NBER WORKING PAPER SERIES

GENDER DIFFERENCES IN RESPONSE TO COMPETITIVE ORGANIZATION? DIFFERENCES ACROSS FIELDS FROM A PRODUCT DEVELOPMENT PLATFORM FIELD EXPERIMENT

Kevin Boudreau
Nilam Kaushik

Working Paper 30062
http://www.nber.org/papers/w30062

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
May 2022

We thank members of supporting organizations and particularly Maria Costa De Sousa, Hugh Courtney, Mavez Dabas, Sylvain Demortier, Anthony Donaldson, Raj Echambadi, Koreen Geisler-Wagner, Austen Keene, Afan Khan, Atif Khan, Abhinav Kharbanda, Dyan Khor, Raghavi Kirouchenaradjou, Sreerag Sreenath Mandakathil, Patrick McGrath, Tucker Marion, Marc Meyer, Robert Hughes, Michael Orr, Olga Ozhereleva, Kaushik Padmanabhan, Edwige Poinssot, Fernando Suarez, Prathamesh Kisan Tajane, Nikin Tharan, Emery Trahan, Maureen Underhill, Robert Whelan, and Katherine Wilhoit. For helpful comments we thank Victor Bennett, Roberto Fernandez, Sharique Hasan, Olga Shurchkov, and seminar participants at Duke University, DRUID, and the Academy of Management. We would also like to acknowledge generous financial support from the Kauffman Foundation and Northeastern DMSB. The views expressed here are those of the authors alone and do not necessarily reflect those of supporting institutions, funders or supporters involved with this study. All errors are our own. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Gender Differences in Response to Competitive Organization? Differences Across Fields from a Product Development Platform Field Experiment
Kevin Boudreau and Nilam Kaushik
NBER Working Paper No. 30062
May 2022
JEL No. D02,D03,D9,D91,J0,J01,J2,M0,M12,M5,O3,O31,O32,O36

ABSTRACT

Prior research, primarily based on lab experiments, suggests that females might be more averse to competition than males and could be more inclined towards collaboration, instead. Were these findings to generalize to adults across the workforce, there could be profound implications for organizational design and personnel management. We report on a field experiment in which 97,678 adults from a wide range of fields and ages were invited to join a product development opportunity. Individuals were randomly assigned to treatments framing the opportunity as either involving competitive or collaborative interactions with other participants. Among those outside of science, technology, engineering, and math fields (STEM), we find significant gender differences in willingness to participate under competition. Among those in STEM fields, we detect no statistical gender differences. These results and broader patterns documented in the study are consistent with significant heterogeneity in competitiveness across both men and women, with field and career sorting resulting in differences (in gender differences) across fields.

Kevin Boudreau
D'Amore-McKim School of Business
Northeastern University
360 Huntington Avenue
Boston, MA 02115
and NBER
k.boudreau@northeastern.edu

Nilam Kaushik
Bannerghatta Road, Bengaluru, India
IIMB
India
nilam.kaushik@iimb.ac.in

1 Introduction

A stream of largely lab-based studies has reported that male subjects have a higher probability of choosing to engage in tasks under competitive conditions than female subjects (e.g., Niederle and Vesterlund 2007; Saccardo et al. 2017; Sutter and Glätzle-Rützler 2015).[1] The gender differences observed in these studies have been interpreted as consistent with gender-based differences in "preferences" for working under competitive conditions (e.g., Croson and Gneezy 2009). Female subjects in the literature have instead exhibited a greater willingness to participate under non-competitive piece-rate regimes and collaborative team-based regimes, relative to male subjects within these studies (e.g., Abramo et al. 2013; Dargnies 2012; Healy and Pate 2011; Kuhn and Villeval 2015). These provocative findings and interpretations are supported by numerous replication studies and have even begun to diffuse to wider popular media audiences (Guo 2015; PBS News Hour 2015; Villeval 2012). Nonetheless, the generalizability of these findings remains little understood (Flory et al. 2015). The preponderance of existing empirical results has been drawn from relatively narrow contexts and study populations (especially children and students). Moreover, there is no theoretical consensus on underlying mechanisms that might be used to extrapolate beyond the existing results (see Section 2). In this paper, we take steps towards further understanding the generalizability and boundary conditions of gender differences in responses to competition across a wide spectrum of working-age adults.

It is important to better understand the extent to which working-age men and women might systematically differ in their responses to competition. For example, scholars in this literature now routinely speculate that differences in competitiveness might explain some part of "why women are underrepresented in many high-profile jobs and in whole professions" (Niederle and Vesterlund 2007). If girls and women are more averse to competition, they might systematically sort out of fields and occupations that are perceived to be competitive. This line of argumentation has been used to speculate about the causes of disparities in gender representation in areas such as senior executive positions, STEM fields (i.e., sciences, technology, engineering, and mathematics), and politics (e.g., Balafoutas et al. 2018; Buser et al. 2014; Niederle and Vesterlund 2010). Clearer understanding might therefore shed light on whether "women's talents are sometimes wasted because they avoid competitive environments" (Kuhn and Villeval 2015).

Attaining a better understanding of systematic behavioral patterns in response to competition could also point to ways of improving organizational designs to better harness workers' talents and efforts. This is because the "levers" of organizational design—internal incentive systems, hiring and selection processes,

[1] See Inglehart et al. (1994) for a discussion of this hypothesis pre-dating the more recent experimental literature.

decision making and political processes, work norms, and values—regulate the intensity of competition experienced by workers (Baumann et al., 2019; Lee and Puranam, 2017; Obloj and Zenger, 2017; Song et al., 2018). Optimizing exposure to competition could be especially relevant for new digital forms of organizing—such as crowdsourcing, gig platforms, open communities, and open innovation—where interactions among participants can be "tuned" to be more or less competitive or collaborative (Hunt and Samman 2019; Liang et al. 2018).

In this paper, we take steps toward theorizing and documenting possible limits and boundary conditions regarding gender differences in competitiveness. At the core of the paper is a field experimental approach that allows us to observe responses to competition across a wide range of college-educated, working-age adults. We proceed by first developing theoretical predictions to guide our investigation. In short, we predict that gender differences in competitiveness should differ according to individuals' fields of training, and that gender differences should be significantly smaller among those who sort into STEM fields.

Given the current lack of theoretical consensus on underlying mechanisms shaping responses to competition and possible gender differences (Section 2.2), our approach is to develop predictions based on just the most consistent claims and assumptions across the prior literature. A first point we build on is that there exists some (yet little understood) degree of heterogeneity in competitiveness across the population, i.e., differences in propensity for or aversion to competition. The second point we build on is that individuals' competitiveness serves as a meaningful basis for *sorting into different fields*. This sorting is especially important as young people develop and make path-dependent choices in education, culminating into longer-term career pathways (Kahn and Ginther, 2017). A simplest possible implication of these presumptions of prior literature is that men and women sorting into STEM fields should be relatively similar in their competitiveness—in the sense that those sorting into STEM might each be relatively competitive.

However, whether and to what extent men's and women's responses to competition converge in STEM fields and do so in a manner consistent with presumptions remains an empirical question. To test for gender differences in responses to competition across working-age adults and to investigate possible differences across fields, we report results from a field experiment covering 97,678 adults from a range of fields of training, and ages spanning college years to retirement.

Each individual in our study population was invited to participate in a three-week, part-time product development activity. This activity was designed with the intent of being both accessible and relevant to all invitees, whatever their (technical or non-technical) field of training, experience, or background (see Section 3.1). The experiment focused on randomly assigning individuals to treatments that framed interactions with other participants as either *competitive* or *collaborative*. We, therefore, derive inferences by investigating how the likelihood of participation depended on assignment to competition versus collaboration treatments, along with gender, field of training (i.e., college majors in humanities, business, sciences, computer science, etc.), and other personal characteristics.

We found that in non-STEM fields, gender differences in responses to competition mapped closely to existing intuitions regarding gender-based differences in responses to competition. Non-STEM men's and women's participation under the collaboration treatment was statistically equivalent, with roughly three percent choosing to participate. Non-STEM men's participation was also statistically the same under the competitive treatment. However, the participation rate of non-STEM women was significantly lower in the competition treatment, by about a half percentage point (or a 25.5% drop).

Baseline participation rates were higher among those in STEM fields, with overall participation at roughly four percentage points. Our primary finding, consistent with predictions, is that responses to competition among men and women in STEM fields was similar. Participation rates for both STEM men and STEM women were about one percentage point lower under the competition treatment than under the collaboration treatment. Apart from simply being more similar, the response to competition among STEM men and women within our context is statistically indistinguishable from zero. (Sections 5.2 and 6 provide a fuller presentation and discussion of results.) The results are robust to a series of alternative controls and specifications. We also validated this finding result in tests of 17 STEM more narrowly-defined subfields.

However, whereas we found evidence of similarities of STEM men and women, we did not find this was because they were both more competitive. Both STEM men and women were in fact more willing to work under collaboration than under competition. This result runs counter to simplest interpretations of the sorting-on-competitiveness hypothesis suggested in prior research. As we discuss in Section 6, willingness to work under competition (and collaboration, too) is inherent to much of the training and work in technical problem-solving in STEM. Building on research describing gender dynamics in STEM fields, we argue the results are consistent with those in STEM fields possessing a willingness and ability to work under either competition or collaboration and teamwork, depending on the demands of the context.

Overall, this study takes steps towards clarifying the generalizability—along with certain limits and boundary conditions—of previously documented gender differences in competitiveness. By doing so, this paper continues to build bridges between research on gender and competition and questions of firms and organizations, outside of the lab. Organizational scholarship considering the implications of possible gender differences in competitiveness remains scarce. (See Brands and Fernandez-Mateo (2017) for an exception.) This paper also contributes to a growing number of studies examining how organizational designs and processes shape gender representation in STEM fields (e.g., Fernandez Campero 2017; Murciano-Goroff 2018; Wynn Correll 2017.)

The paper proceeds as follows: Section 2 reviews the prior literature and develops our theoretical hypothesis. Section 3 describes the field experimental research design. Section 4 describes the data and variables. Section 5 reports the results. Section 6 summarizes and discusses the results. Section 7 concludes.

2 Background and Literature

This section reviews prior research on gender-based differences in responses to competition, clarifying the limits of generalizability and extrapolation beyond the existing studies. We also build upon the existing consensus in the literature to develop our main predictions.

2.1 Existing Evidence Suggesting Gender Differences in Competitiveness

The bulk of existing evidence on gender differences in response to competition comes from lab experiments (e.g., Gneezy et al. 2003; Gneezy and Rustichini 2004; Niederle and Vesterlund 2011). The canonical experimental approaches have involved either comparing whether male and female lab subjects choose to work under competitive or noncompetitive conditions or, alternatively, evaluating the effect of competition on the performance of male and female subjects.

To consider questions of generalizability, it is useful to review details of the existing research. For example, Niederle and Vesterlund's (2007) seminal lab experiment studied the behavior of 80 college students recruited from the University of Pittsburgh. This study has emerged as the canonical experimental protocol in the literature. The subjects sat in groups of two men and two women at a time and were asked to add two-digit numbers for five minutes. The subjects were asked to choose whether they would be paid according to a competitive scheme in which they would be paid $2 for each correct answer if they correctly solved more problems than the other three subjects (i.e., a rank order tournament); alternatively, they could choose a non-competitive piece rate of 50¢ per correct answer. The striking finding in this study was that twice as many male subjects chose competition as did female subjects (73 percent versus 35 percent). To help make the claim that the differences between male and female subjects could be interpreted as related to gender differences rather than unobserved factors correlated with the recruited subjects' gender, the authors pointed out that they found no statistical differences in male and female subjects' ability to solve problems in pre-experimental trials.[2] The authors also showed that a residual mean difference remained between male and female subjects after controlling for measures they constructed to reflect risk tolerance,

[2] As with any study using gender as an explanatory variable, including lab experiments, there are unavoidable challenges in interpreting differences as all-else-being-equal "gender effects." As it is not possible to randomly assign a subject's sex (Holland, 1986) it is therefore not possible to unequivocally rule out spurious differences, even with "control" variables (Oster, 2019).

aversion to receiving feedback, and self-confidence.[3] Additionally, men with low ability were interpreted as being over-confident, as they chose competition too often relative to their odds of winning. Women with high ability choose competition too infrequently relative to their odds of winning, and thus were interpreted as being averse to competition.[4]

These findings of male subjects exhibiting willingness to participate under competition, or otherwise performing better under competitive conditions than female subjects, have been replicated over other experimental tasks, subject pools, and variants on the original protocols (Buser 2016; Comeig et al. 2016; Dato and Nieken 2014; Datta Gupta et al. 2013; Dittrich et al. 2014; Healy and Pate 2011; Iriberri and Rey-Biel 2019; Morin 2015, 2015; Paserman 2010; Reuben et al. 2015; Saccardo et al. 2017; Shurchkov 2012; Sutter and Glätzle-Rützler 2015).

In relation to the questions of this study, there are yet few studies concerning working-age adults and in relation to roles in the workforce. Reuben et al. (2015) found that a lab-based measure of the competitiveness in a set of MBA students differed between men and women and that this measure was correlated with entering high-paying industries after graduation. Additionally, Buser et al. (2020) used online survey data and found that a measure of competitiveness among those responding to the survey was associated with participants' completed level of education, field of study in college, occupation, and income. Research by Flory et al. (2015) is especially notable, as they reported on a field experiment that tested how rates of applications by administrative assistants to online job postings in multiple US cities vary depending on the posted compensation scheme. The study finds that in cities where there are higher wages, women are less likely to apply to competitive pay schemes. The authors pointed out that this result can be interpreted as consistent with female administrative assistants avoiding competitive jobs if they have outside options.

2.2 Existing Evidence on Competition versus Collaboration

While many studies contrast competition with a lack of competition, particularly piece-rate regimes, a closely-related stream of papers has contrasted female subjects' aversion to competition with an inversely-related preference for collaborative or team-based work (Abramo et al. 2013; Dargnies 2012;

[3] NB. From a practical perspective, one's risk tolerance, willingness to receive feedback, and self-confidence might alternatively simply be viewed as part and parcel of one's willingness to compete.

[4] Gneezy et al. (2003) illustrates an alternative canonical protocol, recruiting 324 students to solve maze problems for 15 minutes. Students were randomly assigned to competitive, or piece-rate pay schemes in groups of six. Men outperformed women in all treatments, but the differences were not statistically significant in the noncompetitive regime. The authors interpreted the results as indicating that men are more competitive, providing a possible explanation for "gender differences… in competitive high-ranking positions."

