	(0.0331)
National Diversity*Team Work	-0.670***	-0.634**	-0.630***	-0.171***	-0.176***	-0.168***
	(0.237)	(0.250)	(0.238)	(0.0507)	(0.0533)	(0.0505)
National Diversity	0.125			0.0825**		
	(0.177)			(0.0358)		
Country Pair & Week Fixed Effects	No	Yes	Yes	No	Yes	Yes
Controls	No	No	Yes	No	No	Yes
Observations	162	162	162	162	162	162
Mean Dependant Variable Independent Work		1.062			0.135	
R-squared	0.066	0.191	0.255	0.069	0.225	0.274

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. * significant at 10%; ** significant at 5%; *** significant at 1%

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

VARIABLES	(1) Javascript Feature	(2) PHP Feature	(3) Combined Feature
Team Work	0.263**	-0.138	0.164***
	(0.109)	(0.105)	(0.059)
National Diversity*Team Work	-0.325**	-0.035	-0.277***
	(0.161)	(0.150)	(0.090)
Observations	162	162	162
Mean Dependent Variable Independent Work	0.568	0.444	0.074
R-squared	0.219	0.309	0.243

Notes: An observation is a pair of workers. Robust standard errors in parentheses. Country pair and week fixed effects 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 dependent variable is equal to one if the feature specified in the column headers was successfully implemented, and zero otherwise. * significant at 10%; ** significant at 5%; *** significant at 1%

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

	(1)	(2)	(3)	(4)
	Joint Output		Joint Productivity	
VARIABLES	No Skill Difference	Skill Difference	No Skill Difference	Skill Difference
Team Work	0.0670	0.436	0.0448	0.0923*
	(0.238)	(0.262)	(0.036)	(0.053)
National Diversity*Team Work	-0.242	-1.062***	-0.103*	-0.228***
	(0.386)	(0.328)	(0.0569)	(0.0848)
Observations	75	87	75	87
Mean Dependent Variable Independent Work	1.083	1.044	0.132	0.137
R-squared	0.534	0.437	0.566	0.310

Notes: An observation is a pair of workers. 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. A pair has a skill difference if the individuals in the pair have different knowledge of Javascript and/or PHP. * significant at 10%; ** significant at 5%; *** significant at 1%

Table 8: Effect of Team Work & National Diversity on Coordination

VARIABLES	(1) Difference between Reported and Completed	(2) Both Attempt Same Feature
National Diversity	0.349*	-0.026
	(0.091)	(0.096)
Observations	81	81
Mean Dependent Variable National Homogeneous Teams	0.241	0.177
R-squared	0.385	0.446

Notes: An observation is a pair of workers. 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. The sample is restricted to teams. * significant at 10%; ** significant at 5%; *** significant at 1%

Table 9: Effect of Team Work & National Diversity on Effort

VARIABLES	(1) Javascript Feature Attempt	(2) PHP Feature Attempt	(3) Combined Feature Attempt	(4) No Attempt
Team Work	0.037	-0.060	0.176*	-0.033
	(0.059)	(0.102)	(0.099)	(0.056)
National Diversity	-0.071	0.065	-0.268*	0.134*
*Team Work	(0.106)	(0.147)	(0.142)	(0.074)
Observations	162	162	162	162
Mean Dependent Variable, Independent Work	0.914	0.778	0.222	0.037
R-squared	0.148	0.121	0.272	0.186

Notes: An observation is a pair of workers. 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. Independent variables are dummies that equal one if at least one pair member attempted to work on the feature specified in the column headers. * significant at 10%; ** significant at 5%; *** significant at 1%

Table 10: Effect of National Diversity on Opinion of Teammate

VARIABLES	(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. 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. Independent variables are collected through a survey. * significant at 10%; ** significant at 5%; *** significant at 1%

Table 11: Effect of Team Work & National Diversity on Performance, Pakistan Excluded

