"No Man is an Island": An Empirical Study on Team Formation and Performance

Alessandra Allocca (LMU Munich)

Discussion Paper No. 389

February 21, 2023

"No Man is an Island": An Empirical Study on Team Formation and Performance

Alessandra Allocca*

February 21, 2023.

Abstract

Do self-formed teams perform better than other team structures? Using unique data from Virgo, a Nobel-prize-winning scientific organization with self-formed teams, first, I uncover new evidence on team formation and performance. Then, I develop a structural model to i) estimate which teams perform better controlling for self-formation and ii) evaluate the performance of counterfactual team structures. Regarding i), estimation results show that small teams perform better than large teams. Regarding ii), counterfactual results show that randomly formed teams perform worse than the observed self-formed teams, and teams with a more diverse membership perform better.

Keywords: Teamwork, Entry Game, Structural Estimation, Knowledge Production, Organizational Economics.
JEL Codes: C57, C72, L2, M50, O32.

*LMU Munich, Akademiestrasse 1, 80799 Munich, Germany (alessandra.allocca@econ.lmu.de). This paper is based on the first chapter of my Ph.D. dissertation at the University of Mannheim. I am grateful to my Ph.D. advisors Michelle Sovinsky, Laura Grigolon, and Emanuele Tarantino. Special thanks to Francesco Paolo Conteduca, Marie-Louise Arlt, Luca Braghieri, Mathias Bühler, Aureo de Paula, Sarah Eichmeyer, Florian Englmaier, Miguel Espinosa, Antoine Ferey, Anna Gumpert, Ingrid Haegele, Nicola Lacetera, Martina Magli, Lorenzo Magnolfi, Robin Mamrak, Felix Montag, Andrea Prat, Stephen Ryan, Alina Sagimuldina, Monika Schnitzer, Xuan Teng, Catherine Thomas, Otto Tovainen, Timothy Van Zandt, Fabian Waldinger, and Erina Ytsma for their helpful comments. I also benefited from discussions with Annalisa Allocca, Franco Esteban Cattaneo, Bernhard Ganglmair, Niccolò Lomys, Daniel Savelle, and with seminar participants and discussants at Mannheim, Vienna, Oslo, Carlos III, Amsterdam, Naples Federico II, CREST, LMU Munich, MPI Munich, Bristol, DICE, Bank of Italy, Milano Bicocca, WU Vienna, XXIX School in Economic Theory, MACCI IO day, SAET Conference, JEI, RISE Workshop, Econometric Society Winter Meeting, Economics of Innovation Workshop, ORG Seminar, BGSE Summer Forum, VfS Org Econ Conference, Economics of Teams and Organizations Workshop, CEPR IMO & ENT, ECOBUS. I thank Federico Ferrini for the opportunity of using the data and James Lloyd for excellent research assistance. Financial support by Deutsche Forschungsgemeinschaft through CRC TRR 190 (project number 280092119) is gratefully acknowledged.

> No man is an Iland, intire of itselfe;
> every man is a peece of the
> Continent, a part of the maine [...].
>
> John Donne - 1624

Teams are fundamental for the success of organizations. They may be formed in different ways. In scientific institutions, researchers are usually free to choose their teams of co-authors (Jones, 2021; Guimera et al., 2005; Wuchty, Jones, and Uzzi, 2007). In companies, the management generally determines the composition of teams (Katzenbach and Smith, 2015). However, this is changing: several big firms, such as Google, ING, and IBM, have recently granted their employees flexibility in choosing their working conditions, projects, and teams (e.g., side project time, agile business practices, open workflows).[1]

The increasing relevance of self-formed teams naturally raises a question: how do they perform relative to other team structures? The question hinges on a crucial trade-off. If individuals can choose their teams, their choice might be driven by utility-maximization considerations not aligned with management's objectives. At the same time, they may have access to better information on how to form teams efficiently, and this information might be hard to acquire for the management. Hence, the answer is unclear.

In this paper, I empirically address the above question by exploiting a novel data source from a knowledge production institution, Virgo. Virgo is an institution of about 200 scientific researchers that proved successful in detecting gravitational waves. The founders of Virgo (joint with LIGO, the corresponding experiment in the U.S.) received the Nobel Prize in Physics in 2017.[2] Virgo represents an ideal framework to study team formation because it relies on projects carried out by self-formed teams who report their activities in a diary of work, the Virgo Logbook. The Logbook contains detailed information on the timestamp, name, outcome, teams of researchers

1. Side project time: https://builtin.com/software-engineering-perspectives/20-percent-time. Agile: VersionOne, C. (2020). 14th annual state of the agile report. Open work ows: *Let Employees Choose When, Where, and How to Work*, Harvard Business Review, N. Koloc, 2014.

2. http://www.virgo-gw.eu/. The scienti c structure of Virgo is similar to other major scienti c experiments, like CERN (https://home.cern/science/experiments).

participating, and other project characteristics.

To exploit this rich source, I collect and transform the unstructured textual data of the Logbook into a machine-readable dataset containing information on projects and researchers. With these data, I first uncover new descriptive evidence on teams in science. Self-formed teams are a relevant part of the institution: of about 3,000 projects carried out at Virgo between 2012 and 2016, 66% are team projects, with an average size of 2.3 people. Teams differ in members' characteristics: 60% of the researchers hold a degree in Physics and the remaining in Engineering and other technical fields. Almost 80% of the researchers have a junior position, and only 20% have a senior position. In terms of project outcome, I show that the average project completion is 51%. Moreover, researchers specialized in Physics are associated with a lower probability of project completion relative to those specialized in Engineering.

By looking at this evidence, one might be tempted to conclude that specific characteristics of the researchers cause better performance. However, this conclusion does not consider how teams are formed. To account for this, I develop a two-stage structural model to i) estimate the team performance controlling for team self-formation and ii) determine whether self-formed teams perform better than other team structures.

The first stage (*participation stage*) is an entry game with incomplete information. Similar models have been empirically estimated by Seim (2006) and Aguirregabiria and Mira (2007) to study endogenous market entry. For every given project, researchers decide whether or not to join.[3] By revealed preference, a researcher joins a project if and only if the expected payoff from doing so exceeds the payoff of not joining. The expected payoff depends on preferences regarding potential teammates, the researcher's exogenous characteristics (field of research and professional seniority), and exogenous project characteristics. I also include as a project-specific component the ex-ante project potential to control for unobserved factors related to the project that might influence the decision to participate (such as project complexity), and I allow for match specificity, i.e., a project-researcher-specific component. Solving the game allows me to obtain the equilibrium probabilities of joining a project.

3. Note that I model the decision to join a project instead of a team, as some projects at Virgo are single-authored.

In the second stage of the model (*outcome stage*), the participating researchers work on the project, which may be completed or not. The probability of each outcome depends on a knowledge production function of several inputs: the number and characteristics of team members, the observed project characteristics, and the ex-ante project potential.

Bringing the model to the data poses a major econometric challenge. The ex-ante project potential, unobserved by the econometrician, affects the project outcome directly (as a shock) and indirectly because of the selection of researchers into projects. To overcome endogeneity, I simultaneously estimate the two stages of the model. In other words, I control for selection by estimating the participation stage together with the outcome stage consistently with a control function approach. The approach is similar to that of Olley and Pakes (1996), and more recently Ciliberto, Murry, and Tamer (2021). These papers use it in the context of endogenous market decisions, such as entry or exit. The set of potential entrants represents the main exclusion restriction as it enters only the participation stage. The identification hinges on the fact that, for each project, the set of potential entrants is defined by the pool of available researchers, which is exogenous to the project.

Three key findings emerge. Regarding the participation stage, controlling for researchers' and project characteristics (including project complexity), an additional teammate decreases the probability of another researcher participating in a project by 14%. The finding is consistent with the idea that researchers internalize the coordination and communication costs of working in larger teams (Becker and Murphy, 1992).

Regarding the outcome stage, controlling only for project characteristics, one more researcher in a project is associated with a 1% lower probability of project completion, consistent with higher coordination and communication costs associated with larger teams. Moreover, when I additionally control for selection, the effect becomes ten times larger. The negative effect holds regardless of the field of specialization and the professional position, but the magnitude is heterogeneous across researchers.

To answer the question of whether self-formed teams perform better than other team structures, I compare the estimated probability of project completion of the observed self-formed teams to that of various counterfactual scenarios. The counterfactual results show that 1) randomly composed teams in composition and size have, on average, a 5% lower probability of project completion relative to the observed self-formed teams, and 2) teams that are composed to be more diverse in member characteristics (namely, with members who differ in field of specialization and professional seniority) have a 3% higher probability of completion relative to actually observed teams. The latter holds when team sizes are the same as the observed ones.

The counterfactual results highlight two relevant implications for the organization of knowledge production. When researchers can choose their projects, they internalize some of the costs and benefits of working together as they are aware of efficient team size. At the same time, they also tend to work with (too) similar peers, though working with more diverse peers could increase project efficiency (*homophily bias*). This shows that diversity may boost performance and act as a correction tool for the homophily bias, but it is preferable from a policy perspective if one takes into account decreasing returns to team size.

The paper contributes to several strands of the literature. Starting from the seminal work of Holmstrom (1982), many papers have studied the performance of teams. Hamilton, Nickerson, and Owan (2003) have been the first to evaluate empirically the endogenous formation of teams within a firm. Papers on peer effects have analyzed group interactions and how they affect productivity (Mas and Moretti, 2009; Bandiera, Barankay, and Rasul, 2010). Some of the recent literature has focused on identifying the effect of single team members on team output (Agha et al., 2018; Devereux, 2018; Ahmadpoor and Jones, 2019; Bonhomme, 2021). This paper is the first to provide a tool to analyze empirically the determinants of self-formed teams and test the performance of alternative team structures.

The second contribution is to bring methods developed for firms' endogenous market entry to an organizational setting. Since Bresnahan and Reiss (1990), the empirical literature in industrial organization has developed tools to analyze the determinants of market structures. The main goal is to better inform policymakers on the effect of endogenous changes in market structure on consumers and overall welfare.[4] In organizational economics, because of the increasing relevance of self-formed

4. See Aguirregabiria (2021) and Berry and Compiani (2021) for recent surveys of the literature on market entry.

teams, there is a growing interest in understanding the determinants of endogenous team formation, but no tool has been provided to address this issue yet. To the best of my knowledge, this is the first paper to adapt the methodology of Seim (2006), Aguirregabiria and Mira (2007), and Bajari et al. (2010) to an organizational setting by modeling team formation and analyzing the resulting implications for team performance.

The third contribution is to the literature on innovation economics. Hard science happens in teams (Jones, 2021). Part of the literature studies the performance of collaborations in the development of innovation (Waldinger, 2012; Akcigit et al., 2018; Ganglmair, Simcoe, and Tarantino, 2018; Anderson and Richards-Shubik, 2021). I contribute to this literature by leveraging the wealth of the Virgo data to estimate a model of team formation and performance and hence analyzing the mechanisms behind knowledge creation in science.

The last key contribution is to study which team structures are more desirable for complex organizations. Policymakers have encouraged diversity and interdisciplinarity in scientific institutions.[5] My counterfactual results speak to the literature on team diversity (Morgan and Várdy, 2009; Becker, 2010; Calder-Wang, Gompers, and Huang, 2021; Békés and Ottaviano, 2022) and contribute to the policy-relevant discussion on the value of diversity for science and knowledge production.

The paper proceeds as follows: Section 1 provides a description of the institutional details and the Virgo data. Section 2 presents the model and Section 3 the empirical implementation. Results are discussed in Section 4. Counterfactuals are presented in Section 5 and the conclusion in Section 6.

1 Data and Descriptive Evidence

I use unique data from a scientific institution named Virgo. This institution is particularly interesting for three reasons. First, researchers working at Virgo are free to choose their teams and projects. Hence, Virgo represents an ideal setting to study the determinants of self-formed teams. Second, these researchers report detailed information about their projects, including the outcome. Third, there is

5. ERC Synergy Grant: https://erc.europa.eu/apply-grant/synergy-grant.

considerable variation in team characteristics and outcomes across projects, which is likely associated with selection into projects.

The dataset comprises several sources. I web-scrape information regarding the organization of Virgo, the characteristics of the projects, their outcomes, and participants from the Logbook of Virgo from 2012 to 2016. I supplement the dataset by hand-collecting data on researchers' characteristics from several online public sources, mainly personal websites, available *curricula*, and LinkedIn profiles. I merge these sources into one comprehensive dataset that I pass on to estimate my structural model. The dataset is discussed in turn.

1.1 Institutional Details

Virgo studies the detection of gravitational waves and has two twin institutions in the U.S., LIGO Livingston and LIGO Hanford. Gravitational waves, predicted by Albert Einstein's general relativity, are generated by events happening in the Universe, such as the accelerated masses of orbital binary systems. The first event was recorded by LIGO in 2015 (the merger of a pair of black holes of 36 and 29 solar masses). Subsequently, similar events were recorded by Virgo. The institution has proven successful and the founders have been awarded the 2017 Nobel Prize in Physics.[6]

Virgo was founded by the French National Center for Scientific Research (*Centre National de la Recherche Scientifique*, CNRS) and the Italian National Institute for Nuclear Physics (*Istituto Nazionale di Fisica Nucleare*, INFN)[7] in 1987 and put in operation in 2003. Virgo is located in Italy, on the site of the European Gravitational Observatory (EGO), and run by an international collaboration of about 200 researchers.