Healy and Pate 2011; Kuhn and Villeval 2015). Much of this work follows the same basic outlines of the Niederle-Versterlund (2007) experimental protocol. In addition to comparing forms of competition versus non-competitive (piece-rate) payment, these protocols include tests contrasting collaboration with competition or collaboration with piece-rate.

For example, in a study representative of this work, Dargnies (2012) evaluated 78 people recruited through an online system. She reported that women were more likely than men to choose to "team," where "team" was operationalized as combining (averaging) scores when solving arithmetic problems with an anonymous unobserved "teammate," as opposed to payoffs based on individual performance. Specifically, she found that men with high skills were less likely to accept this version of "teaming" when they were uncertain of the skills of their "teammate." Healy and Pate (2011) carried out a similar protocol in an experiment involving 192 students from Loyola Marymount University, finding similar patterns. Kuhn and Villeval (2015) also implemented a similar protocol in a study of 174 students interacting online at a French University. The protocol was similar, in this case, except the choice between individual pay and pooled average outcomes was made without their being competition among teams. Again, female students disproportionately choose the nominally "cooperative" (shared payoffs) regime. The authors emphasize an interpretation of women experiencing inequity aversion and that "the same confidence deficit that pushes women out of competitions pulls women into teams, where it is beneficial to have an abler teammate." Flory et al.'s (2015) field experiment involving online applicants for administrative assistant jobs included a treatment that mentioning "teams" in the compensation description for an administrative assistant job had a positive influence on the likelihood of women applying.

2.3 Limits to Extrapolating beyond Existing Findings

Notwithstanding a considerable number of replication studies, there are important limitations to generalization that are well-known to authors in this literature. A first challenge to extrapolating beyond the existing results is that the evidence is drawn from narrow contexts and narrow study populations, concentrated on school children and campus lab experiments with synthetic tasks. Important advances in field studies themselves remain somewhat pinpointed in coverage, as the pathbreaking study by Flory et al. (2015), which focused on administrative assistants.

Narrow sampling of evidence places an especially large burden on theory to interpret generalizability. However, there is yet little consensus on underlying theoretical explanations and mechanisms. The prior research has suggested numerous plausible explanations for observed gender differences, emphasizing "attitudes" toward or "preferences" and "tastes" for competition (Croson and Gneezy 2009; Dohmen and Falk 2011; Kleinjans 2009), with an implication of possible biological and/or evolutionary factors playing some role (Apicella et al. 2011; Buser 2009; Kivlighan et al. 2005; Knight 2002; Lippa 1998; Vugt et al. 2007; Wozniak et al. 2014). Other explanations include possible learned, socialized, or innate differences in overconfidence and self-efficacy (e.g., Comeig et al. 2016; Niederle and Vesterlund 2007), socialization of the stereotypes of gender roles (e.g., Gneezy et al. 2006), gender differences in subjective beliefs about the implications and meaning of competition (e.g., Kesebir et al. 2019), and differing psycho-physiological tendencies of men and women to feel stress (Buckert et al. 2017; Holt-Lunstad et al. 2001). We might also interpret related streams of lab studies reporting analogous residual differences in risk attitudes, tastes for high-powered incentives, and aversion to feedback on performance as alternatives or complementary explanations for differences in preferences for competition (e.g., Byrnes et al. 1999; Croson and Gneezy 2009; Eckel and Grossman 2008). Despite the best efforts of the authors in this literature, the inability to randomize subjects' gender means it is not possible to entirely rule out the possibility of spurious factors, especially in small and highly selective study groups.

Apart from the limitations, above, a third challenge is the number of published results that run counter to findings of gender differences (many by the same authors). The authors of these contrarian studies have interpreted results to hypothesize that gender differences could depend on any number of added factors, including socialization, context, individual characteristics, and even how competition is operationalized (e.g., Andersen et al., 2013; Boschini et al., 2018; Dreber et al., 2011; Gneezy et al., 2003, 2009; Günther et al., 2010). More generally, these contrarian results point to a need for further research into possible boundary conditions and contingencies. At the same time, these contrary studies are not themselves immune to the same limitations as the earlier-described studies, described above. A number of these findings that are contrarian to popular intuitions are also contrarian to one another (see Appendix A.)

2.4 Differences Across STEM and non-STEM Fields?

Given the above limits to generalization in the literature, Flory et al. (2015) note: "Whether, and to what extent, gender differences in attitudes toward competition lead to differences in naturally occurring labor markets remains an open question." Our empirical analysis to follow will make progress by reporting more widely sampled evidence on adults across the workforce than has been possible in the prior literature.

However, we first build on earlier research to develop theoretical predictions to better guide and support a generalizable interpretation of the analysis to follow. Below, we enumerate multiple arguments and behavioral mechanisms that suggest that *competitiveness among men and women should be significantly more similar among those in STEM fields than those in non-STEM fields and the wider population.*

Despite a lack of consensus on the theoretical mechanisms (Section 2.2), there are two widely accepted presumptions across much of the prior research that allow us to develop simple predictions. The first presumption is that individuals across the population are heterogeneous and somehow distributed in their propensity toward or aversion to competition. Where prior research has focused on mean gender differences, the presumption also implies there could potentially be considerable variance more generally, both among males and females. Some degree of variance of behavior among men and among women is observed in all prior studies, even small studies.

A second key presumption of the prior literature is that competitive individuals tend to sort themselves into competitive fields (Antonovics et al., 2009; Balafoutas et al., 2018; Buser et al., 2014; Niederle and Vesterlund, 2010a), and particularly within education and within careers thereafter. The greatest available evidence supporting this presumption relates to the mathematics and STEM fields. Specifically, the literature has gone as far to report correlations between measures of competitiveness and sorting into math and STEM fields (Antonovics et al., 2009; Buser et al., 2014; Pregaldini et al., 2020; Reuben et al., 2015). Strictly speaking, this evidence cannot rule out the possibility that sorting occurs based on unobserved characteristics that are merely correlated with competitiveness. This point remains a fundamental limitation in interpreting empirical results, even those of experiments, in this area.

To the extent that these premises are correct and produce economically significant effects, the competitiveness of working-age adults in the economy should reflect some degree of sorting into fields based on competitiveness. Thus, inasmuch as STEM fields are perceived to be more competitive, they should tend to attract more competitive men and women, leading to greater similarities among the men and women selecting these fields. This is an empirical question—and the one studied here.

3 Experimental Research Design

The goal of the latter half of this paper is to empirically investigate gender differences in competitiveness across a widely sampled study population of working-age adults; and to test our prediction that gender differences in response to competition should vary by field (Section 2.3).[5]

3.1 Research Context

This research was carried out in partnership with a large R1 research-oriented university in the United States with roughly 20,000 undergraduate students, 8,000 graduate students, and several hundred

[5] Other experiments and A/B tests were embedded in the launch campaign. However, these assignments were orthogonal to the assignments described here and balanced across all subjects, and therefore have no effect on the results reported here.

thousand alumni in its network. The university is in the top-40 ranked colleges and universities according to the US News Reports list of the best US universities with ranked undergraduate programs in engineering, computer science, and business.

The university runs an online platform-based program to allow alumni and students to gather in teams and compete with each other to design new products related to Internet-of-Things (IoT).[6] The objective of the program is to provide participants with opportunities to design IoT products ("applications"), to gain an understanding of IoT technologies and opportunities, and to learn how these technologies can be relevant to participants' work organizations and sectors. The program also provides experience in increasingly standard product development work practices from "use case" design to technical design and business case design (see Appendix B). These goals are achieved through a series of part-time learning-by-doing product development activities. These exercises are organized as team-based competitive "challenges" or rank-order tournaments. While people sign up as individuals, they are assigned to teams by the platform. To make the program as readily and widely accessible as possible, all activities are hosted on an online platform that allows individuals to participate from any location during off-hours outside of work or school.

The program involves a minimum of three weeks of part-time commitment of one's off-hours, participating in at least one product development exercise. For this opportunity cost, individuals can benefit from participating in interesting projects, learning, interacting with members of the university community, receiving feedback on designs, and receiving cash prizes. Several teams also benefited by taking the ideas and designs they developed on the platform into follow-up entrepreneurial development (off the platform). The opportunity is promoted as highly relevant to all students and alumni, in terms of developing multi-disciplinary product development skills, and in terms of exposing both nontechnical and technical individuals to technical and business opportunities made possible by the Internet-of-Things. The platform work and learning process was designed with the philosophy of "low floors and high ceilings" (Boaler, 2016) to engage a wide range of participants. Thus, relative neophytes could productively engage in learning about the technology and contributing to designs, by following defined steps and platform tools. Meanwhile, the platform also provides pathways for relative experts in the fields to channel their expertise to develop designs. The platform has grown to more than 5,500 participants and includes participants from all fields of training and career stages in all US states and five continents.

[6] The Internet-of-Things or "IoT" refers to a wide set of largely yet-to-be-imagined products, services, and systems that connect machines, infrastructure, consumer products, and other things while making use of the intelligence created by data collection and networking. IoT is widely expected to generate an impact as wide as the Internet, impacting occupations and firms across the economy in a "Fourth Industrial Revolution" or "Second Machine Age" (Brynjolfsson and Mcafee, 2014). These technologies are predicted to affect most sectors across the economy (Patel et al., 2017).

3.2 Timeline and Procedure

Embedding our experiment in the initial launch of the platform allowed us to be part of a one-time process of mass communication with 97,678 potential adopters under highly controlled conditions. By tracking responses of this known set of individuals for whom descriptive data are available, the effect of random assignment to the competition collaboration treatments can be assessed.

3.2.1 Randomized Assignment to Competition vs. Collaboration Treatments

An introductory email introduced individuals to the opportunity, inviting them to the platform to learn more details and join. Invitation emails were sent over 60 business days, with individuals randomly assigned to each day. Spreading invitations over time allowed the engineering team to monitor platform scaling while minimizing the risk that any one day could produce eccentric results.

Individuals were randomly assigned to competition or collaboration treatments with the aim of influencing subjects' perceptions of interactions they would encounter with other participants. Subjects were assigned to receive different emails depending on their treatment. Within the email invitation they could click on a link in the email to send them to the platform ("*Click HERE to learn more and to sign up to the platform*"). The platform landing page also provided additional details including high-level descriptions of the technical nature of IoT and the opportunity to create new products and applications. The version of the platform landing page to which they were sent also depended on whether they were in the competition or collaboration treatments, where these platforms only differed in the sense of repeating the distinct language used in the emails.

The details of both the emails and landing pages varied in the same way, depending on the treatment. What follows are the relevant subset of distinguishing words from both the email and the platform for the competition treatment:

> "*This is a two-sided **competitive** platform to ideate and innovate 'smart' Internet-of-Things (IoT) products and services—using hardware, software, networking, data and algorithms.*
>
> *On one side of the platform, companies seek solutions to their IoT innovation challenges. On the other side, **you will compete with other participants to solve IoT innovation problems** of companies for cash and other benefits*."

For the collaboration treatment, the bolded words were replaced with "*collaborative platform*" and "*you will work within a team to solve IoT innovation.*" Therefore, the content and length were virtually identical for both treatments, except for these key bolded words.

Individuals who did not act on the email were sent a single reminder seven days later (also at 10 am EST). We explicitly added a line in the email stating that the invitation was not transferable or to be forwarded. Nonetheless, there were 26 instances of forwarding of emails; we were able to detect such forwarding, and we removed these instances from our analyses.

Once on the platform, individuals could choose to participate by clicking on a red "join and participate" button prominently displayed on the site. Users were prompted to log in with their LinkedIn passwords. This sign-in and authentication allowed their LinkedIn picture, name, link, and brief blurb to be used to create their platform profiles.

The analytics generated by the email management system (Mailchimp) and the platform, allowed us to determine the number of people among the 97,678 receiving invitations who positively double-clicked to open the invitation they received and then proceeded to the platform. Positive confirmation of opening rates for females and males were statistically identical, with 46.5 percent of females and 45.9 percent of males opening. The true rate at which individuals read emails will be higher, as the email management system tracked active clicking on emails. However, most modern browsers to read an email without double-clicking on it to open the email. Further, even the title of the email mentioned the institution's "competitive platform" or "collaborative platform" for IoT product development. Therefore, even reading the title implied exposure and treatment. We cannot be sure of the true share of emails that were not read. To the extent there is some fraction of our study population who did not, in fact, read the invitation, we should expect weaker statistical results in our analysis.

3.2.2 Controlled and Independent Bilateral Communications

A crucial aspect of running this experiment was to carefully control the information environment. Running the experiment during the initial launch campaign, at a time when prospective adopters had no prior experience or preconceived notions of the platform, allowed us to selectively emphasize the competitive or collaborative aspects of the platform in our communications. By implementing alternative emails and entirely separate platforms with distinct communications, we could present separate and controlled information to each subject, depending on the treatment. Importantly, the information presented in both the competitive and collaboration treatments was also truthful, as the platform and program within which subjects would eventually participate would involve competition among teams.

In controlling the information environment, it was also important that the *only* form of communication and marketing of this initial 60-day campaign be this sequence of bilateral interactions with the platform that would occur prior to opening the platform to development events. For example, there was no simultaneous broad-based marketing campaign. We were also aware of the possibility that individuals within the experiment might potentially, themselves, share messages with their peers. If messages were to be shared with other subjects in our study population, resulting in the mixing of treatments, this would likely weaken any treatment effect by mixing messages. We attempted to minimize such interactions by adding a line to all communications indicating that the subjects should not forward any messages. Using tracking and analytics, we found that among the tens of thousands of communications, there were only 26 instances of forwarding. We removed these instances from our data set; doing so had no effect on the results. We also considered the possibility that any interactions and information diffusion occurring between individuals could be more likely toward the end of the 60-day campaign. However, including time controls, even individual-day fixed effects, showed no effect on results.

3.3 Links with the Prior Experimental Literature

The experimental protocol here is based directly on previous research in the literature on competition and gender (Section 2.1). To observe whether individuals participate conditional on the organizational regime, our experimental protocol follows prior work from Samek (2019), for example. The advantage of this approach is that it allows us to assess preferences to participate under each regime independently (relative to an outside option of not participating). The alternative in the literature is to recruit only those subjects who are willing to participate (as when recruiting to a campus lab experiment) and then ask individuals to choose the regime they would prefer, effectively providing a rank-order preference for one regime versus the other, rather than an independent evaluation of each regime. Another advantage of studying whether people participate or not under either regime independently is that we are able to explicitly observe the selection process into participation in the exercise, rather than having the selection process be part of recruitment to the experiment.