VARIABLES	(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)
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. 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. Pairs with at least one worker from Pakistan are excluded. * significant at 10%; ** significant at 5%; *** significant at 1%

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 Applications	8.358	8.08	0.580
	(0.328)	(0.378)	
oDesk Rating Prior to Hire	4.527	4.62	0.470
	(0.11)	(0.083)	
No Rating Prior to Hire	0.636	0.586	0.397
	(0.042)	(0.040)	
Number of oDesk Contracts Prior to Hire	5.093	3.333	0.134
	(1.036)	(0.542)	
Indicator for having a Profile Picture	0.815	0.87	0.191
	(0.033)	(0.026)	
Level of Education	1.87	1.821	0.649
	(0.077)	(0.076)	
Number of Offline Jobs Listed on Profile	1.16	1.222	0.543
	(0.075)	(0.067)	
Average Score on oDesk Tests	3.416	3.44	0.736
	(0.044)	(0.058)	
Number of oDesk Tests Taken	2.531	2.488	0.905
	(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 Contractor	0.13	0.16	0.440
	(0.029)	(0.028)	
Indicator for Agency Membership	0.228	0.278	0.315
	(0.035)	(0.035)	
Number of Items in Portfolio	3.92	3.821	0.896
	(0.562)	(0.508)	
Number of Observations	80	82	

	Panel B		
Variable	**No National Diversity**	**National Diversity**	**p-value of difference**
Number of Job Posting Applications	8.15	8.287	0.786
	(0.318)	(0.386)	
oDesk Rating Prior to Hire	4.543	4.607	0.635
	(0.112)	(0.080)	
No Rating Prior to Hire	0.619	0.604	0.796
	(0.043)	(0.04)	
Number of oDesk Contracts Prior to Hire	3.625	4.787	0.324
	(0.713)	(0.927)	
Indicator for having a Profile Picture	0.863	0.823	0.356
	(0.029)	(0.031)	
Level of Education	1.875	1.817	0.594
	(0.079)	(0.074)	
Number of Offline Jobs Listed on Profile	1.2	1.183	0.866
	(0.076)	(0.067)	
Average Score on oDesk Tests	3.464	3.394	0.335
	(0.056)	(0.045)	
Number of oDesk Tests Taken	2.175	2.835	0.066*
	(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 Contractor	0.131	0.159	0.495
	(0.029)	(0.027)	
Indicator for Agency Membership	0.231	0.274	0.38
	(0.036)	(0.034)	
Number of Items in Portfolio	4.269	3.482	0.299
	(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

Pairs with Lost Contractors	(1)	(2)	(3)	(4)
	Average Performance Assigned		Dropped from Sample	
VARIABLES	Joint Ouput	Joint Producitivity	Team Output	Team Productivity
Team Work	0.271	0.062*	0.320*	0.070**
	(0.165)	(0.033)	(0.175)	(0.034)
National Diversity*Team Work	-0.576**	-0.159***	-0.697***	-0.182***
	(0.233)	(0.050)	(0.259)	(0.055)
Observations	162	162	150	150
R-squared	0.251	0.268	0.277	0.286

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. * significant at 10%; ** significant at 5%; *** significant at 1%

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

VARIABLES	(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
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)	(2)	(3)	(4)	(5)
	Joint Output	Predicted Values			
VARIABLES	Coefficient Estimates	No Features Added	One Feature Added	Two Features Added	Three Features Added
Team Work	0.804*	0.088	0.513	0.357	0.041
	(0.460)				
National Diversity*	-1.741***	0.356	0.540	0.097	0.008
Team Work	(0.670)				
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 Output, Ordered Logit

	(1)
VARIABLES	Increasing Effort
Team Work	0.390*
	(0.211)
National Diversity	-0.282
*Team Work	(0.285)
Observations	162
R-squared	0.192

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. Increasing effort is equal to one if average effort in the second half of time spent on the job is higher than in the first half. The sum of the two coefficients reported in the table is not statistically different from zero. * significant at 10%; ** significant at 5%; *** significant at 1%