Researchers are employed directly by Virgo and are paid a fixed wage by regulated contracts, in line with the national collective agreements and depending on the

6. \Pioneers Rainer Weiss and Kip S. Thorne, together with Barry C. Barish, the scientist, and leader who brought the project to completion, ensured that four decades of e ort led to gravitational waves nally being observed." Source: https://www.nobelprize.org/prizes/physics/2017/press-release/.

7. National Research Centers in France and Italy.

level of seniority. For juniors, a typical contract is the so-called *assegno di ricerca*.[8] Senior researchers are usually affiliated with a university or a research institution. Crucially, working contracts do not specify any formal obligations in terms of project participation and there exist no bonuses based on performance.[9]

As in many scientific settings, intrinsic motivation plays an important role for the researchers working at Virgo. Moreover, adding this experience to the CV is a positive signal for the job market. Last, there are internal long-term monetary and career incentives for these researchers: future career development might depend, for instance, on making a good impression on a senior researcher while working with her. At Virgo, the assignment of researchers to scientific projects and teams is self-organized: each researcher voluntarily decides what to work on and with whom. These motivating forces impact the decision to join a project.[10]

The scientific purpose of Virgo is to develop, build and put in function a laser interferometer. Interferometers are devices that extract information from interference, hence they work by merging two or more sources of light to create an interference pattern, which can be measured and analyzed (Figure 4 in Appendix A). The interference patterns generated by the interferometers contain information about the phenomenon that produces gravitational waves (e.g. the merger of two black holes).

Building up the interferometer requires an effortful amount of resources and time. The process from the development to the operation of the interferometer is divided into intermediate steps, defined as macro-projects. Macro-projects relate to different phases of the development of the interferometer, from the Infrastructure System, which concerns building the infrastructure of the interferometer, to the Injection System, which takes care of the optics of the high-power laser. Therefore, different sets of skills and knowledge are required depending on the actual task to perform. Macro-projects are then split into projects, which relate uniquely to a phase of the

8. https://it.wikipedia.org/wiki/Assegnista_di_ricerca. There are also other short-term contracts that can be renewed for a limited amount of time, the so-called RTD (*ricercatori a tempo determinato*). For an overview of the current legislation on RTD contracts, see *Legge n. 240/10 del 30 dicembre 2010, art. 24*.

9. Figure 5 in Appendix A.2 shows the general organizational chart of Virgo in 2016.

10. By talking to people at Virgo, it emerged that the main reason for this decentralized allocation is that typically researchers know what to work on and, hence, they do not need to have someone who tells them what to do.

macro-project. A project can consist in analyzing data or building a system of mirrors in a lab. I provide some examples in Appendix A.3.

The dataset spans more than 4 years from June 2012 to September 2016. June 2012 is the starting point of a new phase of the experiment (Advanced Virgo).[11] Projects were set up in advance and written in the *Technical Design Report* in April 2012, hence before the sample period. The Report contains detailed descriptions of Virgo's objectives and how projects should develop. It has been edited and signed by the researchers working in Virgo at that time and covers the whole interferometer construction. The Report is meant to be the project reference document for all the design aspects of the experiment. As explicitly written, each configuration change with respect to what is stated in the document requires a formal Change Request Procedure, and this almost never happened. Hence, the set of projects can be considered fixed and projects pre-determined.

Researchers at Virgo communicate using an online platform: the Logbook.[12] It comprises web pages held by project teams (McAlpine et al., 2006). Other scientific institutions work with electronic Logbooks. For instance, LIGO uses a similar Logbook platform, and CERN in Geneve uses the so-called CMS electronic Logbook.[13]

The Logbook proves useful for the advances of scientific research, as it allows researchers to record information on working projects and experiences such as results of measurements, tests, and data taking. Researchers are obliged to report their work in the Logbook. This obligation facilitates monitoring, as reports are observable and their content is verifiable.[14]

Each web page of the Logbook consists of logs. A log presents a description of a project; it is identified by the title of the macro-project and the project it refers to, the name of the author(s), the time and date, the (chronological) number, the main text and possibly images, comments or other files attached. A screenshot example of a Logbook web page is given in Figure 7 of Appendix A.

11. The phase ended in January 2017, when the interferometer was turned on.

12. https://tds.ego-gw.it/itf/osl_virgo/index.php.

13. https://alog.ligo-wa.caltech.edu/aLOG/, https://alog.ligo-la.caltech.edu/aLOG/, http://cds.cern.ch/record/1272667.

14. Because of these features, free riding is of limited concern in the setting.

1.2 Descriptive Statistics Logbook

Table 1 shows descriptive statistics about the Logbook data for the period of observation (June 2012 – September 2016). The initial dataset contains $3,778$ logs. For the purpose of the analysis, logs that do not belong to the Advanced phase of the experiment (147), logs that only contain external participants (277), and logs that are not related to actual projects (299) are dropped from the sample, together with logs with more than 8 participants (10 projects). The final dataset contains $3,045$ logs that identify unique projects belonging to 16 different macro-projects. Table 8 in Appendix C reports the frequency of the macro-projects in terms of the number of projects. The data cleaning process is discussed in Appendix B.

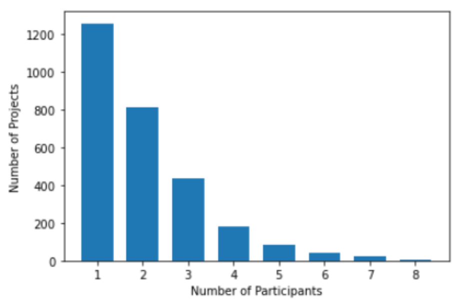
When a project ends, researchers write a final report in a log. There do not exist multiple logs related to one project. However, there can be follow-ups of a project that are easily identified as the titles contain the words "Comment to" and the title of the project they refer to. One can think that projects with follow-ups (defined as "parent" projects) are more complex or more crucial for the development of the experiment; hence I control for these characteristics in the empirical analysis. In particular, I define a dummy for whether a project is a "comment" and a dummy for whether a project is a "parent." In the sample, 18% are parent projects, and 23% are comment projects.

In some projects (8% in total) there are external companies or groups. External companies, for instance, supply Virgo with instruments and tools for lab experiments and help researchers set up those instruments. I account for them in the empirical analysis.

Table 1: Descriptive Statistics

Sample Period	June 12 - Sept. 16	
No. of Projects (obs)	3,045	
No. of Macro-projects	16	
	Mean	**St. Dev.**
Parent Projects	18%	0.38
Comment Projects	23%	0.42
Projects with External Groups	8%	0.27
No. of Projects/Month	62	55.38
Team Projects	66%	0.48
Team Size	2.27	1.41
Max Team Size	8	
Completed Projects	51%	

Figure 1: Project Participants



It is not possible to determine the duration of a project because the logs do not contain information on the initial project date. However, Table 1 shows that on average in a month 62 projects are carried out at Virgo.[15] This value indicates that projects are short-lived. For this reason, it is also unlikely that researchers coordinate beforehand and outside the platform about who is joining projects.[16]

In terms of project participation, around 66% are team projects, and the rest are solo projects. The average team size has 2.27 participants with a maximum of 8. Figure 1 shows the distribution of project participants by the number of projects. Projects with one participant are the most frequent, followed by projects with two and three participants. The frequency decreases substantially for projects with four and more participants.[17]

To evaluate the outcome of a project, I require a measurable output. One possibility would be to use publications that resulted from the projects. Unfortunately, this is not a viable option for two reasons. First, not all projects end with a publication. Second, at Virgo, the general rule is that publications that follow from a project must contain the names of all Virgo researchers in alphabetical order, regardless of the contribution.[18] Therefore publications do not represent adequately the outcome of a single project. Likewise, as projects do not generally yield standards or patents, I cannot use them as an outcome measure.

The logs represent a critical source for this scope. I examine the text to attribute an outcome to each project. In particular, I classify each text into one of two different categories, completed and not completed.[19]

The last row of Table 1 shows that 51% of the projects are completed. Table 9 in Appendix D presents additional descriptive statistics for project completion in terms of the number of participants. Notice that while completion cannot be mapped

15. Figure 8 in Appendix C shows how the average number of logs changes over the sample period.

16. I perform some robustness analysis to address this potential concern. Results are discussed in Section G.

17. Because only one nal log for each project is reported, I only observe the nal list of project participants. One might wonder if in some cases researchers join projects sequentially. The data limitation does not allow me to test this hypothesis. However, given that projects are frequent and short-lived, it is plausible to assume that this is unlikely to happen.

18. By checking the research outputs of some researchers (for instance, on Researchgate), it appears that many publications related to Virgo have above 1,000 authors.

19. Appendix D contains more details about the outcome classi cation.

directly into a standard measure of success, it is still a useful outcome measure in this framework, as it allows Virgo to progress, and in science in general.[20]

Finally, I merge the described data on projects with information about researchers at Virgo, which I discuss in turn. Details on the merging are in Appendix B.

1.3 Descriptive Statistics Researchers at Virgo

I hand-collect data on researchers' demographic characteristics (e.g., nationality, gender, education, professional seniority, field of research) from several online sources, mainly personal websites, available *curricula*, and LinkedIn profiles. Around 160 different researchers write in the Logbook. All the researchers are highly qualified. At the same time, the pool is heterogeneous: researchers can have different levels of seniority, work in various fields, and have different nationalities. In order to coherently classify them in terms of seniority and education, I use the criteria provided on the websites of the main European National Research Centers (Appendix B).[21]

Table 2 provides descriptive statistics of researchers' demographics. Not surprisingly, around 60% of researchers at Virgo are specialized in Physics, and the rest is mainly specialized in the area of Engineering and Electronics (31%), with a smaller percentage (9%) specialized in other fields. Around 14% of the researchers are doctoral students or post-docs, 63% are at a higher level of seniority (Researchers or Assistant Professors), and 23% are Seniors Researchers or Full Professors.[22] In terms of gender and nationality, the majority are male (82%) and are either Italians or French (93%), with a low fraction (7%) belonging to a different nationality.

The last row of Table 2 shows the monthly average of projects per researcher. Conditional on being active at Virgo (namely, being on at least one project before), each researcher works on 3.7 projects per month. I use this piece of information i) to construct a proxy of experience by computing the cumulative number of projects a researcher worked on at each point in time, and ii) to measure the availability of the

20. For instance, in the development of a drug, the nal goal is reached after several steps of trial and error.

21. For some researchers (around 3% of the total number of project participants) I did not nd any information. Most likely they are technicians or seniors that do not have an online identity.

22. Seniors include technicians that do not necessarily hold an academic degree.

	%
Field of Specialization	
Physics	60%
Engineering and Electronics	31%
Informatics and Others	9%
Professional Position	
Ph.D., Post-doc	14%
Researcher, Ass. Prof.	63%
Senior Researcher, Full Prof.	23%
Male	82%
Nationality	
Italian	67%
French	26%
Other	7%
Avg # projects/month/researcher	3.7

Table 2: Descriptive Statistics Researchers

researchers when I define the potential entrants for each project (see Section 3.4 for more details).

As the field of specialization and the professional position exhibit a higher degree of variation than other characteristics, I exploit these characteristics in the structural model. In particular, I assign each researcher to mutually exclusive types, defined as a combination of field of specialization and professional position. This also allows me to reduce the computational burden of the empirical model. I define five researcher types: Physics Senior, Physics Junior, Engineer Senior, Engineer Junior, and Other.[23]

Table 3 shows descriptive statistics in terms of researcher types. Physics Juniors are the majority, accounting for 71 researchers. This is reflected also in the number of projects: Physics Juniors are in 1,851 projects, followed by Physics Seniors (1,467) and the others. Interestingly, by looking at the numbers of completed and not completed projects and at their ratios (last three columns), one can see that researchers who specialize in Physics are associated with a lower ratio of completed to not completed projects than other researchers types: Physics Juniors, for instance, are in 885 completed projects and 996 not completed projects, which gives a ratio of 0.88 completed to not completed projects; Engineer Juniors instead are in 468 completed projects and 363 not completed projects, which gives a ratio of 1.29 completed to not completed projects. Based on this evidence, one would be tempted to conclude

23. The latter includes researchers belonging to any other eld or professional position as well as non-classi ed researchers.

that specific characteristics of the researchers cause better performance. However, this conclusion does not consider how teams are formed. In the empirical analysis, I investigate further what are the determinants of project completion, after controlling for selection.

	Researchers			**Projects**	
		Total	Completed	Not Completed	Completed/Not Completed
Physics Senior	26	1,467	696	771	0.90
Physics Junior	71	1,851	885	996	0.89
Engineer Senior	2	18	10	8	1.25
Engineer Junior	28	831	468	363	1.29
Other	32	876	494	382	1.29

Note: The Table shows in the first column the number of researchers by type, in the second to fourth column the total number of projects, the number of completed and not completed projects and the ratio completed/not completed projects by researcher type.