Our approach also follows the approach of prior studies that compare willingness to work under competition with willingness under collaborative regimes (e.g., Abramo et al. 2013; Dargnies 2012; Healy and Pate 2011; Kuhn and Villeval 2015), as opposed to comparing preferences with a piece-rate regime (e.g., Niederle and Vesterlund, 2007). This approach has natural advantages in our context, as it is simply not feasible to implement a piece-rate for innovation, development, and uncertain learning and problem-solving in our context of developing novel IoT projects. In practice, we also exploited the organization of competitions on the platform between teams as a means of focusing on competitive or collaborative and team-based treatments while scrupulously avoiding deception of any of our subjects in any of our communications.

Prior research tends to operationalize competition and collaboration as a matter of either a small (often two-person) winner-take-all rank-order tournament for competition or averaging performance payoffs for collaboration (see Section 2.1). Here, we simply directly use the words "competitive" and "collaborative" to describe the organizational context, allowing subjects to discern for themselves what should be meant by these terms in an organizational context where they carry out work and learning akin to activities in industry.

The largest and most important departure from the bulk of previous research in the literature focused on gender-based differences in competitiveness is that ours is a field experiment rather than a lab experiment involving synthetic tasks. Among the field experiments in this study, this paper comes closest to that of Samek (2019), in the sense of examining who applies for an office support job, conditional on organizational regime. She focuses on a study population of students. This paper is also arguably more closely related to Flory et al. (2015), as it also involved a large field experiment, involving nearly nine thousand subjects; although that study focused on only a market for administrative assistants. Here, we examine a considerably larger number of college-educated adults from all fields.

4 Data, Variables, and Descriptive Statistics

Our study population included 97,678 people in the risk set of potential platform adopters. (A broader population was targeted by the platform, but this number includes the individuals identified and assigned to treatments during the experiment.) Data for these individuals was kindly provided by the institution, matched, and anonymized before the launch. The data set included data on majors and graduation year—and academic performance (cumulative undergraduate GPA) for 13,926 individuals from cohorts after 2009. Further, we used an online API-based service (https://gender-api.com) to probabilistically assign the vast majority of names in our dataset to genders. Table 1 describes the main variables used in the analysis. Figure 1 summarizes the distribution of individuals and the gender breakdowns by GPA and cohort (reported in terms of age).

<Table 1>

The participation rates of men and women in STEM fields were 3.97 percent and 4.47 percent, respectively; in non-STEM fields, they were 2.62 percent and 1.98 percent, respectively. Participation was higher for men than for women, and this was true across ages and academic performance, shown in Figure 1.[7]

<Figure 1>

[7] Gender differences in participation are also observed when stratifying between undergraduate, graduate, and alumni subgroups.

Therefore, the initial platform launch attracted thousands of adopters from the risk set. According to university officials, these participation rates were (very) high relative to most other university programs and clubs directed at both students and alumni. These levels of participation are also higher than most comparable prior studies. For example, Samek (2019) advertised a research assistant position to 35,462 university students and received applications from 1.38 percent, from which she derived inferences by comparing differences in participation across treatments. These adoption rates might also be regarded to be high in relation to, say, a typical platform launch by a private company, who might typically not expect to achieve multiple percentage points of adoption among the general population in the first couple months of operation. By contrast, those studies that non-randomly select their participants conditional on an expression of interest in participating, of course, tend to find higher eventual participation rates (e.g., Flory et al., 2015; Niederle and Vesterlund, 2007).

Although randomization should ex-ante lead to equal groups in treatment and control, Table 2 tests and confirms that observable characteristics of those assigned to competition and collaboration treatments were virtually identical.

<Table 2>

5 Results

5.1 Baseline Study Population Average Gender Differences

This first subsection begins by replicating findings from previous research.[8] We show that response to competition (in terms of percentage difference in participation rates between the competition treatment and collaboration treatment) is greater for women than for men in the study population average. The magnitude of the gender difference is only marginally statistically significant and smaller than prior lab experimental research (Section 2).

Model (1) of Table 3 reports gender differences in participation rates, regressing the participation indicator on the gender dummy *Female* and constant in a linear probability model. Heteroskedasticity robust standard errors are reported.[9] Coefficients are expressed as percentage points (i.e., 100 = 100 percent). The constant term, or probability of males participating, is 3.19. The coefficient in the female dummy is highly

[8] Instructions for accessing data and code to replicate the analysis reside on the Open Science Framework at the following link https://osf.io/6dsn8/

[9] Linear models ease the interpretation of interactions that are central to this investigation.

significant -0.67 percentage points (p = 0.0001).[10] This result means that women are 21 percent (i.e., -.67 / 3.19) less likely to participate on average than are men, overall.[11]

<Table 3>

Model (2) adds the indicator for the dummy variable for *Competition* to the model, revealing that those assigned to the competition treatment are 0.71 percentage points (p = 0.0001) less likely to participate than those assigned to the collaboration treatment. Model (3) introduces an interaction between competition and gender. The coefficient of this interaction is statistically zero, which means that there is no absolute change difference between men and women. These basic differences between men and women's responses are summarized an equivalent comparison of unconditional participation rates in Panel I of Figure 2.

<Figure 2>

Therefore, the absolute difference in participation differences across competition and collaboration treatments is statistically the same for men and women. However, as baseline levels of participation differ, we also test whether the *percentage* differences across competition and collaboration for men and women statistically differ. A comparison of percentage changes arguably also provides a more direct comparison with prior experimental results (Section 2), in which subjects necessarily proceed to participate, and the percentage of subjects choosing one work regime or another is reported. Figure 2 summarizes the results of model (3) (equivalently, Figure 2) in terms of percentage changes i.e.,: $\frac{Participation(Competition)-Participation(Collaboration)}{Participation(Collaboration)}$. The difference in women's participation from collaboration to competition treatments is $\frac{\beta_{Competition}+\beta_{Female \times Competition}}{constant+\beta_{Female}} = -24.86$ percent. For men, this difference in participation rates is $\frac{\beta_{Competition}}{constant} = -20.08$ percent. A Wald test of the nonlinear restriction that these percentages are the same, $H_0: \frac{\beta_{Competition}}{constant} = \frac{\beta_{Competition}+\beta_{Female \times Competition}}{constant+\beta_{Female}}$, is significant at p = 0.097. These differences are summarized graphically in Panel II of Figure 2

These population-average results would appear to be robust to numerous alternative controls and specifications (aside from accounting for field differences, as in Section 5.2). The pattern of negative coefficient estimates on *Female* and *Competition.,* implying a greater negative percentage response to competition among women, does not change when controlling for a series of dummies corresponding to the

[10] We ran an analysis in which we replaced participation here with a measure of participation in other university-sponsored activities (counts of university clubs and activities listed on Linkedin). The patterns are starkly different, confirming that the patterns observed here do not simply reflect some general inclination to participate in university activities, in general.

[11] A low *adjusted-R^2* statistic is consistent with high variance with a low mean incidence of 1's.

day an individual was invited to join (model 4), whether subjects are students or graduates (model 5), or age (equivalently, cohort) (model 6). We also find similar patterns when controlling for college cumulative grade point average (GPA) (model 8) or all controls at once (model 9). Note that, since tests with college GPA data include only the 13,926 individuals from cohorts after 2009, for which we have GPA data, we also report main model results with just this subset of observations (model 7) to provide a more direct comparison of results in models (8) and (9). The coefficients in model (7), estimated in the subsample and those of model (3), are the same in sign and similar in the relative size ratios between the model coefficients. However, the magnitudes are larger in model (7), as this subsample comes from more recent graduates, where overall participation is higher in younger cohorts (see Figure 1). This is consistent with possible higher returns or lower opportunity costs to participate in this exercise among younger cohorts.

5.2 Main Results: Differences Between Fields

This subsection investigates our main prediction that gender responses to competition differ by field (Section 2.3).

5.2.1 Controlling for Fields

As a first step, we investigate whether simply controlling for a series of dummies for field of training affects the earlier result of Section 5.1. In Table 4, models (1) and (3) re-report population average results, both without and with the *competition* treatment effects. Models (2) and (4) then re-estimate these models with field dummies. Whereas results of Section 5.1 were highly robust to a wide range of alternative controls and tests, here we see that including controls for field fundamentally alters key coefficient estimates. Most striking, comparing models (1) to (2) shows that controlling for field differences leads the coefficient on *Female* to reverse its sign. Comparing models (3) and (4) also finds this reversal of the sign on the *Female* coefficient, once controlling for field-fixed effects.

The coefficients on the field dummies are reported in Table 4 ordered by their magnitude. The estimates indicate that those in the STEM fields of computer science, engineering, and science are more willing to participate than those in other fields. This is consistent with product development in the Internet-of-Things perhaps being more immediately accessible or interesting to those in STEM fields (even despite the considerable efforts to make this activity relevant, accessible, and interesting to all fields, see Section 3.1).

<Table 4>

5.2.2 Gender Differences in STEM Versus Non-STEM Fields

We now test our specific hypothesis of Section 2.3, that gender differences in STEM fields should be significantly smaller than those in non-STEM fields. Our approach is simply to stratify results from the above model by those in STEM versus non-STEM fields. (Recall, *STEM* is defined in this context to include engineering, computer science, and sciences.) Model (1) of Table 5 begins by simply restating overall results, controlling for fields, as a basis for comparison. Models (2) and (3) re-estimate stratified results stratified by subsets of non-STEM and STEM observations. Model (2) reports that in non-STEM fields, the difference between participation under competition versus collaboration for women is 0.47 percentage points more negative than for male counterparts in the same fields, significant at $p = 0.05$ (i.e., the coefficient on *Female* × *Competition*.) Model (3) reports that in STEM fields this difference of participation versus competition is not statistically different between women and men. Note too, in model (3), that those in STEM fields are generally more willing to participate in the collaboration regime than they are in the competition regime (i.e., negative coefficient estimated on the *Competition* variable). These responses, stratified by men and women and by non-STEM and STEM fields are graphically summarized in Panel I of Figure 3. Model (4) summarizes the same information of model (2) and (3) or Panel I of Figure 3 within a single regression specification. (This alternative compact way of summarizing results will be referred in robustness results, below.)

<Table 5>

<Figure 3>

We also test whether the *percentage* differences across competition and collaboration for men and women in STEM and non-STEM fields are significantly different from one another. Panel II of Figure 3 presents our results of models (2) and (3) as percentage differences in participation rates, stratified by gender and STEM versus non-STEM fields, i.e., $\frac{Participation(Competition) - Participation(Collaboration)}{Participation(Collaboration)}$.

As summarized in Figure 3, men and women in non-STEM fields show large differences in their response to the competition treatment, with a difference in participation of -3.93 percent and -25.49 percent in relation to the collaboration treatment, respectively. A Wald test of the nonlinear restriction that these percentages are the same, $H_0: \frac{\beta_{Competition}}{\mu_{Non-STEM\ Female, Collaboration}} = \frac{\beta_{Competition} + \beta_{Female \times Competition}}{\mu_{Non-STEM\ Female, Collaboration}}$ in model (4), is significant at $p = 0.01$.

Consistent with our main prediction, and in stark contrast to the above patterns in non-STEM fields, men and women in STEM fields respond similarly to the competition treatment. The difference in participation for men and women in response to the competitive treatment is -19.62 percent and -23.63

percent, in relation to participation in the collaboration treatment. A Wald test of the nonlinear restriction that these percentages are the same, $H_0 : \frac{\beta_{Competition}}{\mu_{Non-STEM\ Female, Collaboration}} = \frac{\beta_{Competition} + \beta_{Female \times Competition}}{\mu_{Non-STEM\ Female, Collaboration}}$, is not statistically significant in the least (i.e., $p = 0.73$).

5.2.3 Robustness of Main Results

A series of additional tests reported in Table 6 assesses whether the earlier established patterns of results of the earlier section, and particularly those summarized in model (4) of Table 6, are robust to controls for various factors influencing benefits and costs of participating. We indeed find the results to be robust.

For example, model (1) begins by re-reporting the earlier patterns for comparison. Model (2) tests the possibility that the age and cohort distribution of men and women could vary in some way between fields in ways that lead to spurious findings, adding age as a control. There is no effect on the coefficients of the main model. Controlling for a quadratic time trend or distinct time trends for each gender does not affect results (not reported).

We also test whether variation in opportunity costs could be a spurious factor. To test for this possibility, we include a measure of an important source of opportunity costs, being in one's child-rearing years.[12] Model (3) includes an indicator switched on for ages 29 through 35 along with an interaction with gender. The years of child rearing are associated with lower participation, but the main results remain unchanged.

<Table 6>

Model (4) tests whether a measure of aptitude (academic achievement) might somehow be a spurious factor in the model. We first re-estimate the main model on the subsample for which we have GPA data to confirm similar patterns (Model 4). Model (5) then re-estimates the model while including GPA and GPA interacting with gender. A first point to observe is that accounting for GPA does not alter main coefficients, comparing coefficient estimates in models (4) and (5). In this sense the results are robust to including GPA. We remark, too, that the considerably larger coefficient on the three-way interaction, *STEM Field × Female × Competition,* suggests even greater (more positive) differences in response to

---

[12] The average college-educated American women has her first child at the age of 30.3 (Bui and Miller, 2018). We did not find a definitive source of the distribution over time, but results are insensitive to shifting this age range by 3 years on upper or lower bounds. Note, in this cross-section, the age is also the graduation cohort—and therefore open to multiple interpretations.

competition among women in STEM for whom we have GPA data (i.e., younger cohorts).

5.3 Gender Differences Across STEM Subfields

This section re-estimates our model over 17 more narrowly-defined STEM subfields in order to determine whether we continue to see similarities in men's and women's responses to competition at the level of individual subfields.

In short, despite considerable variation in overall participation rates across individual STEM subfields, we see that responses to competition among men and women are similar when examined at the level of subfields. To study patterns at the level of subfields , we worked with university staff familiar with the dozens of degrees offered in these fields to map them to subfields, listed in Table 7. Degrees with ambiguous mapping were removed from this subfield analysis. (Computer science during the study period does not break down into separate degrees or subfields.)[13] We kept within the analysis here all STEM subfields with at least 200 individuals.

Table 7 reports model estimates of participation rates for each subfield. Figure 4 reports the corresponding percent changes in response to competition for both men and women. We must be somewhat cautious in interpreting the results, as this fine-grained stratification of the analysis reduces statistical power, and we might expect large differences in gender dynamics between STEM subfields.