Table A6: Effect of Team Work & National Diversity on Individual Performance

VARIABLES	(1) Individual Output	(2) Individual Productivity
National Diversity	0.046	0.025
	(0.102)	(0.027)
Team Work	0.200*	0.036
	(0.107)	(0.030)
National Diversity*Team Work	-0.390***	-0.104***
	(0.143)	(0.038)
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. * significant at 10%; ** significant at 5%; *** significant at 1%

Table A7: Effect of Geographic Distance & National Diversity on Pair Performance

VARIABLES	(1)	(2)	(3)	(4)
	Joint Output		Joint Productivity	
Team Work	0.313	0.117	0.151**	0.128**
	(0.432)	(0.473)	(0.060)	(0.064)
Log(Distance between Cities)	0.019	0.005	0.022*	0.013
	(0.044)	(0.056)	(0.011)	(0.011)
Log(Distance between Cities)*	-0.047	0.026	-0.030**	-0.018
Team Work	(0.068)	(0.094)	(0.013)	(0.013)
National Diversity*Team Work		-0.494		-0.124**
		(0.414)		(0.054)
Observations	112			
Mean Dependant Variable, Independent Work	0.982		0.129	
R-squared	0.272	0.306	0.302	0.383

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. 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. * significant at 10%; ** significant at 5%; *** significant at 1%

Table A8: Effect of Team Work & National Diversity on Performance by Pair Skill Differences, Interaction

VARIABLES	(1) Joint Output	(2) Joint Productivity
Team Work	0.220	0.035
	(0.224)	(0.034)
Indicator for Whether or not Teammates Have Different Skillsets	-0.391	-0.024
	(0.245)	(0.045)
National Diversity*Team Work	-0.135	-0.103*
	(0.352)	(0.058)
skilld_multi	0.777**	0.034
	(0.374)	(0.078)
skilld_comm	0.197	0.063
	(0.324)	(0.064)
skilld_comm_multi	-0.985**	-0.157
	(0.491)	(0.099)
Constant	0.000	-0.000
	(0.000)	(0.000)
Observations	162	162
R-squared	0.434	0.369

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. * significant at 10%; ** significant at 5%; *** significant at 1%

Appendix B Model Appendix

The proofs for the propositions presented in 2 are as follows.

Proposition 1. The effort of workers in teams is increasing in λ.

Proof. The derivative of e^{TW} with respect to λ is $\frac{1+\delta\lambda^2-\lambda}{(1-\delta\lambda^2)^2}$, which is positive. ■

Proposition 2. Conditional on a sufficiently large marginal cost of effort, there exists a value λ^{TW} such that for all $\lambda > \lambda^{TW}$, effort is higher with teamwork than with independent work. For $\lambda < \lambda^{TW}$, effort is lower with teamwork than with independent work.

Proof.e_i^{IW} is not affected by changes in λ. For $\delta = 1$ all $\lambda < 0.62$, $\frac{\lambda}{1-\lambda^2} < 1$ ■

Proposition 3. The effort of workers in a team is increasing in δ

Proof. The derivative of e^{TW} with respect to δ is $\frac{\delta^3}{(1-\delta\lambda^2)^2}$, which is positive. ■

Proposition 4. Communication that requires a high share of effort can be offset by high complementarities.

Proof. Suppose that $\lambda = \frac{1}{2}$ so that with $\delta = 1$, teamwork would be less productive than independent work. Then, a pair of workers with $\delta > 2$ will be more productive with teamwork than without ($e^{IW} = 1$, $e^{TW,ext} = 1$ when $\lambda = \frac{1}{2}$ and $\delta = 2$). ■

Appendix C Data Appendix

Appendix C.1 Supplementary Job Instructions Information[48]

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

[48] 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.

“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.