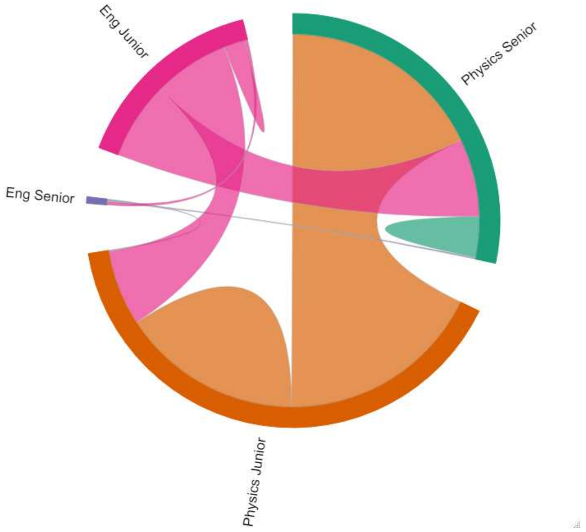
Table 3: Descriptive Statistics Researcher Types

As discussed above, researchers might join projects because of who else is joining. Figure 2 gives a comprehensive illustration of the bilateral project connections among researcher types. The orange flow that links Physics Juniors and Physics Seniors represents the projects in which the two researcher types collaborate. The orange flow that turns back into the orange part represents the projects in which Physics Juniors collaborate with other researchers of the same type. One can easily see that Physics Juniors are working more frequently with researchers specialized in Physics (both Juniors and Seniors) than with Engineers. Moreover, Engineer Juniors collaborate more frequently with others than with researchers of the same type, as suggested by the pink flows. The evidence suggests that the matching of researchers is non-random. In the structural model, I explore these paths to identify the main determinants of project participation.

2 Model

In this Section, I present the two-stage structural model that quantifies the determinants of team performance controlling for endogenous participation. In the first stage (*participation stage*), each researcher type observes the set of exogenous project characteristics and the set of potential entrants, a project-specific shock, and her own

Figure 2: Chord Diagram Bilateral Project Connections Researcher Types



The graph shows the bilateral connections among researcher types. The length of the arches corresponds to the total number of projects with at least one researcher of each type. The bilateral connections are represented by the ows.

idiosyncratic shock. She decides whether to join a working project by comparing post-joining single-period payoffs. In the second stage (*outcome stage)*, the participating researcher types work on the project, which ends with a certain outcome. In this Section, I discuss in detail the two stages.

2.1 Participation Stage

I model the decision to join a project as an entry game with incomplete information (Seim, 2006; Aguirregabiria and Mira, 2002). The model is static and researcher types make their decisions simultaneously.[24] The payoff from joining a project is assumed to be positive, while the payoff from not joining is normalized to zero.[25] For every given project, a researcher type decides whether to join a project by comparing single-period

24. As already discussed in Section 1.1, the project logs contain the names of every team member. This rules out the possibility of modeling the decision as sequential.

25. This assumption is standard in the literature of entry games.

payoffs.

Consider a set of projects $\mathcal{J} = \{1, ..., J\}$ indexed by j over a certain time span, and a set of researcher types i, with $i = 1, ..., I$. For a given project at time t, a researcher type decides whether or not to join. The variable $Participation_{ijt}$ takes on a value of 0 or 1 depending on whether researcher type i is in the project or not. It can be rationalized by the following payoff function:

$$Participation_{ijt} = f\left(\bar{N}_j, X_{it}, D_j, Z_t, q_j\right) + \epsilon_{ij} \tag{1}$$

The function depends first, on the number of potential teammates, $\bar{N}_j$. In the decision to join a project, a researcher type takes into account who else might be joining for, say, personal reasons or for potential externalities in production. In the empirical specification, the coefficient for $\bar{N}_j$ captures how the participation probability changes if an additional teammate joins the project: if the coefficient is positive (negative), it means that adding an additional potential teammate to the project increases (decreases) the probability of the project participation.[26]

Additionally, the function depends on the exogenous characteristics of the researcher type, X_{it}, namely a dummy for the researcher type and the average cumulative number of projects a researcher type is prior to time t. The latter can be viewed as a stock variable and, hence, a measure of experience. The function also depends on project controls D_j and monthly controls Z_t, which account for time trends related to participation. In particular, the vector D_j includes macro-project categories, a dummy for whether the project has external participants, and dummies for whether it is a comment or a parent project.

Finally, the function contains two stochastic components: a project-specific shock q_j, the ex-ante potential (similar in spirit to Hill and Stein, 2021), which captures all unobserved (by the econometrician) factors that influence the decision to join a project including complexity, and a researcher type-project-specific shock ϵ_{ij}, that gives information about the match of the researcher type for the project.

A researcher type might decide not to participate in a project because of no time availability. In this framework, I do not explicitly model the time constraint.

26. I allow the coe cient to be either positive or negative to capture complementarities or substitutabilities.

However, I take it into account first by controlling for the average number of projects a researcher type is in before time t, and second, when defining the set of project potential entrants empirically. I discuss this in detail in Section 3.

I assume that each researcher type observes her own project-specific shock, but only knows the distribution of the others' shocks; therefore, the described entry game is a game of incomplete information. Because of this information structure, i can only form an expectation of the others' optimal choices. Based on the expected teammate distribution across projects, each researcher type chooses whether to join a project by maximizing her expected payoffs.

Assuming that the error terms ϵ_{ij} are iid draws from a continuous distribution, the Bayes-Nash equilibrium probability of joining project j at time t for i, p^*_{ij}, is then:

$$p^*_{ijt} = \Phi\left(X_{it}, D_j, Z_t, p^*_{gjt}, q_j\right) \tag{2}$$

for all i and g, where $\Phi(\bullet)$ is a continuous CDF. Researcher type i's vector of equilibrium conjectures over all projects is given by the set of J equations that define the equilibrium probabilities. The Bayesian Nash equilibrium(a) of the game consists of finding the optimal response(s) that maximizes the researcher type's expected payoff, given her conjecture about others' strategies.[27] Lastly, I define as d_{ijt} the choice of researcher type i for project j at time t. Hence, $d_{ijt} = 1$ if and only if $Participation_{ijt} \geqslant 0$.

2.2 Outcome Stage

Researcher types work on projects which end with a certain outcome. The variable $Outcome_{jt}$ takes on a value of 0 or 1, depending on the project classification (as discussed in Section 1). It underlies a continuous variable $Outcome^*$, which is a latent variable for the degree of project completion (details in Section 3) and can be expressed as a knowledge production function of different "inputs" in the following way:

27. As it is standard for these games, the equilibrium existence is given by Brouwer's Fixed Point Theorem. The types' own conjectures enter the probability simplex and are continuous in others' expected behavior.

$$Outcome_{jt} = f\left(\boldsymbol{N}_j, C_j, Z_t\right) + q_j. \tag{3}$$

The production function includes a vector of researcher types $\boldsymbol{N}_j$, who endogenously participate in the project (similar in spirit to Akcigit et al. (2018)). It also includes project controls C_j, to allow for exogenous project characteristics, and monthly controls Z_t, to allow for time trends in production. The vector C_j includes macro-project categories, a dummy for whether the project has external participants, and dummies for whether the project is a parent or comment project. Moreover, a project might end with a better outcome because of some unobservable factors, such as complexity. These factors fall in the error term q_j.[28]

In the empirical specification, the vector of coefficients for $\boldsymbol{N}_j$ indicates how additional researcher types affect project outcome, and hence can be viewed as measures of researcher types performance: if the coefficient for a particular researcher type is positive (negative), it means that adding a researcher type to the project increases (decreases) the probability of the project completion.

3 Empirical Implementation

Using the Virgo data, I estimate the structural model that allows quantifying the main determinants of project participation and outcome. The model is then used to answer the question of whether self-formed teams perform better than other team structures.

Notice that if one was to estimate the participation stage alone, then that would be analogous to estimating an entry game. If one was to estimate the outcome stage alone, that would mean estimating a production function. Here, a complication to be tackled is that the ex-ante project potential q_j, which is unobserved by the econometrician, affects both the decision to join a project and the final outcome. The latter occurs through two different channels. First, q_j enters the outcome directly

28. Complementarities in production could, in principle, be embedded in the outcome stage as interaction terms. In the two-stage estimation procedure, allowing for the identi cation of complementarities requires additional assumptions on the set of potential entrants and, hence, complicates further the analysis.

as a residual; second, it affects it indirectly through the project participants, who endogenously join the project.

To solve the unobservability issue, one could compute the residuals from the outcome equation and use them to estimate the parameters of the structural model. However, because of selection, the measure of the residual from the outcome equation would likely be biased.

To overcome the selection issue, I estimate the two empirical components jointly. This way I combine both participation and performance and hence I am able to see what would happen under alternative team structures, after controlling for selection. This is similar in spirit to Ciliberto, Murry, and Tamer (2021), that estimate simultaneously entry and pricing decisions of firms, hence accounting for selection in the pricing stage.

I make additional assumptions to parametrize the participation and the outcome functions for the empirical implementation. In the main specification, the functions are linear in the parameters. More details can be found in Appendix E.1.

In this Section, I present the estimation procedure for the project participation stage and the outcome stage, I describe the procedure for the joint estimation and finally, I discuss the identification of the model parameters.

3.1 Participation Stage

Entry games with strategic interactions are likely to lead to multiple equilibria, especially in the presence of strategic complementarities. Solutions to the multiplicity problem have been proposed by, among others, Bresnahan and Reiss (1991) and Berry (1992). Papers on moment inequalities (Ciliberto and Tamer, 2009) allow for general forms of heterogeneity across players providing a methodology for set identification without making equilibrium selection assumptions. However, bounds for the estimated coefficients are likely to give very little information on the kind of strategic interactions among players if their ranges are too broad. This is not well suited in this setting given that one of the goals is to measure the degree of complementarity and substitutability among researcher types. Alternatively, Schaumans and Verboven (2008), for example, imposes assumptions on the sign of the strategic parameter, but in this framework, any assumption would appear to be *ad hoc*.

Part of the literature deals with the multiplicity issue by using a two-step estimation procedure (Aguirregabiria and Mira, 2002, 2007; Bajari et al., 2010), without imposing any further assumptions on the strategic parameter. The method eliminates the need to solve the fixed-point problem when evaluating the corresponding (pseudo) likelihood function that is implied by the structural choice probabilities.

I adapt the two-step method to my static framework and allow for strategic complementarity and substitutability. In the first step, I estimate the probabilities of participation conditional on project observables.[29] In the second step, I find the structural parameters that are most consistent with the observed data and the estimated equilibrium probabilities. A key assumption for the consistency of this approach is that, in the data, two projects feature the same equilibrium conditional on observables.[30]

Let d_{ijt} be the choice of researcher type i for project j at time t and $\Psi_i = \Phi(X_{it}, D_j, Z_t, q_j, \boldsymbol{p}^*)$, where Ψ_i follows a logistic distribution. In line with the literature (Aguirregabiria and Mira, 2007), the Pseudo-Likelihood Function is the following:

$$Q_J(\theta, \boldsymbol{p}) = \frac{1}{T}\frac{1}{J}\frac{1}{I}\sum_{t=1}^{T}\sum_{j=1}^{J}\sum_{i=1}^{I}\log \Psi_i(d_{jt}|X, D, Z, q; \boldsymbol{p}, \theta_1) \tag{4}$$

3.2 Outcome Stage

I estimate the outcome stage using a standard discrete choice model. I assume that the error terms are iid logistically distributed across observations and I set the location and scale parameters equal to 0 and 1, respectively.

As previously discussed, I need to get an estimate of the unobserved ex-ante potential, which is given by the residual. In the case of latent models (probit, logit, ordered probit, etc.) it is not possible to calculate the residuals directly since the latent dependent variable *Outcome** is not observed. Following Gourieroux et al. (1987), I compute an estimate of the conditional distribution of *Outcome** conditioned on the

29. The two-step procedure is embedded in a joint maximum likelihood estimation, therefore I estimate the rst step parametrically to ease the computational procedure.
30. Appendix E discusses this in more details.

observables. From this, I obtain an estimate of the conditional distribution of the error term q_j, from which I construct the generalized residuals $\tilde{q_{jt}} = E[q_{jt}|N, C, Z, \hat{\theta}_2]$. The vector of estimated parameters $\hat{\theta}_2 = (\hat{\tau}, \hat{\boldsymbol{\kappa}}, \hat{\lambda}, \hat{\xi})$ is obtained by maximum likelihood.[31] The residual captures all the unobserved factors that enter the ex-ante project potential. Researchers are likely to sort into projects because of this component. Sorting creates a problem of endogeneity that biases the results of the estimation.

3.3 Joint Estimation

To overcome the endogeneity issue, I estimate the participation probability and the probability of project completion jointly. In a similar framework, Seim (2006) estimates a model of entry with endogenous product-type choices by computing the joint equilibrium prediction for the location probabilities and the equilibrium number of entrants in a market. Likewise, I compute the joint prediction for the probability of project completion and the equilibrium number of project participants. In Seim (2006), however, the location decision does not depend on the market-level unobservable, which influences only the probability of entry.

In this setting, the project-level unobservable q_j affects both the decision to join a project and the project outcome, directly and indirectly through N_j. Therefore, to account for this issue, I express the generalized residual $\tilde{q}_j$ as a function of the outcome variables and I substitute it into the payoff function. By doing so, I estimate the equilibrium parameters of the model of project participation taking into account the project ex-ante potential (unobserved by the econometrician) and I solve for the endogeneity in the outcome equation. The inversion procedure resembles that of Olley and Pakes (1996) in terms of the control-function approach.

For given d_{ijt}, the joint pseudo-likelihood is:

$$\begin{aligned} f(d, outcome) = & \prod_{t=1}^{T}\prod_{j=1}^{J}\prod_{i=1}^{I} Pr(d_{ijt}|X, D, Z, \tilde{q}; P, \theta_1)\times \\ & \times \prod_{t=1}^{T}\prod_{j=1}^{J} Pr(Outcome_{jt}|N, C, Z; \theta_2). \end{aligned} \tag{5}$$

31. See Appendix E.2 for details.

Equation (5) consists of two parts. The first computes the likelihood of observing the participation choices conditional on the project-level unobservable $\tilde{q}$. Recall that $\tilde{q}$ is the random factor that affects also the probability of observing a particular outcome realization. Therefore, to derive the unconditional likelihood, the first component of the joint pseudo-likelihood is multiplied by the probability of observing a certain outcome such that the predicted and actual probability of project completion are equal.