<Table 7>

<Figure 4>

Consistent with previous findings of an average negative response to competition of approximately 20 percent, the coefficient estimates in Table 7 are largely negative, and the percent responses to competition shown in Figure 4 are also largely negative. Examining the percentage changes in response to competition in Figure 4, we found no evidence of statistical significance of the subfields in 16 of 17 subfields. The only gender difference that is statistically significant at conventional levels in computer engineering ($p = 0.01$), and women are much more averse to competition (-83 percent change in response to competition for women, versus -15 percent change for men).[14]

Although we must be cautious in interpreting statistical significance, particularly now as we stratify our analysis into fine-grained subsets, what is most strikingly consistent with the earlier results and

---

[13] Estimating the model separately for computer science, engineering, and sciences shows similar results as STEM, overall, with men and women responding with statistically similar percentage changes in participation under the competition treatment relative to the collaboration treatment.

[14] The next most statistically significant gender difference was in biology, where differences are significant at $p = 0.18$.

theoretical hypotheses is that it is overwhelmingly the case that men's and women's responses to competition appear to be much more similar within subfields than between subfields. As can be seen in Figure 4, in most cases, the responses of men and women follow the same sign, and even the magnitudes of the responses are very similar within the individual subfields. These patterns are broadly consistent with the general claim and earlier documented patterns suggesting that men and women sorting into STEM fields are more likely to respond similarly to competition.

6 Summary and Discussion of Results

6.1 Main Finding: Smaller (Negligible) Gender Differences in STEM Fields

We found that in non-STEM fields, gender differences in responses to competition mapped closely to findings of gender differences reported in the prior research (Section 2). Non-STEM men and women participated at similar rates in the collaboration treatment, but differed significantly in responses to the competition treatment. Whereas non-STEM men participated at statistically similar rates under competition as under collaboration (with participation rates just 3.93 percent lower under competition), non-STEM women significantly reduced their participation rate by 25.49 percent. These patterns are presented in detail in Figure 3 and in Table 5. The observation of relative aversion to competition among women in non-STEM fields is consistent with findings of the prior lab research.

In Section 2.3, we hypothesized that gender differences in STEM fields should be smaller than those in non-STEM fields. Indeed, we found that differences in participation for STEM men and women in the competition treatment, relative to their participation rates in the collaboration treatment were statistically identical (19.62 percent lower participation for STEM men under competition; and 23.63 percent lower participation for STEM women under competition). Again, these patterns are presented in detail in Figure 3 and in Table 5.

Should we interpret the statistically identical responses of men and women in STEM fields as reflecting truly small differences, or just an inability to detect differences with sufficient statistical power? A variety of patterns suggests our tests are sufficiently powered and the differences are indeed negligible. Most starkly, in instances where we do in fact find effects, these effects are large and highly statistically significant—including differences between men's and women's responses in non-STEM fields. In the case where we do not find a significant difference—between men's and women's responses in STEM—this is not because of losing precision of estimates and standard errors "blowing up." Rather, the point estimates are simply small (e.g., compare models (3) and (4) in Table 3). Moreover, given the overall participation rates are higher in STEM, the treatment effect might have been expected to mechanically generate greater variance across women and men in STEM fields, making it easier to detect gender differences had these been large.[15]

Of course, these findings of no statistically discernible gender differences in the responses to competition in STEM fields were established in data covering 31,923 STEM subjects. Just this subset of our data is orders of magnitude larger than the bulk of prior research (Section 2). We also found estimates to be robust to a series of controls (Table 4). Re-estimating the model with 17 narrower STEM subfields also revealed mostly similar results (Section 5.3). Moreover, it is only precisely where we theoretically predict to find small(er) differences that we indeed do find small differences. Therefore, all evidence and arguments point to gender differences being small among STEM subjects.

Nonetheless, we should be cautious in interpreting these results as a finding of true zero average gender differences in STEM, or that this difference is necessarily zero generally. The point estimates of percent changes in participation in response to competition were still lower for women than for men (Figure 4), despite the lack of statistical significance. Further, while the nature of the work was intended to be representative of the sort of work carried out in the industry (Section 3.1), these results must be understood in relation to the activity under study here. The experiment was not designed to examine the effects of variation in the task. Further, while our examination of 17 subfields showed remarkably similar participation rates of men and women within particular fields, there were isolated exceptions.

6.2 Secondary Finding: STEM Subjects Prefer Collaboration over Competition

Although our primary research question focuses on relative gender differences in responses to competition, our research design also allowed us to gain insight on general preferences for working under competition versus collaboration in a widely sampled study population, more generally. Here, we find that *both* men and women in STEM fields have an absolute preference for working under a collaborative regime, as summarized in Figure 3 and Table 5. STEM men participated 19.62 less frequently under competition—a 3.60 percent participation rate under competition versus 4.48 percent participation rate under collaboration. STEM women participated 23.63 percent less frequently under competition—a 4.37 percent participation rate under competition versus 5.56 percent participation rate under collaboration.

[15] We thank the editors and a referee for this point.

6.3 Interpretation of Results

Here we review several explanations for the reported patterns.

6.3.1 STEM Sorting, Socialization, Training and Experience

It is widely presumed sin research on gender and competition that individuals may sort into distinct fields and occupations based on competitiveness. We hypothesized that we should observe smaller gender differences in competitiveness in STEM. A simplest interpretation that follows from discussion in the prior research (Section 2) is that men and women in STEM should be more similar, as they will each tend to have a relatively high preference for competition. But, rather than find that STEM men and women generally preferred competition, we found that opposite: that both STEM men and STEM women had a greater willingness to participate under collaboration in our context. Therefore, a more nuanced explanation is required than simply a "preference" for competition.

One possible explanation for similar responses to competition among STEM men and women, but a general greater willingness to participate under collaboration is suggested by prior research on dynamics in STEM fields. For example, we might expect that student and workers in STEM fields could acquire more moderated or contingent responses to both competition and to collaboration. On the one hand, competition is well documented to be a regular feature of work in industrial firms, such as in laboratories, both within and between organizations (e.g., Ali et al., 2018; Lam, 2011). On the other hand, the demanding nature of problem-solving and the advance of knowledge demands in STEM fields has made it increasingly common for technical problem-solving to be carried out in teams—whether in the lab, in the product development department, within the software development community, and so forth (e.g., Jones 2009; Mortensen 2014; Shah 2006). In this sense, the ability to work under *both* competition and collaboration is a regular aspect of modern STEM work (e.g., Menekse et al., 2017). It is also plausible in a context where both competition and collaboration are regular work modes that STEM workers might perhaps, to some degree, be able to recognize the organizational approach most suited to a certain context and set of circumstances. For example, in our context the task was one of (inherently multi-disciplinary) product conception and development, which could have favored team-based collaboration. Therefore, it is possible that those in STEM could be more competitive in a sense of being open to and able to perform under competition (and collaboration, too), rather than simply inherently "preferring" one work mode over another.

Other research documenting gender dynamics in STEM fields suggests still other reasons for smaller gender differences in STEM. Descriptive case study evidence suggests that STEM women may be more ambivalent to *both* competition and collaboration. For example, female engineers have been reported to exhibit ambivalence or aversion to teams and collaboration where they are subject to gender-stereotyped expectations and work role assignments (Cech et al., 2011; Cheryan et al., 2017; Choi, 2013; Coffman, 2014a, 2014b; Foor et al., 2013; McIlwee and Robinson, 1992b; Moss-Racusin et al., 2018; Silbey, 2015, 2016), and competitions have been reported to be an important means of engaging and socializing women and girls into STEM fields (Friesel and Timcenko 2011; Kuyath and Yoder 2004; Notter 2010). On their own, these arguments imply that STEM females may be relatively moderated or contingent in their willingness to work under collaboration and competition.

Other strands of research directly focused on STEM students and workers argue that STEM fields exhibit cultures and norms promoting practices and norms associated with masculinity, including competitiveness and assertiveness (Cheryan et al. 2009; Etzkowitz et al. 2000; Kvande 1999; McIlwee and Robinson, 1992). Indeed, women scientists and engineers have been documented to sometimes adopt practices to conform with their field's culture (Kvande 1999; Ong 2005; Powell et al. 2009), including emulating competitive behaviors (Cheryan et al. 2017). These arguments, on their own, would suggest that women who remain in STEM fields may then potentially converge towards norms of competitiveness or assertiveness. While there could be some convergence of men and women's predispositions—either through sorting or socialization—again, it is possible that the factors shaping willingness to work under competition and collaboration could be more manifold, nuanced, and contingent, as noted above.

6.3.2 "Intellectually Distant" Non-STEM Men and Over-Confidence

Apart from our main finding related to STEM subjects, our results also allow us to consider the more regular results observed in non-STEM fields and discern plausible explanations. For example, men in non-STEM fields responded to competition and collaboration treatments in statistically identical ways, suggesting an indifference between the two treatments; whereas non-STEM women participated less under competition (model (2) of Table 5). On the one hand, these results might simply be considered consistent with women responding more negatively to competition than men do, as in the prior research (Sections 1 and 2).

However, considering other patterns we find here, we might also observe that this indifference of non-STEM men between competition and collaboration is consistent with over-confidence. This is especially so if there were advantages to collaboration in this context of (inherently multi-disciplinary) product development. Further, working in the Internet-of-Things might have created particular disadvantages for those without technical training. Over-confidence among men has been previously suggested as an explanation for gender differences in competitiveness, particularly so among lower-skilled men (e.g., Niederle and Vesterlund 2007). Of course, men in non-STEM fields cannot as a group be judged to be lower-skilled; but in this instance, they may be less familiar with or more "intellectually distant" from the task at hand. Therefore, it is possible that more intellectually distant males are more susceptible to misconstruing the work to be performed and to exhibit overconfidence in their choices.

This explanation of "intellectually distant" men being over-confident is consistent with prior research finding that domain-specific knowledge and familiarity with a task can be highly important in shaping responses to competition (Wieland and Sarin 2012). Comeig et al. (2016) also show that experience and expertise in an area alters attitudes toward competition by moderating confidence. Kamas and Preston (2012) found in their lab study of 310 students solving mathematical problems that while women were less willing to choose competition than men, the differences were not statistically significant among science and engineering students, who presumably have greater training in mathematics.

6.3.3 Non-STEM Women, Gender-Role Congruity, and Self-Efficacy

Apart from possible over-confidence of non-STEM men, it is possible that non-STEM gender differences also arise from a range of previously hypothesized reasons for an aversion to competition, at least among non-STEM women. For example, women in non-STEM fields might perceive the technologically intensive product development to be a male-oriented task (Nosek et al. 2002, 2009; Su et al. 2009), and consequent stereotype threat could result in a diminution in self-confidence and resulting willingness to participate under competition (Dreber et al. 2014; Woodcock et al. 2012). Gender stereotyping of the task might also have led to an expectation of disproportionate participation by men, which could itself reduce willingness to compete (Booth and Nolen 2012; Gneezy et al. 2003; Lundeberg et al. 1994).

We emphasize, too, that apart from these prior hypotheses from the literature, it is also possible that non-STEM women are simply responding somewhat sensibly to the available choices in our context. Consider again that if there are advantages to collaborating in this product development task, there might especially be advantages to non-technically trained individuals to team-up and collaborate. Under this interpretation, it should not then be a surprising that non-STEM women have a lower participation rate in the competition treatment.

6.3.4 Demand-Side Explanations

Another more general point to raise in interpreting our results is that worker participation in the workforce reflects two-sided matching between worker supply and employer demand. In contrast, this experiment here (as in others in the literature) was designed to focus on the supply-side willingness of individuals to participate or not. (See Murciano-Goroff [2018] and Price [2011] for analyses of demand-side employers.) Nonetheless, it remains plausible that subjects' decisions to participate could have *anticipated* demand-side features of the workplace (Fernandez-Mateo and Kaplan 2018).

For example, Goldin (2014) hypothesized that women's participation and success in the workplace is often curtailed by rigid and long work hours and a need for "facetime." The collaborative product development platform used in our context was designed with the intent to minimize the costs of inflexible time requirements, including an ability to interact asynchronously. Closer to the crux of this experiment, if collaboration was viewed as requiring more inflexible "facetime" costs, this might have had the effect of dissuading female participation. In Table 4 of Section 5.3, we found no evidence that the competition treatment was affected by controlling for a measure of opportunity costs, providing some assurance that there is no such demand-side effect.

More generally, any perceptions of costs or benefits that apply asymmetrically to men and women who interact with the treatment effect could alter our results. For example, women in STEM could potentially have expected higher returns for participating in this developmental activity if women have disproportionately higher returns on education, receiving more call-backs and job offers (Williams and Ceci 2015). However, it is not clear that such an effect should necessarily vary with the competition and collaboration treatments. Nonetheless, we cannot unequivocally rule out this possibility.

6.4 Additional Considerations in Interpreting Results

Of course, there are inherent challenges to interpreting studies on gender and competition, including ours. As per the discussion above, there may be multiple nuanced mechanisms at work, even beyond those typically considered within the prior literature (Section 2) or those reviewed here. We, therefore, emphasize the demonstration of fact of (i) smaller gender differences in STEM; and (ii) STEM men and women preferring collaboration over competition, above all.

It is also important to emphasize distinguishing research design choices here, relative to prior research, and their potential implications on our interpretation. For example, the canonical approach in the prior lab experimental research (Section 2) is to recruit lab participants, who then necessarily participate, and then record under which regimes subjects prefer to work. By contrast, here, we randomly assigned each subject to *either* competition or collaboration. Whereas the prior lab experiments effectively measure the percentages of lab subjects who prefer one regime over the other; here, we measure whether individuals choose to participate—at all—conditional on the regime. Therefore, we observe decisions on a different "margin" than lab subjects who rank order which regime they prefer. Therefore, while our reporting of percentage differences (i.e., Panel II in Figures 2 and 3) may allow slightly closer comparison with prior lab studies, these are not the same outcome being observed.

In attempting to translate ideas of competition and collaboration to a field context from the lab, we also differ in how we operationalized these concepts. For example, the prior experimental research (Section

2) tends to focus on small cash payments depending on rank order performance relative to one or few other subjects to operationalize competition. This prior literature has also tended to operationalize collaboration as a matter of determining payoffs in relation to averaged performance across one or few other subjects. Given that competition and collaboration in the field might imply a wide range of things from the internal culture to payoffs, to modes of work organization—we operationalized these concepts by simply making the word "competition" and "collaboration" salient in the treatments.

Related, when attempting to isolate "preferences" for competition, the prior lab experimental literature at times controls for closely related attributes by controlling, for example, particular measures of risk aversion and other attributes. Controlling for these measures in analyzing lab experimental results is typically done with the goal of showing that doing so does not upend results related to competition. While our analysis goes further in controlling for and stratifying for several characteristics of individuals related to knowledge and training, experience, age, and other factors, it should be noted that we did not control for these sorts of behavioral variables.