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 C.2 Supplementary Job Posting Information

This section describes additional information about the job posting information beyond what was described in section 3. 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 C1: Javascript Job Posting

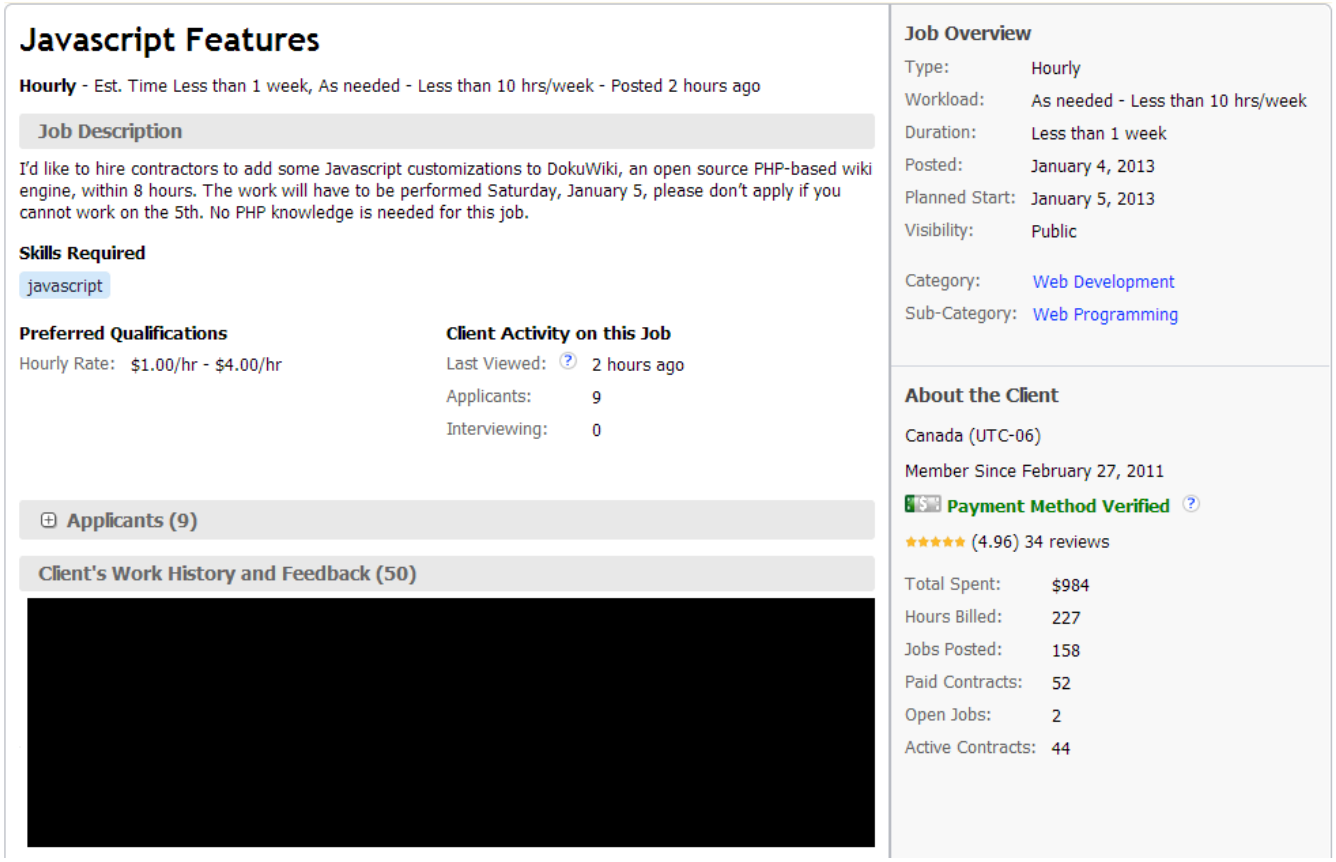
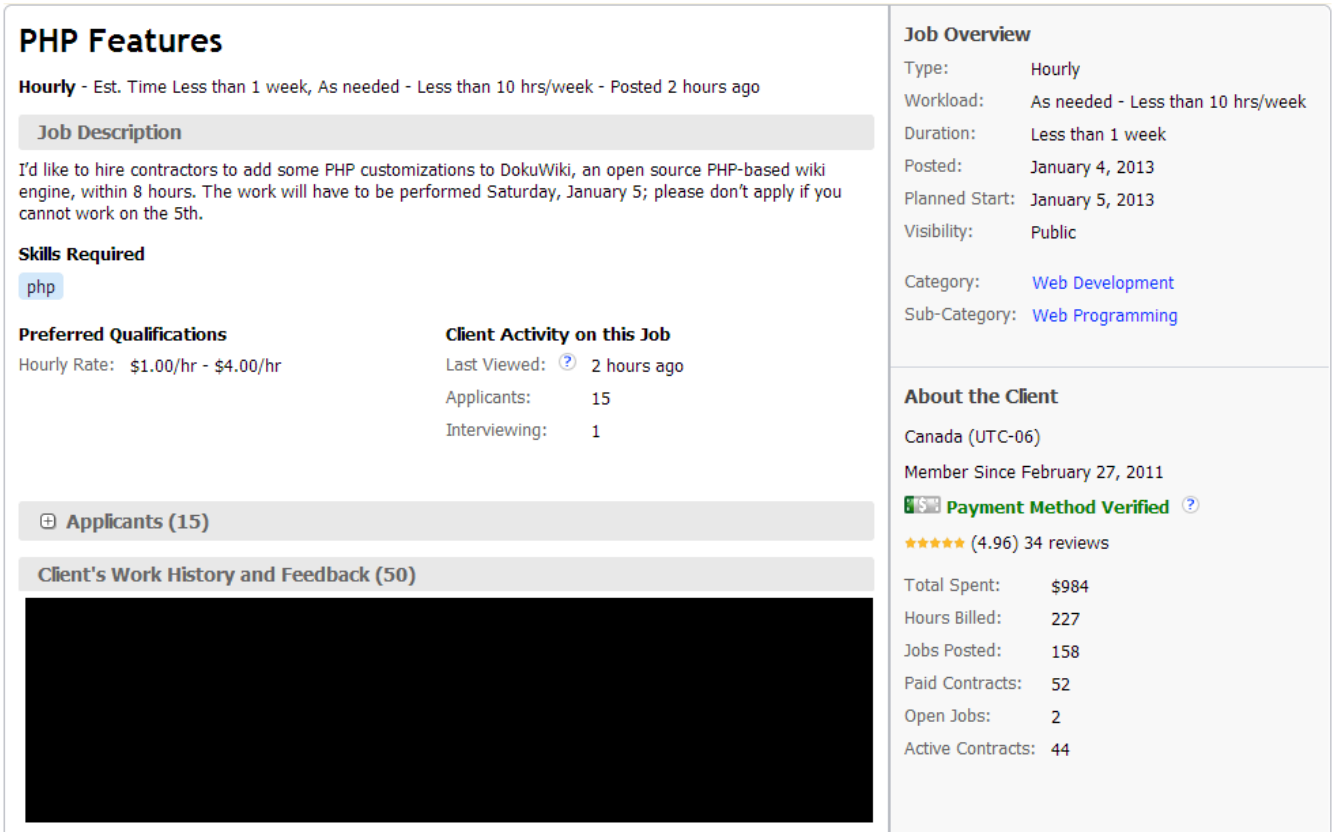


Figure C2: PHP Job Posting



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

Hourly Rate: $1.00/hr - $4.00/hr

Client Activity on this Job

Last Viewed: 2 hours ago
Applicants: 15
Interviewing: 1

Applicants (15)

Client's Work History and Feedback (50)

[redacted]

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

Appendix C.3 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)*.