Because of simultaneity, I derive the unconditional likelihood by expressing $\tilde{q}$ as a function of the outcome parameters and regressors and substitute it into the payoff function of the model of project participation. I follow Olley and Pakes (1996) in inverting the function. I assume that the error terms of the model of project participation and the outcome equation follow a logistic distribution.[32] The joint pseudo-loglikelihood is:

$$\begin{aligned} LL(\theta) = & \frac{1}{T}\frac{1}{J}\frac{1}{I}\sum_{t=1}^{T}\sum_{j=1}^{J}\sum_{i=1}^{I}\log\Psi_i(d_{jt}|X, D, Z, \tilde{q}; \boldsymbol{p}, \theta_1)+ \\ & + \frac{1}{T}\frac{1}{J}\sum_{t=1}^{T}\sum_{j=1}^{J}\log\Psi(Outcome_{jt}|N, C, Z; \theta_2). \end{aligned} \tag{6}$$

In line with the estimation procedure for the model of project participation described above, I perform the joint estimation in two steps. The details of the empirical implementation are discussed in Appendix E.

1. I maximize the joint log-likelihood without the vector $\boldsymbol{p}$ and obtain the reduced-form estimates of the equilibrium probabilities of participation, together with the estimates of $\theta_1^{Step1}, \theta_2^{Step1}$. In this step, I account for the correlation between the project outcome and participation through $\tilde{q}_j$, but not for the endogenous participation as I ignore the strategic interactions.

2. With the probabilities predicted in the first step, I construct the joint pseudo-log-likelihood function expressed in (6) and obtain the final estimates for $\hat{\theta}_1, \hat{\theta}_2$.

32. I restrict the variance-covariance matrix of the joint distribution of the error terms to be the identity matrix.

3.4 Identification

I now discuss the sources of exogenous variation that allow the identification of the parameters of the two stages and the exclusion restrictions for the joint estimation. Recall that the participation stage is at the researcher-project level whereas the outcome stage is at the project level. The variation in the macro-projects and project characteristics identifies the project-level parameters both in the participation and in the outcome stage. For the joint estimation, one needs to include something that affects project participation but not the outcome.

Two important sources of variation only enter the participation stage. First, type-specific characteristics (field or research and professional seniority) that are exogenously fixed and identify the type-dummy coefficients. The second and main exclusion restriction comes from the set of potential entrants. Notice that the payoff function in the participation stage contains the number of *expected* potential entrants in a project, while the outcome equation contains the number of *actual* researcher types in a project. The expectation varies across projects and is formed before the researcher type makes the decision to participate. As an important exclusion restriction, I use the set of characteristics of the potential entrants. The choice of this variable is motivated by the fact that the number of projects a researcher type worked on *before* a project start is fixed in the past and, hence, exogenous to the project. In turn, the set of potential entrants only affects the decision to participate in a project but not directly the project outcome.

Using the universe of researchers as potential entrants is not a plausible assumption. At each point in time, some researchers might not be active at Virgo or might be busy working on other projects. Hence, I define the set based on a measure of non-business of researchers. First, I fix at five projects the threshold for this set, meaning that, at each point in time, researchers already working on five or more projects will not enter the set of non-busy researchers. This threshold is justified by the descriptive evidence that conditional on being active at Virgo, researchers work in a month on average on 3.6 projects, as shown in Section 1.3. With the remaining pool, I make random draws from a uniform distribution to construct the set of potential entrants per project.[33]

Finally, simulation exercises show that a crucial source of exogenous variation that reinforces the identification power of the exclusion restriction is the variation in the identity and number of potential entrants across projects, and hence the variation in the predicted participation probabilities from Step 1 across researcher types. I observe the same researchers working both on solo and team projects, where teams are heterogeneous and can have different sizes. The identification strategy exploits also this variation in team memberships.

4 Results

I present the results in two steps. First, I show reduced-form results from estimating the outcome stage alone without controlling for participation. This exercise represents a meaningful benchmark once I move to the discussion of the results of the structural model. Moreover, it provides an intuition of what are the crucial drivers of project completion. Then, I show the results from the estimation of the structural model that accounts for project participation and ex-ante project potential.

4.1 Reduced-Form Analysis

First, I present the results from reduced-form regressions of the outcome stage (equation (3)). The dependent variable can take values 0 or 1 ("not completed" or "completed", depending on the classification explained in Section 1.2).[34] Results are reported in Table 4.

Column (1) includes only the total number of project participants as a covariate. The coefficient is negative and significant: an additional participant is associated with a decrease in the probability of project completion by 1.5%. Column (2) includes also a dummy for whether there is an external firm or group in the project and

33. With this procedure, 7% of the set of actual entrants are mechanically dropped from the sample. In a robustness analysis, I show that my results are similar also when I use a di erent threshold level.

34. I use Logit speci cations. Results from Probit and Cobb-Douglas speci cations are similar in spirit.

	Outcome Stage					
	(1)	(2)	(3)	(4)	(5)	(6)
No. of Project Participants	-0.015 (0.006)	-0.018 (0.007)	-0.018 (0.007)	0.011 (0.023)		
No. of Project Participants^2				-0.004 (0.003)		
No. of Physics Seniors					-0.042 (0.13)	-0.041 (0.07)
No. of Physics Juniors					-0.034 (0.008)	-0.033 (0.009)
No. of Engineer Seniors					0.005 (0.121)	0.000 (0.12)
No. of Engineer Juniors					0.019 (0.018)	0.018 (0.018)
No. of Others					0.025 (0.014)	0.021 (0.014)
Dummy for External Firm or Group		-0.042 (0.033)	-0.068 (0.034)	-0.065 (0.034)	-0.081 (0.018)	-0.093 (0.035)
Dummy for Parent Project		-0.172 (0.023)	-0.167 (0.023)	-0.166 (0.024)	-0.161 (0.024)	-0.160 (0.024)
Dummy for Comment		-0.060 (0.023)	-0.054 (0.023)	-0.054 (0.023)	-0.054 (0.023)	-0.053 (0.023)
Macro-Project Controls	No	No	Yes	Yes	No	Yes
Time Controls	No	No	Yes	Yes	No	Yes
LL at convergence	-2107.65	-2081.83	-2077.25	-2076.38	-2067.07	-2065.62

Notes: The Table reports the marginal effects from reduced-form regressions of the outcome stage. Column (1) includes only the total number of project participants. Column (2) adds various project controls. Column (3) adds macro-project and time controls. Column (4) adds the square of the total number of participants. Column (5) includes the number of researcher types and Column (6) adds various controls. The number of observations is 3,045. All regressions include the constant. Standard errors in parenthesis. The last row of the Table reports the value of the log-likelihood at convergence.

Table 4: Results Outcome Stage

dummies for whether the project is a parent or a comment project. The effect of an additional team member on the probability of project completion remains negative and significant, and slightly larger. The presence of an external firm or group does not seem to significantly affect the probability of project completion (though the coefficient is negative) while being a parent project or a comment is associated with a negative probability of project completion. In particular, parents or comments are, respectively, 17% and 6% less likely to be completed relative to standard projects. This is not surprising as these variables can be interpreted as proxies for complexity. Results are similar in column (3), where I add macro-project and time controls. To sum up, from columns (1)-(3) of Table 4 it clearly emerges that projects with more participants are associated with a lower probability of completion. Notice that one cannot give a causal interpretation to these results because of the likely selection into projects.

To explore the presence of non-monotonicities in the number of participants, in Column (4) I include all the covariates previously specified together with a quadratic term for the number of project participants. Interestingly, though not significant, the linear coefficient is now positive while the coefficient of the quadratic term is negative. Though there is not enough power to give statistical meaning to this result, it points to the existence of a non-monotonic relationship. Two main rationales can explain non-monotonicity. First, decreasing returns to scale in team production. In particular, the marginal contribution of an additional researcher of a given type can be decreasing as the improvement on the pre-existing stock of skills already present in the project can shrink. Alternatively, free-riding in teams can imply that, as the number of researchers increases, some researchers can exploit the work of the other teammates. Free-riding does not play an important role in this setting, as discussed in Section 1, hence the first mechanism seems to be in place.

Researchers with different characteristics might affect differently the probability of project completion. To further investigate the heterogeneous effects related to researcher types, in columns (5) and (6) of Table 4 I present results from specifications where I include the number of project participants for each researcher type. Column (5) shows that an additional participant specialized in Physics is associated with a lower probability of project completion, whereas the contrary holds for participants specialized in Engineering and other fields. For example, an additional Physics researcher in the team is associated with a decrease in the probability of project completion of around 3%. These results remain also when adding macro-project and time controls (column (6)). Appendix F shows additional reduced-form results from estimating the participation stage alone.

To sum up, the results from Tables 4 show that larger teams are associated with a lower probability of project completion. Moreover, it is important to notice that the effect on the probability of project completion is not randomly distributed across researcher types. It is not possible to give an economic interpretation to the results as researchers might select into projects with better ex-ante potential or because they want to work with certain teammates.

4.2 Structural Model

I now present the results from the joint estimation of the full structural model, as discussed in Section 3. First, I show the results from a simpler specification where the selection effect is homogeneous across different researcher types. Then, I turn to the results with heterogeneous selection effects. Finally, I discuss the results and the potential mechanisms.

Homogeneous Selection Effects Across Researcher Types
The first column of Table 5 reports the results from Step 1 of the procedure described in Section 3.3. I allow for correlation in the project ex-ante potential both in the outcome and in the participation by estimating the joint likelihood expressed by equation (6), but I ignore the effects of the potential teammates on the probability of participation (namely, there are no strategic effects). In the model of participation, the number of potential entrants is 10 for every project.[35]

The first set of variables refers to the outcome stage. Controlling for correlation in ex-ante potential has a substantial impact on the estimates of the outcome stage: an additional project participant decreases the probability of project completion by 10%, i.e., ten times more than the effect found in the reduced-form analysis (Table 4). This result shows that ignoring correlation in ex-ante potential leads to an overestimation of the effect of teams on the probability of project completion. Moreover, the effects for the presence of external firms or groups and that for being a comment almost double in size relative to those of Table 4.

The second set of variables refers to the participation stage (equation (1)). *Ceteris paribus*, being a Physics Senior increases the probability of participation by 4% relative to the left-out category (Others). Vice versa for the other researcher types, which exhibit a negative probability of participation: being an Engineer Senior, for instance, decreases the probability of participation by 31%. By comparing these re-

35. I perform robustness analyses with di erent sets of potential entrants, as discussed in Appendix G.

	Step 1 No Endogenous Participation	**Step 2** Two-Step Pseudo-Likelihood
Outcome Stage		
No. of Project Participants	-0.099 (0.001)	-0.108 (0.001)
Dummy for External Firm or Group	-0.082 (0.006)	-0.068 (0.006)
Dummy for Parent Project	-0.117 (0.006)	-0.115 (0.006)
Dummy for Comment	-0.121 (0.005)	-0.100 (0.005)
Participation Stage		
Dummy for Physics Senior	0.039 (0.002)	0.088 (0.002)
Dummy for Physics Junior	-0.034 (0.002)	0.017 (0.002)
Dummy for Engineer Senior	-0.312 (0.004)	-0.255 (0.005)
Dummy for Engineer Junior	-0.312 (0.004)	-0.037 (0.003)
Dummy for External Firm or Group	-0.037 (0.004)	-0.039 (0.005)
Dummy for Parent Project	-0.014 (0.002)	-0.039 (0.003)
Dummy for Comment	-0.099 (0.003)	-0.129 (0.004)
No. of Potential Teammates		-0.146 (0.002)

Notes: The Table reports the marginal effects at the mean from the structural model. Column (1) reports the results from Step 1 of the two-step procedure. Column (2) reports the results from Step 2. The number of observations is 3,045. The number of potential entrants is 10. All regressions include macro-project and time controls, and a constant. Bootstrapped standard errors in parenthesis.

Table 5: Results Structural Model

sults to those found in a reduced-form analysis of the participation stage (Table 10 in Appendix F), one can see that, both in terms of sign and magnitude, controlling for selection into the project ex-ante potential has an impact on the probability of participation.

The second column of Table 5 reports the results from Step 2 of the procedure described in Section 3.3, namely the full structural model where I control for the correlation in the project ex-ante potential and endogenous participation. The effects in the outcome stage remain stable and significant. However, the results of the participation stage are affected. The effect of being a Physics Junior turns out positive compared to that of the first column: holding other things fixed and controlling for selection and endogenous participation, Physics Juniors are 2% more likely to participate in projects relative to researchers in fields other than Physics and Engineering. Intuitively, junior researchers are willing to join projects once they take into account who else might be in the project. This is likely because they want to learn from others and gain experience. Moreover, the negative effect of Engineer Juniors is only one-tenth of the effect reported in the first column (-0.037). Hence, there is evidence of endogenous selection. The second and most important finding relates to the effect shown in the last row. On average, an additional teammate decreases the probability of participation by 14% and this effect is statistically significant: researchers dislike working with large groups.