As reviewed in Section 2.3, there are also inherent limits to studies on gender and competition, even those featuring experiments, in unambiguously establishing the causal effect of gender. This is because it is not possible to randomly assign gender, all-else-being equal. Although we benefitted here from a wide range of controls and tests, and the ability to stratify on especially narrow subsamples, our own study cannot avoid this criticism and should be interpreted with caution, as all studies in this literature. Therefore, strictly speaking, we cannot rule-out the possibility that other factors causing sorting to STEM fields (e.g., confidence, aptitude, interest in puzzle solving, interest in "things," etc.) are imperfectly controlled. This limitation is necessarily true of all research on gender sorting and competition, and might be especially important in highly selected study groups.

Perhaps most fundamentally, there are inherent limitations to making claims about gender differences based on studies on small subsets of humanity. Although our study population is unusually large and widely sampled, and covers adults rather than children or students, it is nonetheless only a population of tens of thousands of individuals, with a relatively narrow range of human experience and circumstances, educated and working in a particular society, in a particular period in history.

Finally, while we chose a product development activity and process that closely maps to such activities in industry, we should underline that our field experimental protocol was not geared to examining effects of variation in work task.

7 Conclusion

Are females more averse to working within competitive organizations than males? We investigated this question by studying the willingness of individuals across all fields to participate in a platform-based product development activity depending on whether individuals were randomly assigned to competition or collaboration treatments. These treatments were based on framing the event in ways that either emphasized competitive or collaborative interactions with other participants.

Our main novel contribution was to present evidence (Section 5) that gender differences in competitiveness differ by field (Sections 5.2). Specifically, we theorized and found evidence that gender differences are smaller in STEM fields than in other fields. In our context, not only were gender differences appreciably smaller in STEM than in non-STEM fields, but these differences were statistically indistinguishable from zero (see Section 6.1 summary). Our results were robust to cohort and age and various controls for factors shaping benefits and costs of participation (Section 5.2 Table 4), and similar patterns were found when examining finer-grained STEM subfields (Section 5.3).

It is also the case that STEM men and women were both more willing to participate under collaboration than under competition. This result is inconsistent with a simplest characterization of sorting on "preferences" for competition, where individuals who simply prefer competition sort into STEM fields. As we discussed in Section 6, these results might reflect a more nuanced combination of, say, tolerance and skills for competing under competition—rather than a simple preference. Further, given that work and problem-solving and training in STEM fields routinely involves both competition and collaboration, it is possible that we might expect a more nuanced, moderated and contingent willingness to work under either competition or collaboration. In our case of an inherently multi-disciplinary product development task, those in STEM perhaps expected superior performance when organizing collaboratively.

These findings begin to outline systematic boundary conditions in relation to the increasingly well-known literature on gender and competition (Section 2.1 and 2.3). While our results contrast with many prior replication studies showing gender differences in competitiveness in smaller studies, our theoretical predictions were in fact derived from generally accepted presumptions of the prior research (Section 2.4). Further, our results are consistent with prior results finding systematic gender differences in lab-based studies if a majority of those subjects came from non-STEM fields of training. Indeed, even in broad study population averages in our study, we found that competition effects were more negative for women. It was only when accounting for fields that we observed striking differences. We detailed caveats and cautions regarding our analysis and emphasized distinctions in our research approach in Section 6.4.

7.1 Organizational Implications

We began the introduction by observing that any behavioral regularities in competitiveness across working-age adults could—if better understood—create opportunities for unlocking the talents of workers through organizational design. Prior research, for example, has shown that in the context of gender differences in a lab experiment, affirmative action can increase female productivity along with representation (Niederle et al. 2013), and priming females subjects or alternative framings of participation choices can also increase female participation under competitive treatments (Balafoutas et al., 2018; He let al., 2019). We ourselves suggested that, in field settings, "tuning" the competitiveness of internal organizational environments might also better harness workers' talents and willingness to participate (Section 1).

Our findings suggest that any practical organizational interventions will need to consider contingencies (across fields, discussed in Section 6.1) and possible nuances (regarding mechanisms, discussed in Section 6.3), as we find evidence contrary to claims of *general* gender differences. For example, our results provided no indication that altering the competitiveness of a work environment in the case of STEM workers will affect gender composition—at least among those already sorted into STEM fields of training and careers. It remains plausible that managing the intensity of the competition faced by boys and girls in STEM education earlier in life could affect sorting into STEM fields and thus affect the composition of the STEM pipeline and workforce (Buser et al. 2014, 2016; Niederle and Vesterlund 2010).

References

Abramo G, D'Angelo CA, Murgia G (2013) Gender differences in research collaboration. *Journal of Informetrics*. 7(4):811–822.

Almås I, Cappelen AW, Salvanes KG, Sørensen EØ, Tungodden B (2016) Willingness to compete: Family matters. *Management Science*. Aug;62(8):2149-62.

Altonji JG, Arcidiacono P, Maurel A 2016 The analysis of field choice in college and graduate school: Determinants and wage effects. *In Handbook of the Economics of Education* Jan 1 (Vol. 5, pp. 305-396). Elsevier.

Andersen S, Ertac S, Gneezy U, List JA, Maximiano S (2013) Gender, competitiveness, and socialization at a young age: Evidence from a matrilineal and a patriarchal society. *Review of Economics and Statistics*. 95(4):1438–1443.

Antonovics K, Arcidiacono P, Walsh R (2009) The effects of gender interactions in the lab and in the field. *Review of Economics and Statistics*. 91(1):152–162.

Apicella CL, Dreber A, Gray PB, Hoffman M, Little AC, Campbell BC. Androgens and competitiveness in men. *Journal of Neuroscience, Psychology, and Economics*. 2011 Feb;4(1):54.

Balafoutas, L, Fornwagner, H and Sutter, M, 2018. Closing the gender gap in competitiveness through priming. *Nature Communications*, 9(1), pp.1-6.

Baumann, O, JP Eggers, N. Stieglitz. "Colleagues and competitors: How internal social comparisons shape organizational search and adaptation." *Administrative Science Quarterly* 64, no. 2 (2019): 275-309.

Boaler J (2016) *Mathematical Mindsets: Unleashing Students' Potential Through Creative Math, Inspiring Messages and Innovative Teaching* (John Wiley and Sons, New York).

Booth A, Nolen P (2012) Choosing to compete: How different are girls and boys?. *Journal of Economic Behavior Organization*. Feb 1;81(2):542-55.

Boschini A, Dreber A, von Essen E, Muren A, Ranehill E (2019) Gender, risk preferences and willingness to compete in a random sample of the Swedish population☆. *Journal of Behavioral and Experimental Economics*. Dec 1;83:101467.

Breda T, Jouini E, Napp C, Thebault G (2020) Gender stereotypes can explain the gender-equality paradox. *Proceedings of the National Academy of Sciences*. 2020 Dec 8;117(49):31063-9.

Brenoe AA, Zölitz U. (2020) Exposure to more female peers widens the gender gap in stem participation. *Journal of Political Economy* 38(4).

Brynjolfsson E, McAfee A (2014) *The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies* (W. W. Norton and Company).

Buckert M, Schwieren C, Kudielka BM, Fiebach CJ (2017) How stressful are economic competitions in the lab? An investigation with physiological measures. *Journal of Economic Psychology*. Oct 1;62:231-45.

Bui Q, Miller C (2018) The age that women have babies: how a gap divides America. *The New York Times*.

Buser T (2009) The Impact of Female Sex Hormones on Competitiveness. Discussion Paper, Tinbergen Institute.

Buser T (2016) The impact of losing in a competition on the willingness to seek further challenges. *Management Science*. 62(12):3439–3449.

Buser T, Niederle M, Oosterbeek H (2014) Gender, competitiveness, and career choices. The Quarterly *Journal of Economics*. 129(3):1409–1447.

Buser T, Niederle M, Oosterbeek H (2020) Can competitiveness predict education and labor market outcomes? Evidence from incentivized choice and survey measures. Evidence From Incentivized Choice and Survey Measures (January 20, 2020).

Buser T, Yuan H (2019) Do women give up competing more easily? Evidence from the lab and the Dutch math Olympiad. *American Economic Journal: Applied Economics*. 11(3):225–252.

Byrnes JP, Miller DC, Schafer WD (1999) Gender differences in risk taking: A meta-analysis. *Psychological Bulletin*. 125(3):360.

Casari M, Ham JC, Kagel JH (2007) Selection bias, demographic effects, and ability effects in common value auction experiments. *American Economic Review*. Sep;97(4):1278-304.

Cech E, Rubineau B, Silbey S, Seron C (2011) Professional role confidence and gendered persistence in engineering. *American Sociological Review*. 76(5):641–666.

Cheryan S, Ziegler SA, Montoya AK, Jiang L (2017) Why are some STEM fields more gender balanced than others? *Psychological Bulletin*. 143(1):1.

Choi KS (2013) Evaluating gender significance within a pair programming context. 2013 46th *Hawaii International Conference on System Sciences*, 4817–4825.

Comeig I, Grau-Grau A, Jaramillo-Gutiérrez A, Ramírez F (2016) Gender, self-confidence, sports, and preferences for competition. *Journal of Business Research*. 69(4):1418–1422.

Croson R, Gneezy U (2009) Gender differences in preferences. *Journal of Economic Literature*. 47(2):448–474.

Dargnies M-P (2012) Men too sometimes shy away from competition: The case of team competition. *Management Science*. 58(11):1982–2000.

Dittrich M, Knabe A, Leipold K (2014) Gender differences in experimental wage negotiations. *Economic Inquiry*. 52(2):862–873.

Dohmen T, Falk A (2011) Performance pay and multidimensional sorting: Productivity, preferences, and gender. *American Economic Review*. Apr;101(2):556-90.

Dreber A, Von Essen E, Ranehill E (2011) Outrunning the gender gap—Boys and girls compete equally. *Experimental Economics.* 14(4):560–582.

Dreber A, von Essen E, Ranehill E (2014) Gender and competition in adolescence: task matters. *Experimental Economics*. Mar 1;17(1):154-72.

Eckel CC, Grossman PJ (2008) Men, women and risk aversion: Experimental evidence. *Handbook of Experimental Economics*. 1:1061–1073.

Etzkowitz H, Kemelgor C, Uzzi B. (2000) *Athena unbound: The advancement of women in science and technology*. Cambridge University Press.

Fernandez, R. M., Campero, S. (2017). Gender sorting and the glass ceiling in high-tech firms. *ILR Review*, 70(1), 73–104.

Fernandez-Mateo I, Kaplan S (2018) Gender and organization science: Introduction to a virtual special issue. *Organization Science*. 29(6):1229–1236.

Flory JA, Gneezy U, Leonard KL, List JA (2018) Gender, age, and competition: A disappearing gap? *Journal of Economic Behavior and Organization.* 150:256–276. https://doi.org/10.1016/j.jebo.2018.03.027

Flory JA, Leibbrandt A, List JA (2015) Do competitive workplaces deter female workers? A large-scale natural field experiment on job entry decisions. *Review of Economic Studies*. 82(1):122–155.

Foor CE, Walden SE, Shehab RL, Trytten DA (2013) "We weren't intentionally excluding them… just old habits": Women (lack of) interest and an engineering student competition team. 2013 *IEEE Frontiers in Education Conference* (FIE), 349–355.

Friesel A, Timcenko O (2011) Competition-motivated teamwork and narratives to motivate girls for engineering. *22nd EAEEIE Annual Conference.*

Gee LK (2019) The more you know: Information effects on job application rates in a large field experiment. *Management Science*. 65(5):2077–2094.

Gneezy U, Leonard KL, List JA (2009) Gender differences in competition: Evidence from a matrilineal and a patriarchal society. *Econometrica*. 77(5):1637–1664.

Gneezy U, Niederle M, Rustichini A (2003) Performance in competitive environments: Gender differences. *Quarterly Journal of Economics*. 118:1049–1074.

Gneezy U, Rustichini A (2004) Gender and competition at a young age. *American Economic Review*. 94(2):377–381.

Goldin C (2014) A grand gender convergence: Its last chapter. *American Economic Review*. 104(4):1091–1119.

Günther C, Ekinci NA, Schwieren C, Strobel M (2010) Women can't jump?—An experiment on competitive attitudes and stereotype threat. *Journal of Economic Behavior and Organization*. 75(3):395–401.

Guo J (2015, January 2) Why do some studies show that women are less competitive than men? *The Washington Post*, https://www.washingtonpost.com/news/storyline/wp/2015/01/02/why-do-some-studies-show-that-women-are-less-competitive-then-men/

Healy A, Pate J (2011) Can teams help to close the gender competition gap? *Economic Journal*. 121(555):1192–1204.

Holland PW (1986) Statistics and causal inference. *Journal of the American Statistical Association.* 81(396):945–960.

Holt-Lunstad J, Clayton CJ, Uchino BN (2001) Gender differences in cardiovascular reactivity to competitive stress: The impact of gender of competitor and competition outcome. *International Journal of Behavioral Medicine*. Jun 1;8(2):91-102.

Hunt A, Samman E. Gender and the gig economy. *ODI Working Paper* 546, 2019.

Inglehart M, Brown DR, Vida M (1994). Competition, achievement, and gender: A stress theoretical analysis. Student Motivation, Cognition, and Learning: *Essays in Honor of Wilbert. J. McKeachie,* 311–330.

Iriberri N, Rey-Biel P (2019) Competitive pressure widens the gender gap in performance: Evidence from a two-stage competition in mathematics. *Economic Journal*. 129(620):1863–1893.

Jetter M, Walker J. (2016) Gender in Jeopardy!: The Role of Opponent Gender in High Stakes Competition. *IZA Discussion Paper* No. 9669, Available at SSRN: https://ssrn.com/abstract=2725030

Lundeberg MA, Fox PW, Punćcohaŕ J (1994) Highly confident but wrong: Gender differences and similarities in confidence judgments. *Journal of Educational Psychology*. Mar;86(1):114.

Jurajda Š, Münich D (2011) Gender gap in performance under competitive pressure: Admissions to Czech universities. *American Economic Review*. 101(3):514–518.

Kahn S, Ginther D 2017 Women and STEM. *National Bureau of Economic Research*; Jun 15.

Kamas L, Preston A (2012) The importance of being confident; gender, career choice, and willingness to compete. *Journal of Economic Behavior and Organization*. 83(1):82–97.

Kesebir S, Lee SY, Elliot AJ, Pillutla MM (2019) Lay beliefs about competition: Scale development and gender differences. *Motivation and Emotion*. 43(5):719–739.

Kivlighan KT, Granger DA, Booth A (2005) Gender differences in testosterone and cortisol response to competition. *Psychoneuroendocrinology*. 30(1):58–71. https://doi.org/10.1016/j.psyneuen.2004.05.009

Kleinjans KJ (2009) Do gender differences in preferences for competition matter for occupational expectations?. *Journal of Economic Psychology*. 2009 Oct 1;30(5):701-10.

Knight J (2002) *Sexual Stereotypes* (Nature Publishing Group).

Kuhn P, Villeval MC (2015) Are women more attracted to co-operation than men? *Economic Journal*. 125(582):115–140.

Kuyath SJ, Yoder L (2004) Diversity in engineering technology: Competitions. *Age*. 9:1.

Kvande E (1999) In the Belly of the Beast' Constructing Femininities in Engineering Organizations. *European Journal of Women's Studies*. Aug;6(3):305-28.

Lazear EP, Malmendier U, Weber RA (2012) Sorting in experiments with application to social preferences. *American Economic Journal: Applied Economics*. Jan;4(1):136-63.

Lee, S, P Puranam. "Incentive redesign and collaboration in organizations: Evidence from a natural experiment." *Strategic Management Journal* 38, no. 12 (2017): 2333-2352.

Liang C, Hong Y, Gu B, Peng, J., 2018. Gender wage gap in online gig economy and gender differences in job preferences. Available at *SSRN* 3266249.

Lippa R (1998) Gender-related individual differences and the structure of vocational interests: The importance of the people–things dimension. *Journal of Personality and Social Psychology*. 74(4):996.

McIlwee JS, Robinson JG (1992) *Women in Engineering: Gender, Power, and Workplace Culture* (SUNY Press).

Moakler Jr MW, Kim MM (2014) College major choice in STEM: Revisiting confidence and demographic factors. *The Career Development Quarterly*. Jun;62(2):128-42.

Morin L-P (2015) Do men and women respond differently to competition? Evidence from a major education reform. *Journal of Labor Economics*. 33(2):443–491.

Mortensen M. (2014) Constructing the team: The antecedents and effects of membership model divergence. *Organization Science*. Jun;25(3):909-31.

Moss-Racusin CA, Sanzari C, Caluori N, Rabasco H (2018) Gender bias produces gender gaps in STEM engagement. *Sex Roles*. 79(11–12):651–600.

Murciano-Goroff R (2018) Missing women in tech: The labor market for highly skilled software engineers. Work in progress. http://Stanford. Edu/~ Ravivmg/Papers/JMP. Pdf.

Niederle M, Vesterlund L (2007) Do women shy away from competition? Do men compete too much? *Quarterly Journal of Economics*. 122(3):1060–1101.

Niederle M, Vesterlund L (2010) Explaining the gender gap in math test scores: The role of competition. *Journal of Economic Perspectives*. 24(2):129–144.

Niederle M, Vesterlund L (2011) Gender and competition. *Annu. Rev. Econ.* 3(1):601–630.

Notter KB (2010) Is competition making a comeback? Discovering methods to keep female adolescents engaged in STEM: A phenomenological approach.

Nosek BA, Smyth FL, Sriram N, Lindner NM, Devos T, Ayala A, Bar-Anan Y, Bergh R, Cai H, Gonsalkorale K, Kesebir S. (2009) National differences in gender–science stereotypes predict national sex differences in science and math achievement. *Proceedings of the National Academy of Sciences*. Jun 30;106(26):10593-7.

Obloj, T, Zenger, T., 2017. Organization design, proximity, and productivity responses to upward social comparison. *Organization Science*, 28(1), pp.1-18.

Ong M (2005) Body projects of young women of color in physics: Intersections of gender, race, and science. *Social problems*. Nov 1;52(4):593-617.

Ors E, Palomino F, Peyrache E (2013) Performance gender gap: Does competition matter? *Journal of Labor Economics*. 31(3):443–499.

Oster E (2019) Unobservable selection and coefficient stability: Theory and evidence. *Journal of Business and Economic Statistics*. 37(2):187–204.

Paserman MD (2010) Gender Differences in Performance in Competitive Environments? Evidence from Professional Tennis Players. *NBER Working Paper* 27154.

PBS News Hour (2015) Why we think the way we do about men, women and work | *PBS NewsHour*. https://www.pbs.org/newshour/nation/the-lasting-impact-of-the-plow-on-our-attitudes-about-gender-and-work

Powell A, Bagilhole B, Dainty A (2009) How women engineers do and undo gender: Consequences for gender equality. *Gender, work organization*. Jul;16(4):411-28.

Pregaldini D, Backes-Gellner U, Eisenkopf G (2020) Girls' preferences for STEM and the effects of classroom gender composition: New evidence from a natural experiment. *Journal of Economic Behavior and Organization*. 178:102–123.

Price CR (2011) Gender, competition, and managerial decisions. *Management Science*. 58(1):114–122. https://doi.org/10.1287/mnsc.1110.1384

Puranam, P. (2018) *The microstructure of organizations*. Oxford University Press.

Reuben E, Sapienza P, Zingales L (2015) Taste for competition and the gender gap among young business professionals. *NBER Working Paper* No. 21695.

Saccardo S, Pietrasz A, Gneezy U (2017) On the size of the gender difference in competitiveness. *Management Science*. 64(4):1541–1554. https://doi.org/10.1287/mnsc.2016.2603

Samek A (2019) Gender differences in job entry decisions: A university-wide field experiment. *Management Science*. 65(7):3272–3281.

Shurchkov O (2012) Under pressure: Gender differences in output quality and quantity under competition and time constraints. *Journal of the European Economic Association*. 10(5):1189–1213.

Silbey SS (2015) The elephant in the room: Constraints and consequences of a four-year undergraduate engineering degree. *Engineering Studies*. 7(2–3):164–160.

Silbey SS (2016) Why do so many women who study engineering leave the field. *Harvard Business Review. 1–2.*

Slack Website. (2019). *In good company: How to create a culture of collaboration at work*. https://slackhq.com/culture-collaboration-work

Su R, Rounds J, Armstrong PI (2009) Men and things, women and people: a meta-analysis of sex differences in interests. *Psychological bulletin*. Nov;135(6):859.

Sutter M, Glätzle-Rützler D (2015) Gender differences in the willingness to compete emerge early in life and persist. *Management Science*. 61(10):2339–2354.

van Tuijl C, van der Molen JH 2016 Study choice and career development in STEM fields: an overview and integration of the research. *International journal of technology and design education*. May 1;26(2):159-83.

Vugt M van, Cremer DD, Janssen DP (2007) Gender differences in cooperation and competition: The male-warrior hypothesis. *Psychological Science*. 18(1):19–23.

Wieland A, Sarin R (2012) Domain specificity of sex differences in competition. *Journal of Economic Behavior and Organization*. 83(1):151–157.

Williams WM, Ceci SJ (2015) National hiring experiments reveal 2: 1 faculty preference for women on STEM tenure track. *Proceedings of the National Academy of Sciences*. Apr 28;112(17):5360-5.

Woodcock, A., Hernandez, P. R., Estrada, M., Schultz, P. (2012). The consequences of chronic stereotype threat: domain disidentification and abandonment. *Journal of personality and social psychology*, 103(4), 635.

Wozniak D, Harbaugh WT, Mayr U (2014) The menstrual cycle and performance feedback alter gender differences in competitive choices. *Journal of Labor Economics*. 32(1):161–198.

Wuchty S, Jones BF, Uzzi B (2007) The increasing dominance of teams in production of knowledge. *Science*. May 18;316(5827):1036-9.

Wynn, A. T., Correll, S. J. (2017). Gendered perceptions of cultural and skill alignment in technology companies. *Social Sciences*, 6(2), 45.

Zafar B 2013 College major choice and the gender gap. *Journal of Human Resources*. Jul 1;48(3):545-95.

Tables

Table 1 Descriptive Statistics of Main Variables

Variable	Mean	Std. Dev.	Description
Participation	0.03	0.17	An indicator variable that is switched to one for those subjects who chose to participate
Competition	0.50	0.50	An indicator variable that is switched to one for those randomly assigned to the competitive treatment (as opposed to the collaborative comparison group)
Female	0.44	0.50	An indicator variable that is switched to one for female subjects
STEM	0.33	0.47	An indicator variable that is switched to one for those in Computer Science, Engineering, or Sciences
Student	0.08	0.27	An indicator variable that is switched to one for subjects currently enrolled as students
CohortYear	1998.7	16.0	Year of matriculation
GPA	3.01	1.17	Cumulative undergraduate college grade point average
Day	30.5	17.3	Count variable from 1 to 60 for each of the days in which invitations were sent (not including weekends and
Engineering	20%	40%	Indicator variable for Business major
Law	4%	19%	Indicator variable for Computer Science major
Computer Science	4%	19%	Indicator variable for Design & Media major
Health & Nursing	17%	37%	Indicator variable for Engineering major
Design & Media	7%	25%	Indicator variable for Health & Nursing major
Business	25%	43%	Indicator variable for Humanities major
Humanities	14%	35%	Indicator variable for Law major
Sciences	9%	28%	Indicator variable for Sciences major

Note. Num Obs. = 97,678 for all variables with the exception of *CohortYear* for which we have 90,087 observations; data for *GPA* were available for only a subset of 13,926 observations in post-2009 data.

Table 2 Balance of Assignment Groups

	Competition Treatment		Collaboration Comparison Group	
Variable	Mean	Std. Dev.	Mean	Std. Dev.
Female	44.63%	49.71%	44.33%	49.68%
STEM	32.68%	46.91%	32.68%	46.91%
Student	7.88%	26.95%	7.65%	26.57%
CohortYear	1998.8	16.0	1998.7	16.0
GPA	3.01	1.17	3.00	1.18
Day	30.5	17.3	30.4	17.3
Engineering	20.44%	40.32%	20.17%	40.13%
Law	3.70%	18.87%	3.76%	19.02%
Computer Science	3.65%	18.76%	3.62%	18.67%
Health & Nursing	16.82%	37.40%	16.95%	37.52%
Media & Design	6.67%	24.95%	6.57%	24.77%
Business	24.74%	43.15%	25.02%	43.31%
Humanities	14.22%	34.93%	13.98%	34.68%
Sciences	8.59%	28.03%	8.89%	28.46%

Table 3 Baseline OLS Model Estimates of Participation Rates for Study Population-Averages

Dep. Var.:	*Participation* [Percentage Points]								
Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Gender	Treatment	Interaction	Day Controls	Student Control	Cohort Controls	Subsample with GPA	GPA Controls	All Controls
Female	-0.67***	-0.67***	-0.67***	-0.67***	-0.81***	-1.04***	-1.15***	-1.13***	-1.31***
	(0.11)	(0.11)	(0.16)	(0.16)	(0.16)	(0.15)	(0.28)	(0.28)	(0.28)
Competition		-0.71***	-0.71***	-0.71***	-0.73***	-0.61***	-1.22***	-1.22***	-1.09***
		(0.11)	(0.15)	(0.15)	(0.15)	(0.14)	(0.27)	(0.27)	(0.27)
Female × Competition			0.00	-0.01	0.02	-0.06	0.30	0.30	0.19
			(0.21)	(0.21)	(0.21)	(0.20)	(0.37)	(0.37)	(0.37)
Student					6.75***				2.95***
					(0.33)				(0.50)
CohortYear						0.078***			-0.09***
						(0.00)			(0.01)
GPA								-0.16*	-.030
								(0.08)	(0.08)
Day FEs				Y					Y
Constant	3.19***	3.54***	3.54***		3.09***	-151.89***	4.35***	4.81***	
	(0.08)	(0.10)	(0.11)		(0.11)	-(5.83)	(0.21)	(0.33)	
Adjusted R^2	0.000	0.001	0.001	0.001	0.012	0.011	0.002	0.002	0.008

Note. OLS linear probability model coefficients are reported as percentage points. Coefficients reported as percentage points (i.e., 100 percent = 100). Heteroskedasticity-robust standard errors reported in brackets. *** p<0.01, ** p<0.05, * p<0.1.

Table 4 OLS Model Estimates of Participation Rates, Controlling for Fields

Dep. Var.:	*Participation* [Percentage Points]			
Model:	(1)	(2)	(3)	(4)
	Population Avg.	Field Controls	Population Avg.	Field Controls
Female	-0.67***	0.25**	-0.67***	0.40**
	(0.11)	.12	(0.16)	(0.17)
Competition			-0.71***	-0.43***
			(0.15)	(0.15)
Female × Competition			0.00	-0.29
			(0.21)	(0.21)
Computer Science		7.03***		7.25***
		(0.43)		(0.44)
Engineering		3.97***		4.18***
		(0.14)		(0.16)
Sciences		2.98***		3.18***
		(0.20)		(0.22)
Business		2.75***		2.96***
		(0.11)		(0.14)
Design & Media		2.59***		2.80***
		(0.22)		(0.24)
Humanities		2.07***		2.28***
		(0.14)		(0.16)
Law		1.43***		1.63***
		(0.22)		(0.23)
Health & Nursing		1.06***		1.27***
		(0.13)		(0.16)
Constant	3.19***		3.54***	
	(0.08)		(0.11)	
Adjusted R^2	0.000	0.030	0.000	0.030

Note. OLS linear probability model coefficients are reported as percentage points. Coefficients reported as percentage points (i.e., 100 percent = 100). Heteroskedasticity-robust standard errors reported in brackets. *** p<0.01, ** p<0.05, * p<0.1.

Table 5 OLS Model Estimates of Participation Rates Stratified by STEM and non-STEM

Dep. Var.:	*Participation* [Percentage Points]			
Model:	(1)	(2)	(3)	(4)
		Non-STEM Subset	STEM Subset	Pooled Interactions
Female	0.40**	0.20	1.13***	0.20
	(0.17)	(0.19)	(0.40)	(0.19)
Competition	-0.43***	-0.11	-0.87***	-0.11
	(0.15)	(0.18)	(0.26)	(0.18)
Female × Competition	-0.29	-0.47**	-0.32	-0.47**
	(0.21)	(0.23)	(0.50)	(0.23)
STEM Field × Female				0.93**
				(0.44)
STEM Field × Competition				-0.77**
				(0.32)
STEM Field × Female × Competition				0.16
				(0.55)
Field FEs	Y	Y	Y	Y
Adjusted R^2	0.03	0.02	0.05	0.03

Note. OLS linear probability model coefficients are reported as percentage points. Coefficients reported as percentage points (i.e., 100 percent = 100). Heteroskedasticity-robust standard errors reported in brackets. *** p<0.01, ** p<0.05, * p<0.1. Note: Among the 97,678 observations, 31,923 are associated with those trained in STEM fields and 65,755 are associated with non-STEM fields.