To conclude, two are the main takeaways from the results of Table 5. First, larger teams decrease the probability of project completion. As selection into projects is non-random, controlling for ex-ante potential and endogenous project participation matters for obtaining unbiased estimates of team performance (the effect is ten times larger than without controlling for selection). Second, the larger the number of potential teammates, the lower the probability of joining a project. Hence, there seem to be strategic substitutabilities in teaming up.

Heterogeneous Selection Effects Across Researcher Types

Selection might differ by researcher type. The results in Table 6 explore the effect of heterogeneity of researcher types on the probability of project completion and participation. The first column corresponds to the first column of the previous (Table 5) in terms of covariates; the only difference is that now the outcome depends on the number of project participants of each researcher type.

The first set of covariates relates to the outcome stage. An additional participant (of any type) lowers the probability of project completion. One can see that this effect is not randomly distributed across researcher types. Moreover, it is important to notice that in the reduced-form analysis (columns (5) and (6) of Table 4), holding other things constant, an additional Engineer Junior was associated with an increase in the probability of completion. The results from this Table show that it is no longer the case. Controlling for selection, now an additional Engineer Junior decreases the probability of completion by 8%. The effects are also larger for the other researcher

	Step 1 No Endogenous Participation	**Step 2** Two-Step Pseudo-Likelihood
Outcome Stage		
No. of Physics Seniors	-0.108 (0.003)	-0.107 (0.003)
No. of Physics Juniors	-0.110 (0.002)	-0.110 (0.002)
No. of Engineer Seniors	-0.098 (0.143)	-0.098 (0.140)
No. of Engineer Juniors	-0.084 (0.005)	-0.084 (0.005)
No. of Others	-0.076 (0.003)	-0.076 (0.003)
Dummy for External Firm or Group	-0.096 (0.007)	-0.096 (0.007)
Dummy for Parent Project	-0.112 (0.004)	-0.112 (0.004)
Dummy for Comment	-0.120 (0.004)	-0.120 (0.004)
Participation Stage		
Dummy for Physics Senior	0.037 (0.002)	0.021 (0.002)
Dummy for Physics Junior	-0.036 (0.002)	-0.044 (0.002)
Dummy for Engineer Senior	-0.313 (0.065)	-0.209 (0.064)
Dummy for Engineer Junior	-0.313 (0.003)	-0.104 (0.002)
Dummy for External Firm or Group	-0.037 (0.003)	-0.037 (0.003)
Dummy for Parent Project	-0.013 (0.003)	-0.013 (0.003)
Dummy for Comment	-0.099 (0.002)	-0.099 (0.002)
No. of Potential Teammates for Physics Senior		0.016 (0.003)
No. of Potential Teammates for Physics Junior		0.008 (0.003)
No. of Potential Teammates for Engineer Senior		-0.112 (0.002)
No. of Potential Teammates for Engineer Junior		0.012 (0.002)

Notes: The Table reports the marginal effects at the mean from the structural model. Column (1) reports the results from Step 1 of the two-step procedure. Column (2) reports the results from Step 2. The number of observations is 3,045. The number of potential entrants is 10. All regressions include macro-project and time controls, and a constant. Bootstrapped standard errors in parenthesis.

Table 6: Results Structural Model – Heterogeneity

types. An additional Physics Senior, for instance, lowers the probability of completion by 11% (compared to 3% of the previous Table).

In the second column of Table 6, I estimate the full structural model accounting for heterogeneity in researcher types. The effects of the outcome stage remain similar to those of the first column. The coefficients for the researcher-type dummies in the participation stage change in magnitude but not in sign: being an Engineer Junior, for instance, lowers the probability of participation by 10% (previously 31%).

More importantly, the effect of an additional potential teammate affects the probability of participation differently for each researcher type. This can be seen in the last four rows of column 2. *Ceteris paribus*, with an additional teammate the probability of project participation increases by 2% for a Physics Senior, and by 1% for a Physics Junior and an Engineer Junior. For Engineer Seniors instead, an additional potential teammate decreases the probability of participation by 11%. Table 5 showed that, on average, the higher the number of potential teammates, the lower the probability of joining a project. Table 6 shows that this effect differs across researcher types. This suggests that heterogeneity in researchers' characteristics plays an important role in explaining selection into projects and that there are both strategic complementarities and substitutabilities in teaming up. I discuss the potential mechanisms behind the results below. Other robustness results are discussed in Section G.

Discussion

The two key findings from the estimation of the structural model are that first, controlling for project characteristics and endogenous selection, one more researcher in a project decreases the probability of project completion. Second, controlling for researchers' and project characteristics (including the ex-ante potential), one more teammate decreases the probability of participating in a project.

As already discussed, the first result can be interpreted in terms of congestion and coordination costs. The existence of coordination costs that increase with team size represents an important obstacle for collaborative work (Becker and Murphy, 1992). In fact, lowering coordination costs can increase the returns to collaborative work. Agrawal and Goldfarb (2008) for instance show that a decrease in collaboration costs through the adoption of Bitnet facilitated increased research collaboration between US universities and the specialization of research tasks. Hence, the first finding is in line with the mentioned papers and contrasts with Anderson and Richards-Shubik (2021). Optimal team size hinges on the trade-off between the benefits of specialization and division of labor and the increased coordination costs (Adams et al., 2005); in this setting, the second component plays a prominent role.

Two possible mechanisms can explain the second finding. First, researchers might anticipate that larger teams are allowing for less information about individual contributions. As Jones (2021) discusses, individuals need to take credit when working in teams. There are career concerns and internal signaling, especially for juniors (Meyer, 1994; Jeon, 1996). To find empirical support for this mechanism, it helps to look at the heterogeneity results (Table 6).

Should the mechanism of internal signaling be in place, one would need to find that the costs of working in larger teams are bigger for juniors. As shown in the previous Section, Physics Seniors and Juniors exhibit complementarities in participation. This is plausible as 1) they work in the most relevant field and, in the case of Seniors, 2) they are most likely taking up a managerial position in the team. Engineer Juniors also seem to benefit from working in larger teams. Engineering Seniors instead seem to be disincentivized by working in larger teams. Hence, signaling does not seem to explain the results.

As a second potential mechanism, researchers might internalize the congestion and coordination costs of working in larger teams, and hence suffer more from these costs. This can be particularly true for Engineers as they have less expertise in handling projects related to Physics and, in fact, they exhibit a negative coefficient for the strategic effect. This piece of evidence supports the second mechanism.

To conclude the discussion, the evidence strongly suggests that team size plays a crucial role in the self-formation and performance of teams, as larger teams have higher costs of communication and coordination.

5 Counterfactuals

Ultimately, we are interested in how self-formed teams perform relative to other team structures. It is hard to empirically assess the performance of the same individuals within one organization under alternative team structures. This would require either quasi-random variation in team composition, which is not easy to find in real-world settings, or an experiment that randomly assigns individuals to different teams, which can be costly to implement.

The approach of this paper is ideally suited for this purpose: I use the parameter estimates of the structural model (Table 5) to investigate the performance of alternative team structures with counterfactual exercises. Crucially, I use the unbiased residuals from the regressions of the structural model, i.e., the estimates of the ex-ante project potential that take into account the endogenous selection.

Keeping the projects fixed, I consider teams that instead of self-forming, are exogenously formed by a hypothetical manager. In doing so, I simulate different scenarios leveraging two dimensions: team composition and team size. To assess the performance of the observed self-formed teams relative to that of counterfactual team structures, I compare the average predicted probability of project completion in the data (0.57) with the average predicted probability in the various scenarios.

As a benchmark, I first simulate teams randomly. Random matching of individuals represents a standard benchmark when studying the causal effects of certain outcomes.[36] In the context of an organization, for a manager it might be costly to acquire information about team members and how they efficiently team up. Hence, in terms of organizational practices, one could interpret the random assignment as having no information on team members or having high costs of extracting this information. Under random assignment, I simulate two scenarios: one with random team sizes and another keeping the team sizes the same as the observed ones.

As a second counterfactual exercise, I simulate teams to maximize diversity in terms of field of specialization and professional position. Suppose for instance that a project has four researchers. In this case, I simulate the team being composed of four researchers belonging to four different researcher types (Physics Junior and Senior, and Engineers Juniors and Seniors). Theoretical work has highlighted the main trade-off of diversity: in an environment diverse in terms of skills and knowledge, individuals benefit from information gains; at the same time, they suffer from less efficient communication among members, and this might lower productivity (Lazear, 1999). Hence, it is not clear which effect prevails. Again, under this assignment, I simulate two scenarios: one with random team sizes and another keeping the team sizes the same as the observed ones.

36. For instance, random assignment of students to teams (Calder-Wang, Gompers, and Huang, 2021).

Figure 3 shows the distributions of the average probability of project completion for the four simulated team structures described above: random teams with size kept the same as the observed one (Panel 3a), teams with maximum diversity with size kept the same as the observed one (Panel 3b), random teams with random size (Panel 3c) and teams with maximum diversity and random size (Panel 3d). The blue line in the four panels indicates the average completion in the observed data (0.57).[37]

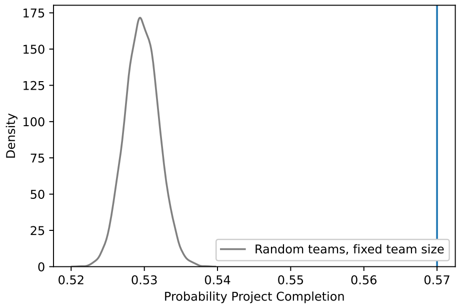
Interestingly, all the distributions except for that in Panel 3b lie to the left of the blue line: all team structures except for one perform worse than the observed self-formed teams in terms of project completion. The distributions in panels 3a and 3c show that randomly composed teams in size and composition have, on average, between 5% and 1% lower probability of project completion relative to actually observed teams. These benchmark results show that self-formed teams exhibit some optimality in teaming up relative to a pure random assignment. Panel 3b shows instead that team diversity improves project completion by about 2% on average when the team size is the same as the observed one. In contrast, in Panel 3d one can see that teams with diverse members and random size perform worse than any other scenario, with a probability of project completion of around 20%. The results are likely a direct consequence of the increased coordination costs associated with larger teams. To sum up, teams that are more diverse in member characteristics perform better than actually observed teams. This holds when team sizes are the same as the observed ones.

The counterfactual results highlight two relevant implications for the organization of scientific production. When researchers can choose their projects, they internalize some of the costs and benefits of working together as they are aware of efficient team size. At the same time, they also tend to work with similar peers, though working with more diverse peers could increase project efficiency (*homophily bias*).

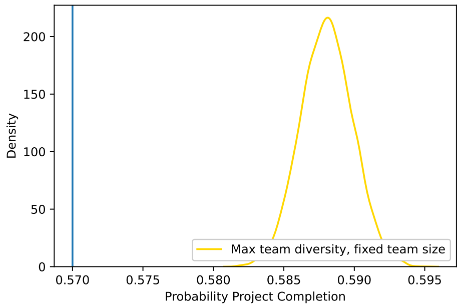
Diversity is often considered to be a crucial condition for radical innovation (Nelson and Winter, 1982; Singh and Fleming, 2010). From an empirical perspective, it has been shown that structurally diverse teams are more likely to produce breakthroughs (Guimera et al., 2005; Jones, Wuchty, and Uzzi, 2008; Banal-Estañol, Macho-Stadler, and Pérez-Castrillo, 2019) and that diversity in endogenously formed teams reduces

37. 10,000 simulations of researcher types drawn from a uniform distribution in $[0, 1]$.

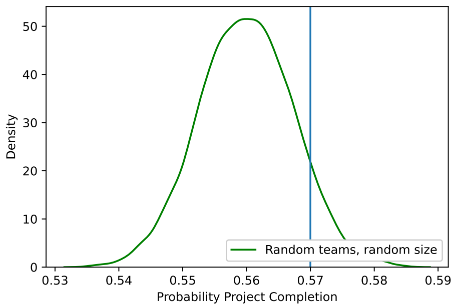
Figure 3: Counterfactual Results Exogenous Team Structure



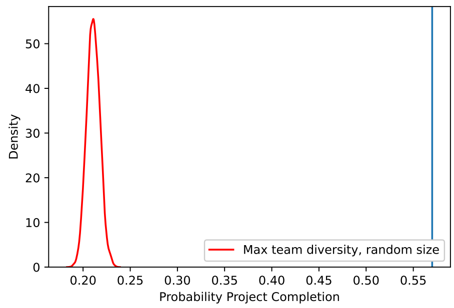
(a) Random Team Structure, Same Size as Actual Teams



(b) Maximum Team Diversity, Same Size as Actual Teams



(c) Random Team Structure, Random Size



(d) Maximum Team Diversity, Random Size

The graphs show the distributions of the probabilities of project completion from 10,000 simulations under the four counterfactual scenarios. The blue bar represents the average predicted probability of project completion in the data (0.57). The left panels show the counterfactuals with researcher types allocated randomly to teams that have the same sizes as the one observed in the data (Panel 3a) and random sizes (Panel 3c). The right panels show the counterfactuals with researcher types allocated to maximize diversity in characteristics with the same team sizes as the one observed in the data (Panel 3b) and random team sizes (Panel 3d).

the negative effect on team performance relative to the case of randomly-assigned teams (Calder-Wang, Gompers, and Huang, 2021). Policymakers have encouraged diversity and interdisciplinarity in scientific institutions.[38] These counterfactual re-

38. ERC Synergy Grant https://erc.europa.eu/apply-grant/synergy-grant.

sults suggest that diversity may increase performance and correct the *homophily bias*. At the same time, one needs to account for decreasing returns to team size as coordination costs increase with larger teams.[39]

6 Conclusion

Leveraging new data from a scientific institution consisting of self-formed teams, this paper develops and estimates a structural model to quantify the main drivers of team formation and performance. The key methodological innovation is to provide an econometric framework that combines both team self-formation and performance into one empirical model that accounts for selection of researchers into projects and allows for testing the performance of alternative counterfactual team structures. There are three main takeaways from the results. First, the larger the number of teammates, the lower the probability of joining a project. Second, the bigger the team, the lower the probability of completion, controlling for project characteristics and selection into projects. These two effects are heterogeneous across researchers. Finally, counterfactual exercises show that randomly composed teams in composition and size have a lower probability of project completion relative to the observed self-formed teams, and teams that are composed to be more diverse in member characteristics (namely, with members who differ in field and seniority) have a higher probability of completion relative to the observed teams, provided that team sizes are the same as the observed ones. Taken together, these results suggest that i) self-formation of teams might be efficient under certain circumstances and ii) diversity may increase performance but it is crucial to account for team size as coordination costs increase when working in larger teams.

The procedure of this paper is flexible enough to be adapted to several other settings. One straightforward application would be to study the mechanisms behind endogenous alliances and partnerships (e.g., R&D joint ventures) and, importantly,

39. Another interesting counterfactual would be to compute the team structure that maximizes the overall probability of project completion. The implementation of this counterfactual might not be straightforward in this setting. The objective function of the manager might not necessarily be to maximize only project completion; moreover, there might be timing decisions that I do not take into consideration here.

their consequences on outcomes. Using counterfactual experiments, one could analyze the consequences of policy restrictions targeted at joint-venture participants on developing patents and more in general on innovation. Another application could be to other team contexts in organizations with some (at least partial) flexibility in team formation. One could test for instance, under which circumstances specific team structures achieve better performance.

So far, the empirical literature has mainly focused on the performance of teams that are exogenously formed. However, many organizations are moving toward a more flexible allocation of workers to teams. This paper provides the first step toward understanding whether self-formed teams are desirable for complex organizations. The approach proposes a tractable empirical framework using a novel source of data and allows for evaluating the performance of counterfactual team structures. This can shed a light on our comprehension of teams' organization and allows us to understand why no man is an island.

References

Adams, James D, Grant C Black, J Roger Clemmons, and Paula E Stephan. 2005. "Scientific teams and institutional collaborations: Evidence from US universities, 1981–1999." *Research Policy* 34 (3): 259–285.

Agha, Leila, Keith Marzilli Ericson, Kimberley H Geissler, and James B Rebitzer. 2018. *Team formation and performance: Evidence from healthcare referral networks.* Technical report. National Bureau of Economic Research.

Agrawal, Ajay, and Avi Goldfarb. 2008. "Restructuring research: Communication costs and the democratization of university innovation." *American Economic Review* 98 (4): 1578–90.

Aguirregabiria, Victor. 2021. "Empirical industrial organization: models, methods, and applications." *University of Toronto, Preliminary version.*

Aguirregabiria, Victor, and Pedro Mira. 2002. “Swapping the nested fixed point algorithm: A class of estimators for discrete Markov decision models.” *Econometrica* 70 (4): 1519–1543.

———. 2007. “Sequential estimation of dynamic discrete games.” *Econometrica* 75 (1): 1–53.

———. 2019. “Identification of games of incomplete information with multiple equilibria and unobserved heterogeneity.” *Quantitative Economics* 10 (4): 1659–1701.

Ahmadpoor, Mohammad, and Benjamin F Jones. 2019. “Decoding team and individual impact in science and invention.” *Proceedings of the National Academy of Sciences* 116 (28): 13885–13890.

Akcigit, Ufuk, Santiago Caicedo, Ernest Miguelez, Stefanie Stantcheva, and Valerio Sterzi. 2018. *Dancing with the stars: innovation through interactions.* Technical report. National Bureau of Economic Research.

Anderson, Katharine A, and Seth Richards-Shubik. 2021. “Collaborative production in science: An empirical analysis of coauthorships in economics.” *The Review of Economics and Statistics,* 1–45.

Bajari, Patrick, Han Hong, John Krainer, and Denis Nekipelov. 2010. “Estimating static models of strategic interactions.” *Journal of Business & Economic Statistics* 28 (4): 469–482.

Banal-Estañol, Albert, Inés Macho-Stadler, and David Pérez-Castrillo. 2019. “Evaluation in research funding agencies: Are structurally diverse teams biased against?” *Research Policy* 48 (7): 1823–1840.

Bandiera, Oriana, Iwan Barankay, and Imran Rasul. 2010. “Social incentives in the workplace.” *The Review of Economic Studies* 77 (2): 417–458.

Becker, Gary S. 2010. *The economics of discrimination.* University of Chicago press.

Becker, Gary S, and Kevin M Murphy. 1992. “The division of labor, coordination costs, and knowledge.” *The Quarterly Journal of Economics* 107 (4): 1137–1160.

Békés, Gábor, and Gianmarco IP Ottaviano. 2022. "Cultural homophily and collaboration in superstar teams."

Berry, Steven T. 1992. "Estimation of a Model of Entry in the Airline Industry." *Econometrica: Journal of the Econometric Society,* 889–917.

Berry, Steven T, and Giovanni Compiani. 2021. "Empirical models of industry dynamics with endogenous market structure." *Annual Review of Economics* 13:309–334.

Bonhomme, Stéphane. 2021. "Teams: Heterogeneity, sorting, and complementarity." *University of Chicago, Becker Friedman Institute for Economics Working Paper,* nos. 2021-15.

Bresnahan, Timothy F, and Peter C Reiss. 1990. "Entry in monopoly market." *The Review of Economic Studies* 57 (4): 531–553.

———. 1991. "Entry and competition in concentrated markets." *Journal of Political Economy* 99 (5): 977–1009.

Calder-Wang, Sophie, Paul A Gompers, and Kevin Huang. 2021. *Diversity and performance in entrepreneurial teams.* Technical report. National Bureau of Economic Research.

Ciliberto, Federico, Charles Murry, and Elie Tamer. 2021. "Market Structure and Competition in Airline Markets." *Journal of Political Economy* 129 (11): 2995–3038.

Ciliberto, Federico, and Elie Tamer. 2009. "Market structure and multiple equilibria in airline markets." *Econometrica* 77 (6): 1791–1828.

De Paula, Aureo. 2013. "Econometric analysis of games with multiple equilibria." *Annu. Rev. Econ.* 5 (1): 107–131.

De Paula, Aureo, and Xun Tang. 2012. "Inference of signs of interaction effects in simultaneous games with incomplete information." *Econometrica* 80 (1): 143–172.

Devereux, Kevin. 2018. *Identifying the value of teamwork: Application to professional tennis.* Technical report. Working Paper Series.

Efron, Bradley, and Robert J Tibshirani. 1994. *An introduction to the bootstrap.* CRC press.

Ganglmair, Bernhard, Timothy Simcoe, and Emanuele Tarantino. 2018. *Learning When to Quit: An Empirical Model of Experimentation.* Technical report. National Bureau of Economic Research.

Gourieroux, Christian, Alain Monfort, Eric Renault, and Alain Trognon. 1987. "Generalised residuals." *Journal of Econometrics* 34 (1-2): 5–32.

Greene, William H. 2003. *Econometric Analysis.* Pearson Education India.

Guimera, Roger, Brian Uzzi, Jarrett Spiro, and Luis A Nunes Amaral. 2005. "Team assembly mechanisms determine collaboration network structure and team performance." *Science* 308 (5722): 697–702.

Hamilton, Barton H, Jack A Nickerson, and Hideo Owan. 2003. "Team incentives and worker heterogeneity: An empirical analysis of the impact of teams on productivity and participation." *Journal of Political Economy* 111 (3): 465–497.

Hill, Ryan, and Carolyn Stein. 2021. *Race to the bottom: Competition and quality in science.* Technical report. Working Paper.

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

Jeon, Seonghoon. 1996. "Moral hazard and reputational concerns in teams: Implications for organizational choice." *International Journal of Industrial Organization* 14 (3): 297–315.

Jones, Benjamin F. 2021. "The rise of research teams: Benefits and costs in economics." *Journal of Economic Perspectives* 35 (2): 191–216.

Jones, Benjamin F, Stefan Wuchty, and Brian Uzzi. 2008. "Multi-university research teams: Shifting impact, geography, and stratification in science." *Science* 322 (5905): 1259–1262.

Katzenbach, Jon R, and Douglas K Smith. 2015. *The wisdom of teams: Creating the high-performance organization.* Harvard Business Review Press.

Lazear, Edward P. 1999. "Globalisation and the market for team-mates." *The Economic Journal* 109 (454): 15–40.

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

McAlpine, Hamish, Ben J Hicks, Greg Huet, and Steve J Culley. 2006. "An investigation into the use and content of the engineer's logbook." *Design Studies* 27 (4): 481–504.

Meyer, Margaret A. 1994. "The dynamics of learning with team production: Implications for task assignment." *The Quarterly Journal of Economics* 109 (4): 1157–1184.

Morgan, John, and Felix Várdy. 2009. "Diversity in the Workplace." *American Economic Review* 99 (1): 472–85.

Nelson, Richard R, and Sidney G Winter. 1982. "The Schumpeterian tradeoff revisited." *The American Economic Review* 72 (1): 114–132.

Olley, Steven G., and Ariel Pakes. 1996. "The Dynamics of Productivity in the Telecommunications Equipment Industry." *Econometrica* 64 (6): 1263–1297.

Schaumans, Catherine, and Frank Verboven. 2008. "Entry and regulation: evidence from health care professions." *The RAND Journal of Economics* 39 (4): 949–972.

Seim, Katja. 2006. "An empirical model of firm entry with endogenous product-type choices." *The RAND Journal of Economics* 37 (3): 619–640.

Singh, Jasjit, and Lee Fleming. 2010. "Lone inventors as sources of breakthroughs: Myth or reality?" *Management Science* 56 (1): 41–56.

Waldinger, Fabian. 2012. "Peer Effects in Science: Evidence from the Dismissal of Scienceists in Nazi Germany." *Review of Economic Studies* 79 (2): 838–861.

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

Yu, Zhengfei. 2021. "Detecting multiple equilibria for continuous dependent variables." *Econometric Reviews* 40 (7): 635–656.

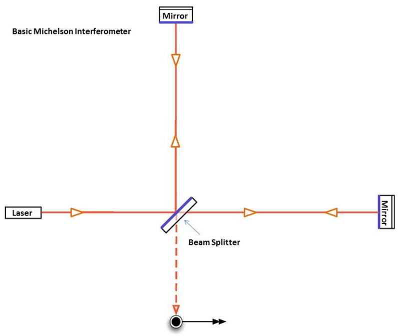
Appendix for Online Publication

A Institutional Setting

A.1 The Interferometer

Figure 4 shows the configuration of an Interferometer (on the left) and the actual Virgo Interferometer (on the right).

Figure 4: Laser Interferometer

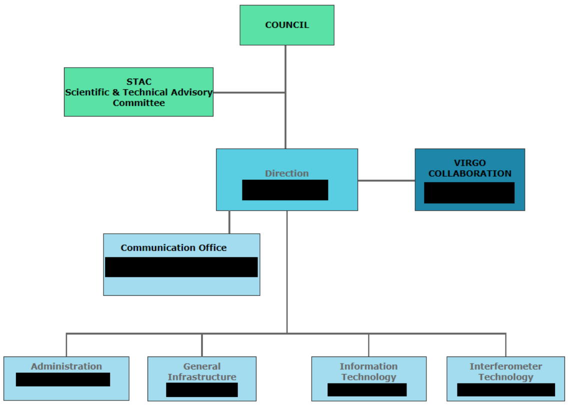


The Figure on the left represents the basic con guration of a Michelson laser interferometer. It consists of a laser, a beam splitter, a series of mirrors, and a photodetector (the black dot) that records the interference pattern (LIGO website: https://www.ligo.caltech.edu/page/what-is-interferometer). The picture on the right represents the Interferometer at Virgo. It is the largest ultra-high vacuum installation in Europe, with a total volume of 6,800 cubic meters. It is placed in Cascina, Italy.

A.2 Organizational Chart Virgo

Figure 5 represents the organizational chart of Virgo for the year 2016. The Director is the head of the organization and is sided by a Spokesperson, who is responsible for external communication together with the Communication Office. The Administration Office takes care of all the bureaucratic and administrative duties related to the institution. The scientific core identifies with three scientific branches: General Infrastructure, Information Technology, and Interferometer Technology. A responsible person is assigned to each branch. The scientific activities of Virgo used for the empirical analysis lie within these three areas.

Figure 5: Organizational Chart Virgo



The Figure presents the general Organizational Chart of Virgo as of 2016. To keep anonymity, the names of the individuals responsible for each o ce have been blacked out.

A.3 Examples of Projects

Projects at Virgo are heterogeneous in terms of skills and work required. Figure 6 shows two images that are related to two different projects.

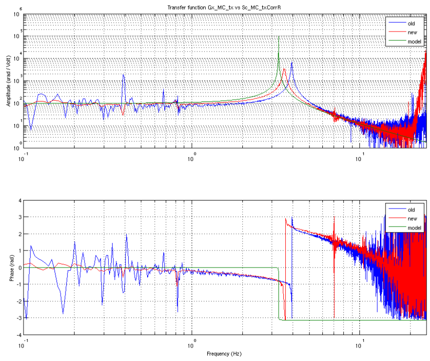
In project 1 (Panel 6a), researchers compare actual measurements with simulations. The text of the log to which the image is attached says the following: “To verify some simulations, we measured some resonances of the MC payload. [...] The measurements agree pretty well, the only big difference seems the change in frequency of the mirror mode from around 4 Hz to 3.5 Hz.” It is clear that this kind of task requires analytical skills and knowledge of simulation analysis.