Table 6 Robustness Tests

Dep. Var.:	*Participation* [Percentage Points]				
Model:	(1)	(2)	(3)	(5)	(6)
	Main Result	Cohort	Opportunity Costs	GPA Subsample	GPA
Female	0.20	-0.10	0.01	-0.17	0.79
	(0.19)	(0.19)	(0.20)	(0.35)	(0.60)
Competition	-0.11	-0.01	-0.01	-0.40	-0.42
	(0.18)	(0.18)	(0.18)	(0.34)	(0.34)
Female × Competition	-0.47**	-0.57**	-0.58**	-0.73*	-0.71
	(0.23)	(0.23)	(0.23)	(0.44)	(0.44)
STEM Field × Female	0.93**	0.44	0.38	-0.89	-0.87
	(0.44)	(0.44)	(0.44)	(0.66)	(0.66)
STEM Field × Competition	-0.76**	-0.84***	-0.83***	-1.33**	-1.31**
	(0.32)	(0.31)	(0.31)	(0.56)	(0.56)
STEM Field × Female × Competition	0.15	0.20	0.22	2.10**	2.09**
	(0.55)	(0.55)	(0.55)	(0.84)	(0.84)
Childrearing Years			-0.98***		
			(0.22)		
Childrearing Year × Female			-0.43		
			(0.27)		
GPA					0.12
					(0.12)
GPA × Female					-0.33**
					(0.16)
Age		-0.09***	-0.09***	-0.08***	-0.08***
		(0.00)	(0.00)	(0.01)	(0.01)
Field FEs	Y	Y	Y	Y	Y
Adjusted R^2	0.03	0.04	0.04	0.04	0.04
No. Obs.	97,696	97,696	97,696	36,345	36,345

Note. OLS linear probability model coefficients are reported as percentage points. Coefficients reported as percentage points (i.e., 100 percent = 100). Heteroskedasticity-robust standard errors reported in brackets. *** p<0.01, ** p<0.05, * p<0.1.

Table 7 OLS Model Estimates of Participation Rates, Separately Estimated on STEM Subfields

Dep. Var.:		*Participation* [Percentage Points]							
		Female		*Competition*		*Female* × *Competition*		*Constant*	
	No. Obs.	coeff.	s.e.	coeff.	s.e.	coeff.	s.e.	coeff.	s.e.
Behavioral Neuroscience	603	1.49	(2.01)	-0.95	(1.98)	-0.77	(2.56)	2.20	(1.54)
Biology	2420	-0.79	(0.96)	-2.26	(0.84)	1.55	(1.15)	3.13	(0.75)
Chemical Eng	1337	1.11	(1.48)	-0.03	(0.88)	1.20	(2.24)	1.94	(0.61)
Chemistry	782	1.57	(1.78)	0.45	(1.37)	-1.58	(2.39)	1.95	(0.87)
Civil Eng	2567	3.52	(1.65)	-0.61	(0.45)	-2.16	(1.99)	1.42	(0.36)
Computer Eng	1621	1.10	(2.59)	-0.87	(1.26)	-4.98	(2.86)	5.98	(0.93)
Computer Science	3662	0.87	(1.55)	-2.24	(0.99)	-0.84	(2.00)	8.11	(0.75)
Electrical Eng	4756	1.64	(1.31)	-0.86	(0.47)	-0.60	(1.70)	2.80	(0.36)
Eng Mgmt	680	2.21	(3.18)	-0.72	(1.80)	-0.27	(4.21)	4.45	(1.32)
Environmental Sciences	298	2.12	(3.23)	-1.57	(2.48)	-2.19	(3.74)	2.94	(2.06)
Industrial Eng	1172	2.71	(2.14)	0.79	(1.16)	0.62	(3.29)	2.72	(0.78)
Information Systems	1156	-0.91	(2.00)	0.99	(1.94)	-2.48	(2.74)	5.97	(1.33)
Mathematics	770	-2.49	(1.47)	-0.94	(1.64)	6.40	(2.89)	3.59	(1.25)
Mechanical Eng	4418	0.56	(1.17)	-0.20	(0.51)	0.02	(1.61)	2.64	(0.37)
Physics	469	-4.26	(1.48)	-2.65	(1.74)	4.87	(2.81)	4.26	(1.48)
Psychology	2121	-0.82	(1.08)	-1.90	(1.11)	1.38	(1.31)	2.89	(0.95)
Telecommunications Eng	505	-3.18	(3.69)	3.32	(3.64)	-1.73	(5.77)	10.40	(2.33)

Note. The table reports percent participation changes in the competition treatment for both men and women, stratified by STEM subfield. Statistical significance is reported as *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Figures

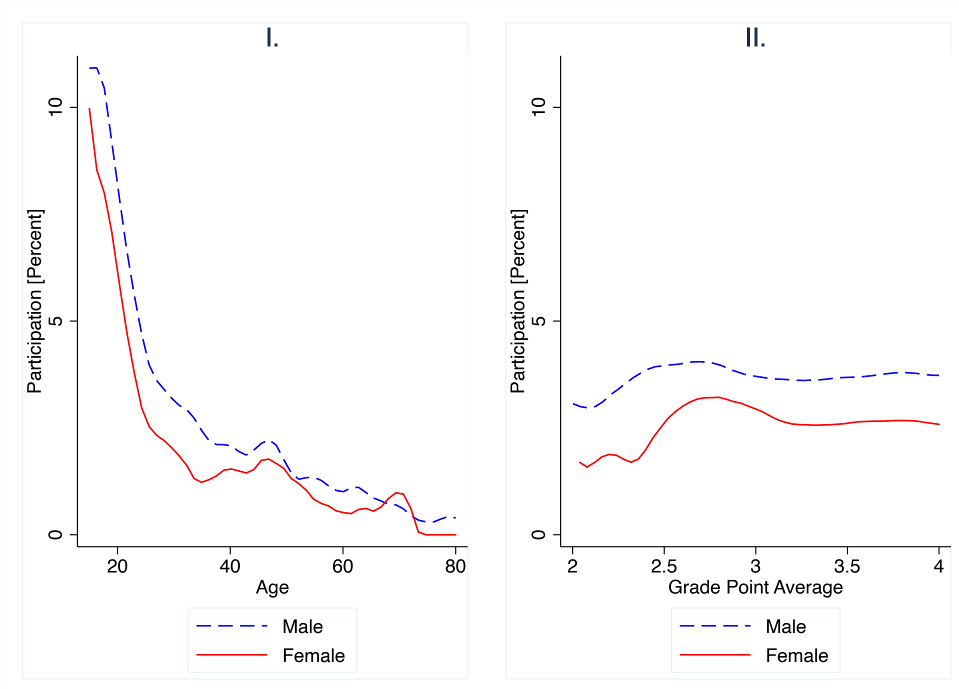

Figure 1 Gender Differences in Participation by Age and Academic Performance

Note. In Panel I, age is used rather than graduation year in the interest of more direct interpretation. Age is based on graduation year and approximated according to the approximation that graduation occurs at the age of 22 or thereabouts. Panel III summarizes differences in participation across males and females in relation to GPA with a quadratic fitted curve.

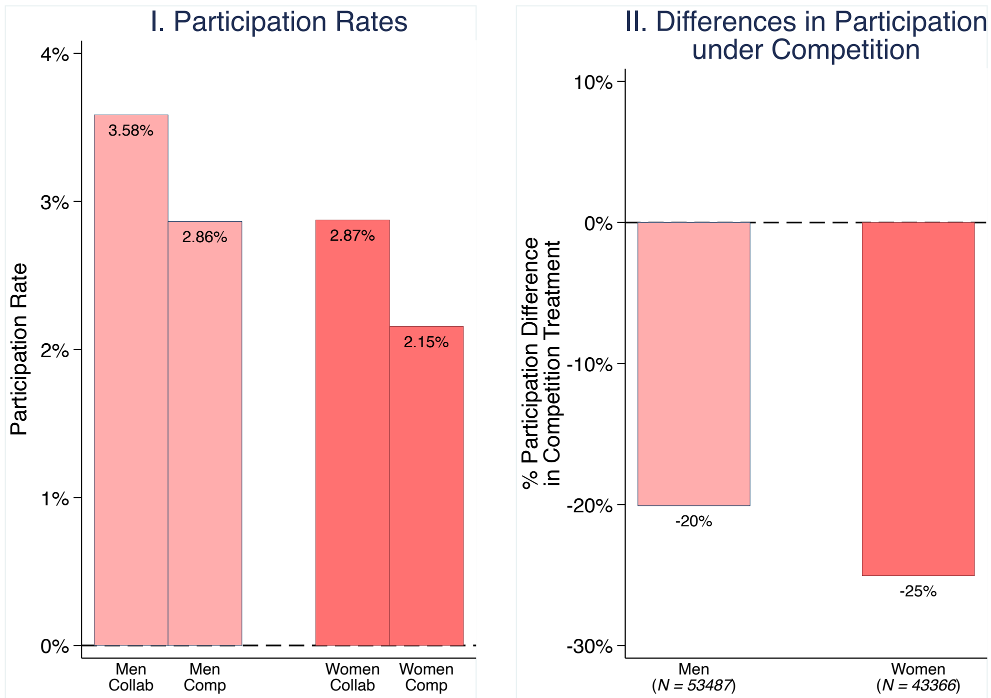

Figure 2 Differences in Participation and Response to Treatment by Sex, Overall Averages

Note. Panel I is the graphical equivalent of regression results presented in model (3) of Table 3. Panel II presents the equivalent information in terms of percentage differences in participation under competition, i.e.,: $\frac{Participation(Competition)-Participation(Collaboration)}{Participation(Collaboration)}$. The differences between percentage differences in responses of men and women in Panel II is weakly significant at $p = 0.097$, based on a non-linear Wald-test of their equivalence.

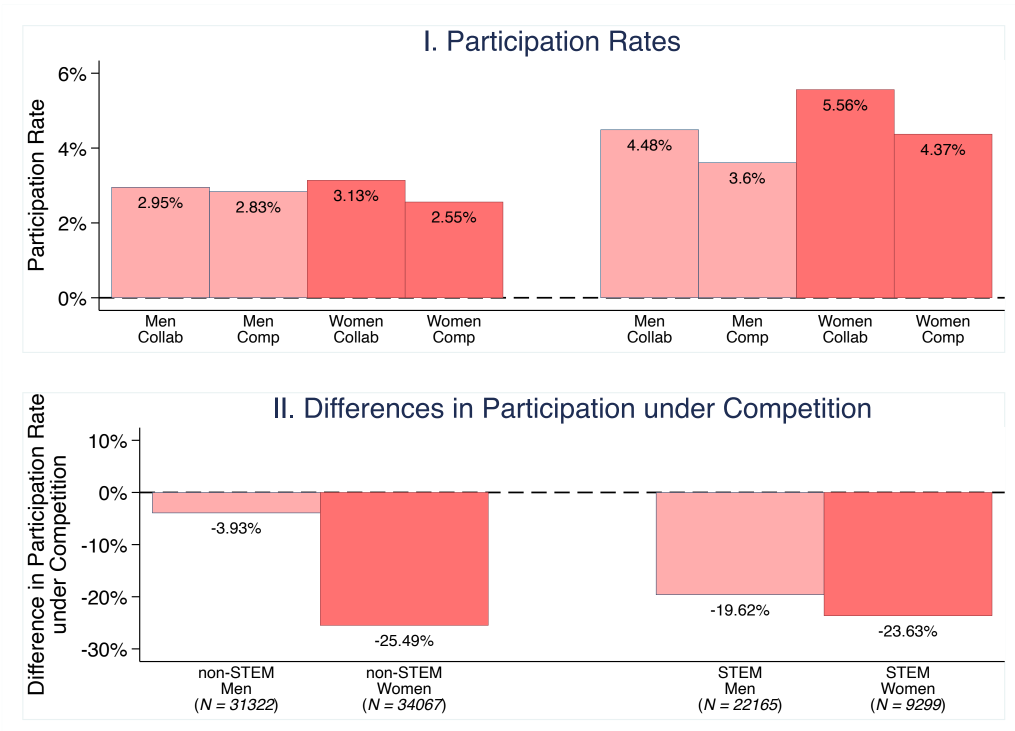

Figure 3 Treatment Effects, Conditional on Field, by STEM and Non-STEM

Note. Panel I is the graphical equivalent of regression results presented in models (2) and (3) or, equivalently, model (4) of Table 6. Panel II presents the equivalent information in terms of percentage differences in participation under competition, i.e.,: $\frac{Participation(Competition)-Participation(Collaboration)}{Participation(Collaboration)}$. The difference in non-STEM men's and women's percentage differences between competition and collaboration (i.e., Panel II) is significant at $p = 0.01$, based on a non-linear Wald-test of their equivalence. This gender difference in percentage differences STEM fields is significant at $p = 0.73$—or, not statistically significant in the least.