Figure 2 (Panel 6b) represents a dehumidifier researchers have constructed. The text of the log to which the image is attached says the following: “We have realized and put in operation a system to keep under control the humidity around the SAT mechanical filters stored on the CB high terrace. The system (see fig1) consists of [...]”. Compared to the other project, here a more technical set of skills is required.

A.4 The Logbook

Figure 7 shows a screenshot of a Logbook web page. Each web page of the Logbook consists of logs. A log presents a description of a project; it is identified by the title

Figure 6: Examples of Projects



(a) Project 1

(b) Project 2

The Figure shows two examples of projects at Virgo. The left gure shows two graphs in comparison. The top graph is the result of simulation analysis, while the bottom graph comes from actual measurements. The right gure shows a tool that has been built as a dehumidi er in a room lab.

of the macro-project and the project it refers to, the name of the author(s), the time and date, the (chronological) number, the main text and possibly images, comments or other files attached.

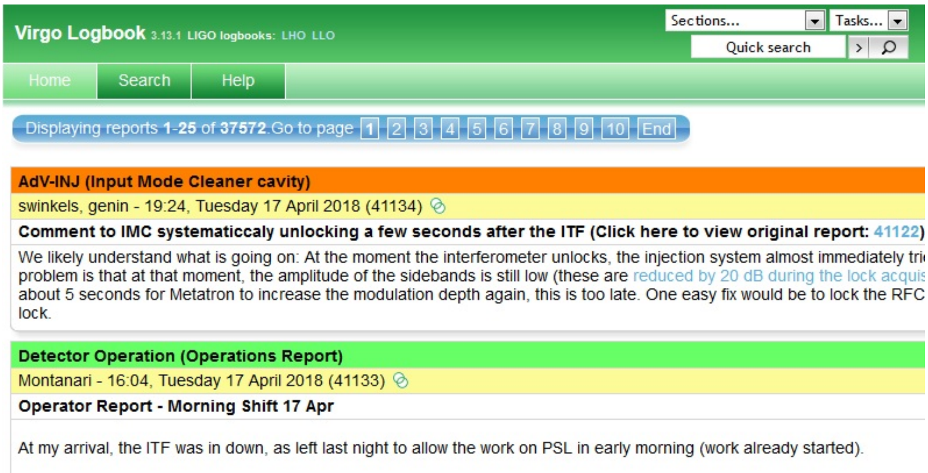
B Data Processing

For the final dataset, I merge two main data sources. The first includes information from the Logbook. The other consists of information about the researchers. In this Section, I discuss in detail the data cleaning, the merging process, and the construction of the final dataset.

B.1 Logbook

The Logbook dataset initially contains 3,778 logs. Each log includes the project title, timestamp and date, macro-project groupings, and the names of the authors.

Figure 7: Example of a Logbook Web page



This web page example consists of two logs belonging to di erent projects. For each log, the rst row identi es the title of the macro-project; the second row identi es the name(s) of the project participants, together with the time and day of the log; the third row identi es the project; the fourth part identi es the actual text of the project. In this example, the rst is a project with two participants, the second is a single-author project.

B.1.1 Authors

The authors of a given project enter their names into the database in plain text; hence, there could be name variations, typos, and potential mistakes in the raw data. For example, a researcher named "John Smith" may have entered their name as "jsmith" on one project, "john smith" on the next, and "jogn smith" on the third. The first step of the data cleaning is to minimize variations of an author's name to get a single consistent version that could then be matched to a known researcher.[40]

After an initial cleaning to remove white spaces and punctuation and to set all logs to the same (lower) case, fuzzy matching is used to identify matches between author logs that may be different versions of the same name. The most basic form of fuzzy matching uses the Levenshtein distance ratio to calculate the similarity between two strings. The Levenshtein distance measures the minimum number of edits (insertions, deletions, or substitutions) used to replace one string with another. The Levenshtein

40. Throughout this Section, \researcher" will refer to a known, identi ed person and the set of personal characteristics in the researcher dataset, whereas \author" will refer to the version(s) of a name (belonging to a researcher, team, or company) entered in plain text in the Logbook logs.

ratio is a similarity calculation between the two strings based on the Levenshtein distance.

The software package used to process the text logs is the “fuzzywuzzy” package (since renamed “TheFuzz”) for Python. The most basic “ratio” method is mathematically identical to the Levenshtein ratio. However, the package also contains more advanced fuzzy matching methods, *partialratio* and *tokensortratio* (a fourth method, *tokensetratio*, was not used). The partial ratio allows for matching only substrings; for example, “JohnSmith” and “Smith” represents a 100% partial ratio match as the substring ”Smith” is identical in both, rather than the lower ratio of a basic Levenshtein ratio given the Levenshtein distance added by the substring “John.” Mathematically, the package takes the number of characters in the shorter string and compares it via the Levenshtein ratio to each possible substring of the same length in the longer string. The token sort ratio method deals with strings of similar length, but with substrings in a different order (for example, “JohnSmith” vs “SmithJohn”). In this case, the package tokenizes the strings and organizes the tokens alphabetically before performing a basic Levenshtein ratio calculation.

The appropriate method depends on how a researcher has entered her name into the data. Hence, all three methods are used to identify different variations of the same name, with any matches of less than 100 but greater than 90% similarity in at least one method of fuzzy matching identified as a potential match. Because of the exponential computational complexity of comparing every entry of a dataset to every other entry directly, the matching is performed on a list of unique entries.

Out of the 3,778 Logbook logs, there are 8,616 different author-name-entries. In some cases, two names are entered into a single cell. After correcting this, there are a total of 8,625 author entries. Of these, 307 unique names are found. Comparing each of the 307 entries to every other entry in the data, 81 matches are found at a threshold of at least 90% similarity with at least one of the three matching methods. Once these matches are reviewed manually, a single name for all possible variants of an author’s name is chosen and the alternative variants are replaced throughout the dataset. After the manual review, the remaining number of unique names is 258, with 5 remaining fuzzy matches, which are either false positive matches or ambiguous situations with unknown authors.

B.1.2 Parent and Comment Projects

A number of projects in the dataset are follow-ups to previous projects. These projects are denoted by the format of the project name, which is "Comment to [name of original project]" followed by an internal reference number. Partial ratio matching from the fuzzywuzzy package can again be used to identify both "parent" projects (projects that have later follow-ups) and "comment" projects (the projects following up the initial parent project). In order to do this, each project identified as a comment project is compared using partial ratio fuzzy matching to every project in the dataset that did not contain "Comment to." In total, there are 947 comment projects. All but 14 can be easily matched to a parent project using a partial ratio match threshold of 95%. Of the remaining 14 unmatched comment projects, 7 are identified manually, leaving only 7 total unmatched comment projects. In total, 616 projects are identified as parent projects by being matched with a comment (note that a parent project may have multiple comments).

After parent-comment pairs are created, dummy variables are constructed for comments (contains "comment to") and parents (matched via partial ratio fuzzy matching to a comment project). Then, to retain additional information about the pair, an additional variable is constructed for each comment project containing the parent project's project number (being NA for every project that is not a comment project), alongside variables for each parent project identifying each comment project (of which there are in some cases up to 9).

B.1.3 Macro-Projects

In addition to time, authors, and title, each log contains a broader group to which the project belongs. In most cases, this is denoted by AdV-XXX and followed by a more detailed clarification in parentheses. For example, one group is "AdV-SBE (MultiSAS for end benches commissioning)." For simplification, a variable containing only the "AdV-XXX" component and not the more detailed clarification is constructed and it is defined as "Macro-project." In total, across the 3,778 logs in the dataset, there are 24 unique "Macro-projects". Of the 3,778 logs, 3,258 belong to a group with contains the format "AdV-XXX", as opposed to projects which belong to a group of a different format. The latter refers to projects related to the previous phase of the experiment and hence are dropped from the final sample.

B.2 Researcher Characteristics

B.2.1 Professional Position and Field Variables

Before merging the researcher characteristics data to the projects, a number of additional researcher characteristic variables are constructed. In order to investigate the role of seniority in team formation, variables representing the seniority of different researchers are of prime interest.

To be able to compare the positions of researchers belonging to different nationalities, information is collected from the following websites: http://www.differencebetween.net/miscellaneous/difference-between-technician-and-technologist, http://www.guide-des-salaires.com/fonction/technicien-datelier, http://www.cnrs.fr/en/join/engineer-technician-permanent.htm, https://cadres.apec.fr/Emploi/Marche-Emploi/Fiches-Apec/Fiches-metiers/Metiers-Par-Categories/Etudes-recherche-et-developpement/charge-de-recherche,https://www.dgdr.cnrs.fr/drhchercheurs/concoursch/chercheur/carriere-en.htm. Figure 7 shows the table of conversion for professional positions.[41]

Each researcher's position is initially categorized into one of four categories depending on the seniority of their position: Senior, Junior, Senior without a degree, and Junior without a degree. The "no degree" category refers to non-academics, such as technicians. The positions classified as Junior are "Ph.D.," "Post-doc," "Researcher, Ingénieur de recherche, Chargé de recherche," and "Ingénieur d'Etudes, Engineer, Technologist," and "Assistant Professor." Positions classified as Senior are: "First Researcher," "Director of research," "Associate Professor," "Full Professor," and "Director Technologist." The Junior without degree position is "Technician, Technicien d'atelier, Assistant ingénieur" and Senior without degree position is "First Technician." Then, a two-category version of the classification system is constructed combining all juniors and all seniors without regard to degree.

41. When I am not able to nd a professional position, I deduce it from the age, h-index, or eld of research. When two di erent levels of seniority are stated, I take the highest.

Table of conversion professional seniority			
Academia	Research Institution (Italy)	Research Institution (France)	Technical Profession (no degree)
PhD		Engineer	
		Technologist	
		Ingénieur d'études	
Post-doc	Post-doctoral fellow		
Researcher/Assistant Prof	Researcher	Ingénieur de recherche	Technician
		Chargé de Recherche	(Technicien d'atelier) Assistant ingénieur
Associate Prof	First Researcher	First Engineer	First technician
Full Prof	Director of Research	Diriger des Recherches	
		Director technologist	

Table 7: Table of Conversion Professional Positions

In addition to seniority, information about the field of a researcher is collected. The researcher's characteristic data contains detailed information regarding the research field, for example, Astrophysics, Cosmology, Optics, Interferometry, etc. These are grouped together into "Engineering," "Physics," and "Others," with the vast majority of researchers (and, when later matched, authors) being in "Physics" or "Engineering." Similarly to the procedure used for the professional position, each researcher's field of specialization is categorized into one of three categories: Physics, Engineers, and Others.

B.3 Datasets Merging

B.3.1 Matching Authors, Researchers, and External Companies

Unique names in the cleaned author dataset have to be matched to a known researcher so that the personal characteristics of the researchers are attached to each of the 3,778 logs in the dataset. This is accomplished using the *tokensortratio* method. Each known researcher's name is compared to each of the cleaned plain text entries in the cleaned author dataset. For each known researcher, the match with the highest similarity is reported. In most cases, this similarity ratio is 100% as the plain text entry is typically the author's last name. The list is reviewed manually to check imperfect and missing matches. With 169 known researchers, all had positive 100% matches except 12. Of these 12, 6 were partial matches determined by manual review to be correct matches, with most partial matches being caused by accented characters in the researcher's name that do not appear in the author dataset. In 3 different cases, authors with similar or identical last names also caused false 100% matches, leaving

9 known researchers unaccounted for in the dataset. These researchers with very similar names are then identified with an "ambiguous match" characteristic variable.

Out of 258 unique author names identified in the data, 150 are accounted for by being matched to a known researcher, leaving 108 unmatched author names. Of these, 52 are known to belong to either a third-party contractor company or a specific predefined group at Virgo, such as the vacuum or electronics teams. This leaves a total of just 56 author names that cannot be identified.

Out of 8,625 author entries, 7,744 can be matched to a specific known researcher. Of the remaining 881 author entries, 665 can be matched to a known company or group, with more than half (395) belonging to the vacuum team. Thus, the total number of unknown author entries in the dataset is 216, namely 2.5% of the total dataset.

B.3.2 Merging Researcher Characteristics to Projects

After each researcher is matched to an author's name, a dataset in wide format associating each project with the characteristics of the researchers involved is constructed. First, variables are constructed for each researcher's characteristics in the researcher data, with an initial value of NA for each observation/project. The researcher data consists of dummy variables for gender, nationality, field, and position.

Then, a nested loop cycles through each of the 3,778 logs and author positions, in order to ultimately check each of the 8,625 author entries against the list of matched researchers detailed in the previous section. When a match is found, the characteristics of that research are attached to the project. This results in a dataset in which each observation is a project that contains all of the variables already attached to each project (project name, time, task, supertask, etc.) and binary dummy variables for each possible characteristic of each possible author position. Based on the classification in terms of professional position and field described in the previous section, for each researcher, a dummy variable given by the combination of these two characteristics is constructed.

Finally, among the 3,778 logs in the merged dataset in wide format, those that do not belong to the Advanced phase of the experiment (147), those that only contain external participants (277), and those that are not classifiable (299) are dropped from the sample, together with logs with more than 8 people (10 projects). The final dataset contains 3,045 logs associated with unique projects.