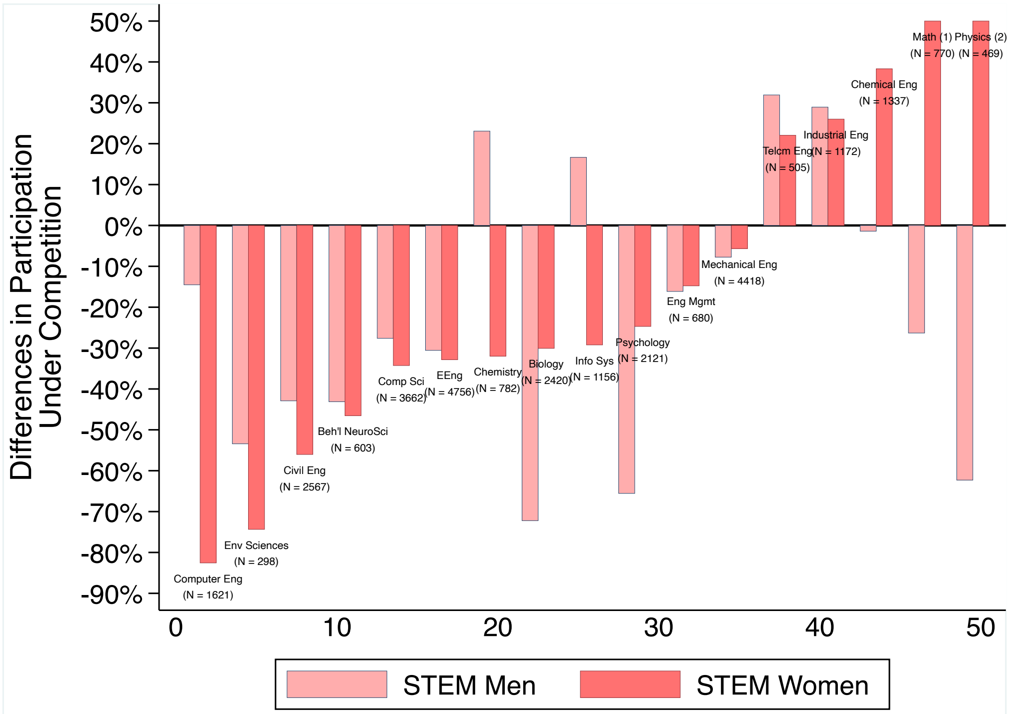

Figure 4 Percent Difference in Participation Under Competition Treatment in STEM Subfields

Note. The graph presents percent differences in participation between those in the competition treatment and those in the collaboration comparison group for both men and women, i.e.,: $\frac{Participation(Competition) - Participation(Collaboration)}{Participation(Collaboration)}$, stratified by both sex and STEM subfields. Only gender differences for computer science are found to be statistically significant at conventional levels (*p = 0.01*), based on nonlinear Wald test of $H_0\colon \frac{\beta_{Competition}}{constant} = \frac{\beta_{Competition} + \beta_{Female \times Competition}}{constant + \beta_{Female}}$, where $Participation_i = constant + \beta_{Female \times Compete} \cdot \text{Female}_i \times \text{Compete}_i + \beta_{Female} \cdot \text{Female}_i + \beta_{Compete} \cdot \text{Compete}_i + \varepsilon_i$. The estimated percent change for females in Physics (1) and Mathematics (2) are truncated in the graph. That for Math is 497 percent (p = 0.27) and that for Physics (2) is undefined (zero women participated in the collaboration treatment).

Appendix A: Examples of Studies on Gender and Competition Reporting Contradictory Findings

As noted in Section 2.2, one challenge to extrapolating beyond existing results suggesting that females might be relatively averse to competition is a growing number of published results that run counter to these original findings. This appendix provides a brief review to illustrate the important contributions of this research, while also highlighting how it is difficult to derive a simple set of lessons from these studies.

For example, whereas Gneezy et al.'s (2003) seminal study found that 30 male students solving mazes at an Israeli university were more positively affected by competition than 30 female Israeli university students who were recruited to the experiment, Günther et al. (2010) find similar results for solving similar maze problems but not for problems involving writing out words beginning with a certain letter or for memory games in a group of Spanish students. Günther et al. (2010) interpret their results as possibly explained by a theory of "stereotype threat," whereby gendered associations of tasks negatively impact task performance of the negatively stereotyped group.

However, whereas Gneezy and Rustichini's (2004) highly influential study of 140 fourth grade Israeli schoolchildren finds that the boys ran faster when facing competition, Dreber et al. (2011) find no such gender differences in a group of 149 Swedish school children in either running, dancing, or skipping rope—clearly gender stereotyped activities. We might then perhaps consider possible differences between Swedish versus Israeli culture. However, Boschini et al. (2018) had a market research firm randomly telephone Swedish adults, performing laboratory-style verbal and math tasks to assess the effects of competition; collecting the responses of 997 Swedes between the ages of 18 and 73, they find some evidence of gender differences in responses to competition in the case of the math task while controlling for income, age, and level of education.

Gneezy et al. (2009) further examine the possible role of cultural context, recruiting 155 people from the Maasai in Tanzania and the Khasi in India, asking them to throw tennis balls into a bucket, choosing either competitive or noncompetitive monetary regimes, i.e., the Niederle-Vesterlund protocol. The Maasai men more frequently chose competition, whereas the Khasi women more frequently chose competition. The authors interpret the results as evidence that gender differences are tied to societal culture. They speculate further that gender differences in prior research (i.e., Sections 2.1 and 2.3) could be a fixture of Western societies.

Other results suggest that narrower socialized differences could shape gender differences. Andersen et al. (2013) recruit 318 Indian school children, 9–15 years old, to similarly throw tennis balls into a bucket in a Niederle-Vesterlund protocol. Most subgroups (by gender, age, and village type) do not statistically differ, when stratifying the data by sex, village, and for ages 9–12 or 13–15. They find differences between boys and girls in the case of patriarchal villages for the 13–15 group. The authors interpret the results to mean that gender differences emerge only in the older girls within patriarchal societies are socialized for some years. In other research, Gneezy and Rustichini (2004) report gender differences occurring before the age of 13.

(A series of findings on gender differences in auctions, where female subjects tend to bid higher than males (e.g., Casari et al. 2007), also appears to run counter to the findings in research on gender and competition.)

Appendix B: Product Development Platform

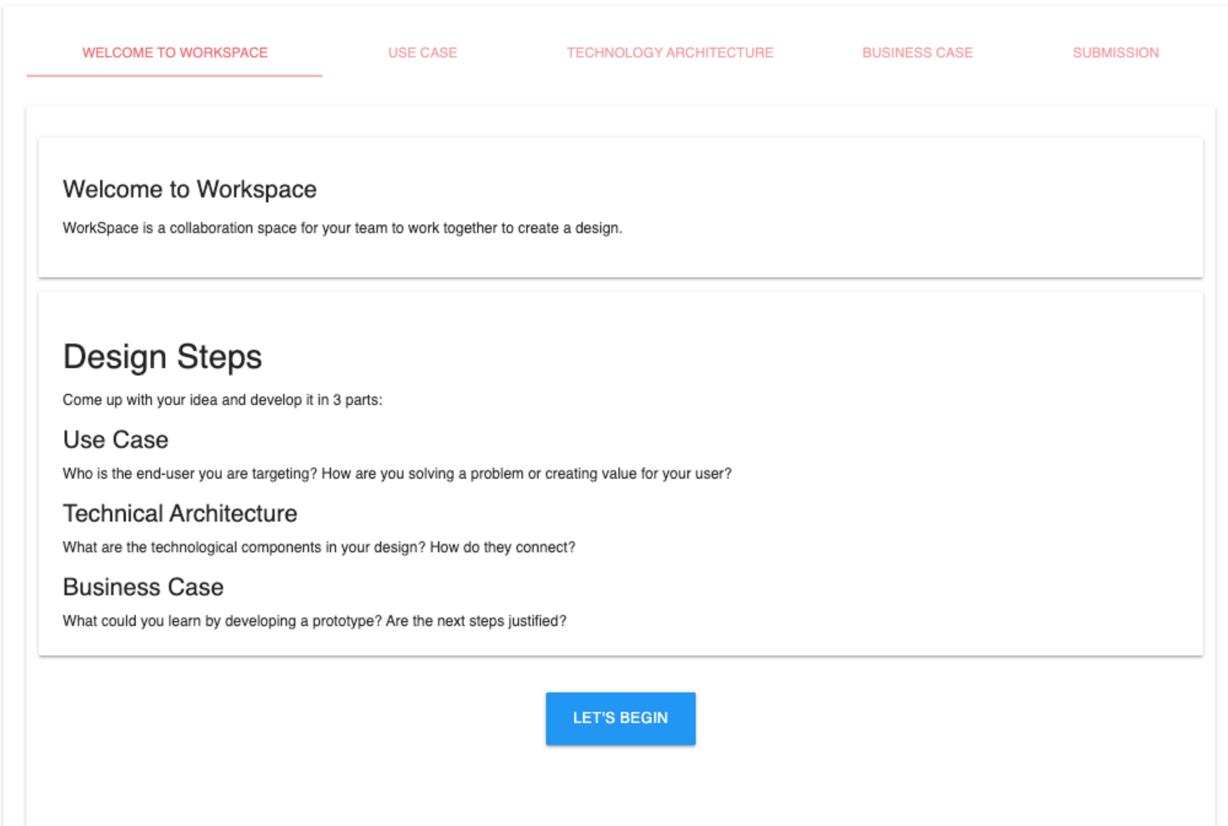

Fig. A1 Description of IoT Project Work: Design of Use Case, Technology Architecture, and Business Case for Prototype

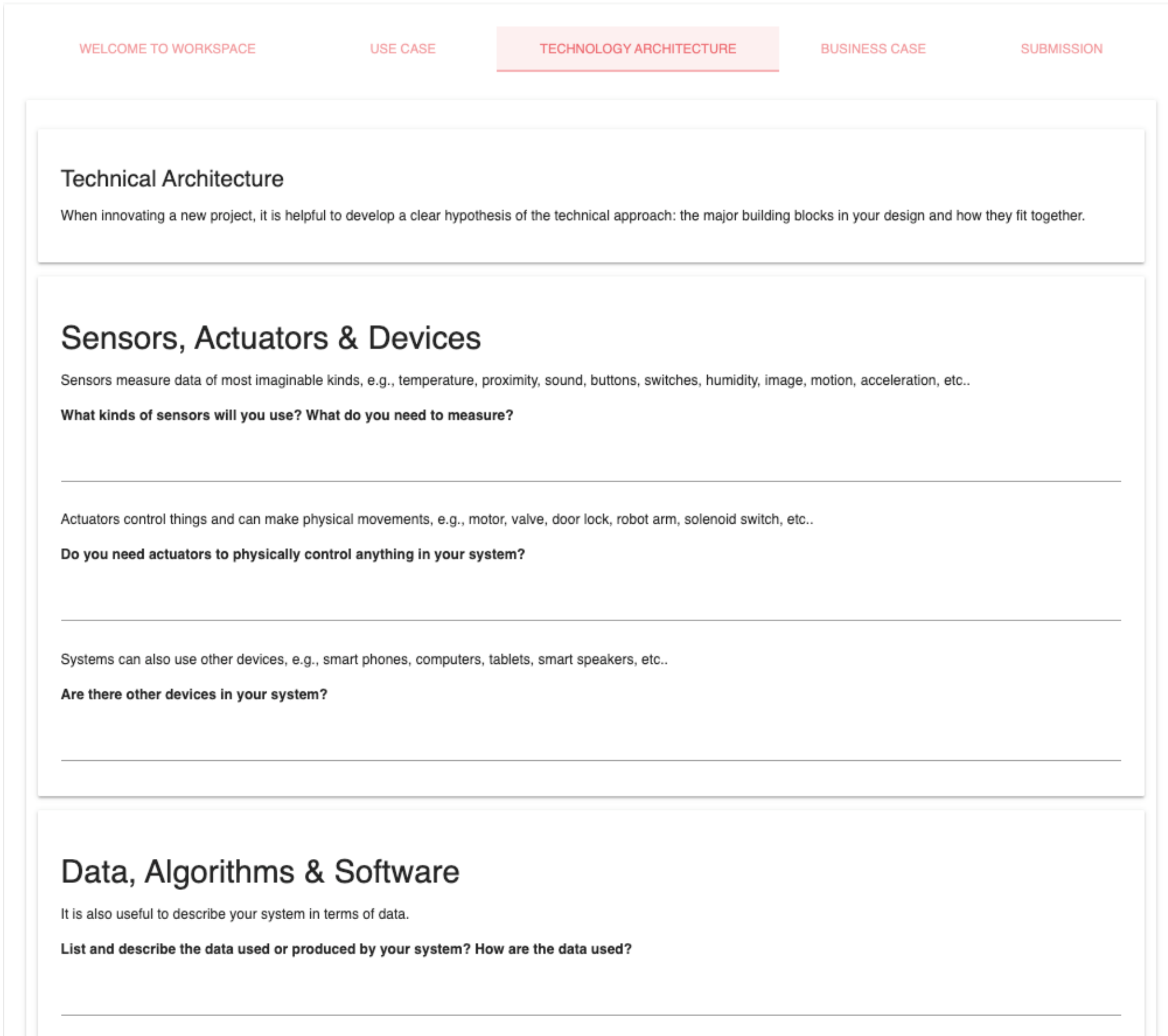

Fig. A2 Sample Section of Technology Architecture Design Steps

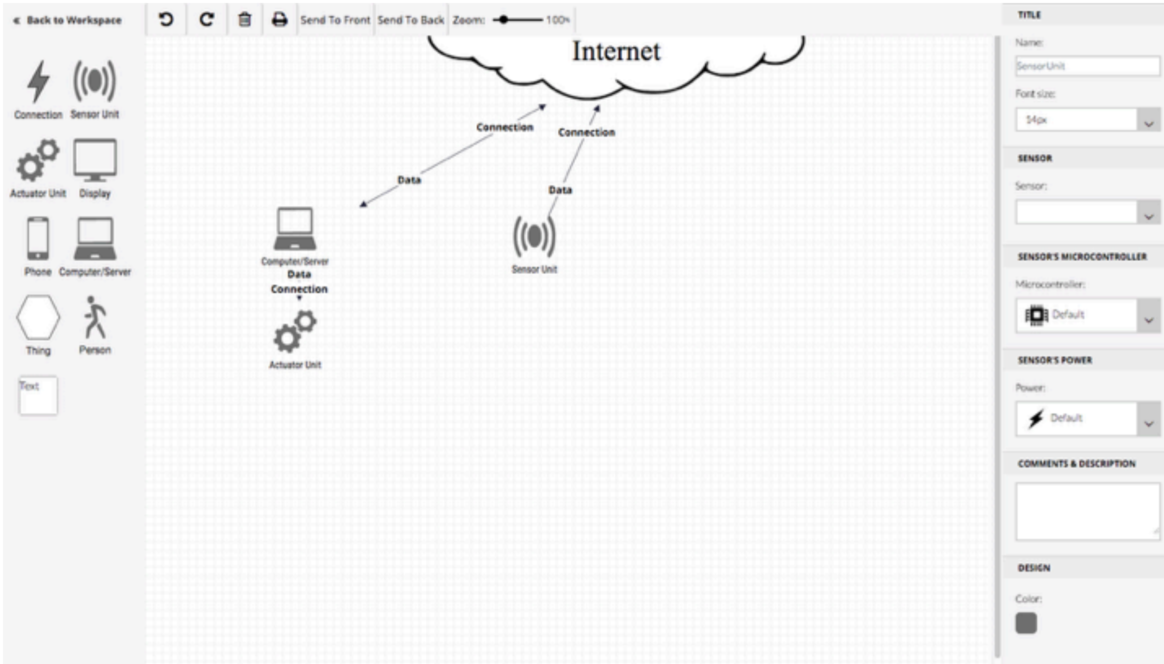

Fig. A3 Sample Technology Design Drag-and-Drop Design Tool