C Other Descriptive Statistics

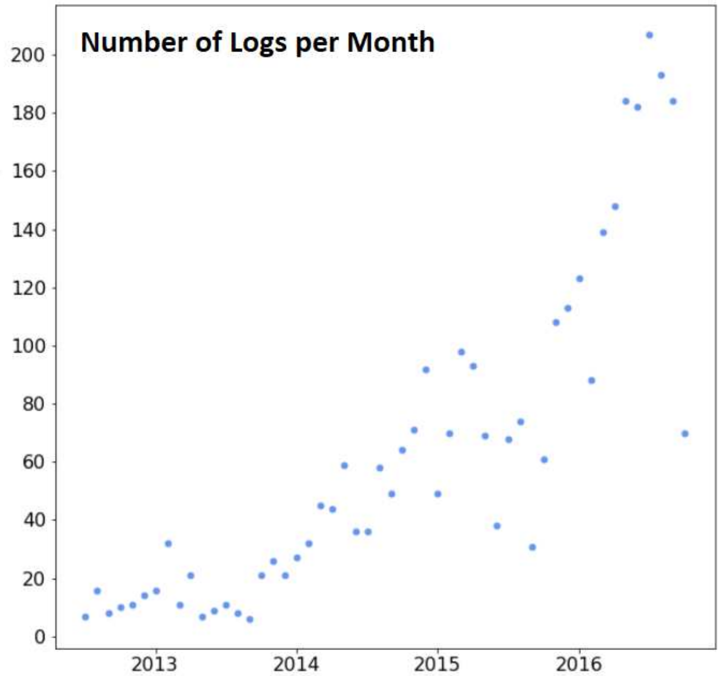
In this Section, I present more descriptive evidence related to the Logbook data. Table 8 shows the frequency of macro-projects in terms of the number of projects. One can see that there is substantial variation, with macro-projects appearing between 23 and 425 times in the sample (Macro-Project 9 only appears once). I construct a categorical variable to account for macro-project controls in the empirical specification.

	Number of Projects
Macro-Project 1	298
Macro-Project 2	424
Macro-Project 3	234
Macro-Project 4	94
Macro-Project 5	214
Macro-Project 6	591
Macro-Project 7	49
Macro-Project 8	23
Macro-Project 9	1
Macro-Project 10	196
Macro-Project 11	124
Macro-Project 12	425
Macro-Project 13	66
Macro-Project 14	70
Macro-Project 15	90
Macro-Project 16	137
Total	3,045

Table 8: Frequency of Macro-projects

Figure 8 reports the average number of monthly logs in the Logbook between June 2012 and October 2016. Over time, a clear upward trend emerges: at the beginning of the period, on average less than 20 projects are carried out monthly. The monthly average goes above 200 at the end of the observation. Interestingly, one can see that there are some forms of seasonality: around December/January, for instance, fewer logs are written. I account for this time trend in the empirical specification by adding month-specific controls.

Figure 8: Monthly Evolution of Logs (2012-2016)



Note: The Figure reports the average number of monthly logs written at Virgo (blue dots) between June 2012 and October 2016.

D Project Outcome

Two categories are identified for the outcome classifications. The categories are the following:
Category 0: Describe a problem or a task proposing possible solutions (with no actual intervention); fix or understand a problem or perform a task temporarily/partially, do a measurement still in progress.
Category 1: Fix or understand a problem, successfully perform a task, and complete or improve a measurement or survey.

For the classification procedure, I make sure that these are sensitive classification criteria also with researchers at Virgo. Methods from Supervised Machine Learning (in particular, classification methods) to determine measures of completion proved less fruitful than manual classification because the jargon of the text is very detailed;

therefore any set of features I gave as inputs to the classifiers was not improving the classification. Hence, I implement the classification manually. Two examples of classified projects are reported below.

Project with classification 0:

Looking at *NARM_LOCK_state* it seems that the lock could hold until around 10 UTC this morning. From that time a series of relocks attempts (with lock periods of different duration) has triggered until around 14:20 UTC were the lock could not be achieved anymore [...].

Project with classification 1:

As foreseen after the completion of Long Towers scaffolding [...] also the DET Tower has been equipped with a Frigerio Style scaffolding. The installation could be completed, yesterday, in a single day [...].

Table 9 presents descriptive statistics for project completion in terms of the number of participants. Notice that 36% of completed projects have one participant (similarly for not completed projects - 37%). Projects with two participants are associated with a probability of completion of 32%, and this percentage decreases to 27% for not completed projects. Finally, projects with more than two participants have a 32% probability of being completed, compared to 36% for not completed projects.

	Number	**%**
Completed Projects	*1,541*	*51%*
One Participant	561	36%
Two Participants	488	32%
More Than Two Participants	482	32%
Not Completed Projects	*1,504*	*49%*
One Participant	561	37%
Two Participants	403	27%
More Than Two Participants	546	36%

Notes: The Table shows the number and % of completed and not completed projects in total and by number of project participants.

Table 9: Descriptive Statistics Project Completion

E Details on Empirical Implementation

E.1 Parametrizing the Model

Following the standard literature on entry (Seim, 2006), equation (1) can be parametrized as:

$$Participation_{ijt} = \alpha + \beta_i X_{it} + \delta_i \bar{N}_j + \eta' D_j + \zeta' Z_t + q_j + \epsilon_{ij} \tag{7}$$

I allow the vector of strategic coefficients δ_i to differ across researcher types. The vector of parameters to estimate for the participation stage is defined as $\theta_1 = (\alpha, \boldsymbol{\beta}, \boldsymbol{\delta}, \boldsymbol{\eta}, \boldsymbol{\zeta})$.

Given the assumption on the information structure, researcher type i joins project j at time t if and only if:

$$\mathbb{E}[Participation_{ijt}] = \alpha + \beta_i X_{it} + \delta_i \, \mathbb{E}[N_j] + \eta' D_j + \zeta' Z_t + q_j + \epsilon_{ij} \geqslant 0. \tag{8}$$

Likewise, the outcome function expressed in equation (3) takes the following form:

$$Outcome_{jt} = \tau + \sum_{g=1}^{G} \kappa_g N_{gj} + \lambda' C_j + \xi' Z_t + q_j. \tag{9}$$

for every researcher type g in the project. The vector of parameters to estimate for the outcome stage is defined as $\theta_2 = (\tau, \boldsymbol{\kappa}, \boldsymbol{\lambda}, \boldsymbol{\xi})$.[42]

E.2 Generalized Residuals

Gourieroux et al. (1987) show that for the logistic distribution, the score vector can be expressed in terms of generalized errors. In the following, I ignore the index t

42. For sake of simplicity and in line with part of the literature on entry games of incomplete information, I assume that the number of potential teammates enters the payo linearly. In alternative speci cations, I allow the outcome to be a quadratic function of the number of researcher types.

for simplicity. Define the log-likelihood as:

$$\ln L = \sum_{j=1}^{J} \log \Psi(Outcome_j | N, C, Z; \theta_2). \tag{10}$$

The first order derivative (score function) with respect to the constant (Greene, 2003) produces the generalized residual. Let ϕ and Φ be respectively the pdf and the CDF of the logistic distribution. Given the parametrization discussed in the previous section, for $Outcome_j = 0$:

$$\tilde{q}_j = E[q_j | Outcome_j = 0, N, C, Z, \hat{\theta}_2] = \frac{-\phi(\hat{\tau} - \hat{\kappa} N_j - \hat{\lambda}' C - \hat{\xi}' Z)}{1 - \Phi(\hat{\tau} - \hat{\kappa} N_j - \hat{\lambda}' C - \hat{\xi}' Z)}. \tag{11}$$

For $Outcome_j = 1$:

$$\tilde{q}_j = E[q_j | Outcome_j = 1, N, C, Z, \hat{\theta}_2] = \frac{\phi(\hat{\tau} - \hat{\kappa} N_j - \hat{\lambda}' C - \hat{\xi}' Z)}{\Phi(\hat{\tau} - \hat{\kappa} N_j - \hat{\lambda}' C - \hat{\xi}' Z)}. \tag{12}$$

E.3 Estimation and Optimization Routine

In the estimation, I minimize the log-likelihood function expressed in equation (5) by comparing the likelihood of the data to the model-predicted analog. First, I define a function that computes the equilibrium probabilities of having 1, 2, etc. potential teammates as the convolution of independent probabilities (in log terms). This results in a matrix of probabilities, where each row is the vector of probabilities for each project. Then, I define the joint likelihood of Step 1. I use as an initial guess of the parameters the estimated parameters of the separated regressions for the participation and the outcome stage. The log-likelihood of Step 1 is the log-likelihood of the participation stage with exogenous entry where the project's unobservable term is expressed as a function of the outcome equation. In other words, I invert the generalized residuals from the outcome stage and replace them in the participation stage. Then, I minimize the joint log-likelihood.

After I compute the predicted probabilities of participation in Step 1, I initialize Step 2 using the parameters and the predicted probabilities from Step 1 to construct the expected number of potential entrants for each project. Following the same procedure I use to construct the log-likelihood of Step 1, I define the log-likelihood of Step 2, with the additional term related to endogenous participation. I calculate the marginal effects at the mean by using the predicted probabilities of Step 2. For the bootstrapped standard errors, I initialize the estimation of the bootstrapped sample at the global minimum found in the main estimation. I fix the random seed and I make 50 draws of sub-samples with 500 projects (Efron and Tibshirani, 1994).

For the main specification, I use the Sequential Least SQuares Programming optimizer (SLSQP). It is an iterative method used for constrained non-linear optimizations, and in case of unconstrained problems (like this one) it reduces to the Newton method. The level of tolerance is set at 1e-8 and the optimal minimum for Step 2 is found after 26 iterations and 642 function evaluations. It takes approximately 2 hours to find the global minimum for the main specification, but the computational time increases exponentially with the number of potential entrants, as it needs to compute 2^n, where n is the number of potential teammates. The same results are found with the Newton Conjugate-Gradient Trust-region method and with different sets of initial values (all zeros, all ones). The log-likelihood goes from an initial value of 19002.29761 in the first iteration of Step 1 to 18291.398057 at the minimum.

E.4 Multiplicity of Equilibria

Several papers discuss the issue of multiplicity of equilibria (Bajari et al., 2010; De Paula, 2013; Aguirregabiria and Mira, 2019). De Paula and Tang (2012) propose a test for the signs of state-dependent interaction effects that do not require parametric specifications of players' payoffs, the distributions of their private signals, or the equilibrium selection mechanism. Yu (2021) provides as a detecting criterion for multiple equilibria the possibility to check discontinuities in the density function of the conditional choice probabilities.

F Other Results

F.1 Reduced-form Results

In this paragraph, I discuss the results from reduced-form regressions of the participation stage (equation (1)) without including the number of potential teammates and without controlling for the project ex-ante potential (the unobserved project component that accounts also for complexity). In other words, I estimate equation (1) ignoring the strategic interactions and not conditioning on q_j. The dependent variable is equal to 1 if a researcher joins a project and 0 otherwise.

	Participation Stage (1)	**Participation Stage (2)**
Dummy for Physics Senior	0.085 (0.009)	0.076 (0.009)
Dummy for Physics Junior	0.003 (0.007)	0.014 (0.007)
Dummy for Engineer Senior	-0.165 (0.011)	-0.156 (0.013)
Dummy for Engineer Junior	-0.049 (0.007)	-0.039 (0.008)
Dummy for External Firm or Group	-0.024 (0.008)	-0.027 (0.008)
Dummy for Parent Project	-0.007 (0.006)	-0.006 (0.006)
Dummy for Comment	-0.081 (0.005)	-0.080 (0.005)
Average Stock of Projects by Type	0.000 (0.000)	0.001 (0.000)
Macro-Project Controls	No	Yes
Time Controls	No	Yes
LL at convergence	-15593	-15588

Notes: The Table reports marginal effects from preliminary regressions of the participation stage. Column (1) includes various research-type and project controls. The number of observations is 3,045. The number of potential entrants is 10. All regressions include the constant. Standard errors in parenthesis. The last row reports the value of the log-likelihood at convergence.

Table 10: Results Participation Stage: No Strategic Interaction

In Table 10 column (1), I include researcher types' dummies (Physics Seniors and Juniors, and Engineer Seniors and Juniors) as well as other project covariates. The excluded dummy for researcher type is Others. One can see that *ceteris paribus*, being specialized in Physics is associated with a positive probability of joining a project (relative to the left-out researcher type Others). For example, for a Physics Senior, this probability is around 9%. The contrary holds for Engineers: being an Engineer Junior is associated with a negative probability of participating of 5%. This is plausible as many projects are in physics and hence require specialized skills in that field.

A project with an external firm or group is associated with a negative and significant probability of participation and the same holds for parent and comment projects. This is plausible as these variables are proxies for complexity. Last, the coefficient for the average stock of projects by researcher type exhibits a positive coefficient, but the effect is very small and insignificant. Results remain similar when adding macro-project and time controls (column (2)), except for the fact that the dummy for Physics Juniors is now significant.

To sum up, the results from Table 10 show that researchers may have different incentives to participate in a project, depending on their characteristics. Though it is hard to interpret these results because I do not control for endogenous participation and selection on unobservables, one could already have an intuition of what may drive the decision to join a project.

F.2 Results Structural Model

In Table 11, I show the results of a specification of the structural model where I include the quadratic term for the number of potential teammates. I perform the analysis to account for non-monotonicities in the strategic interaction coefficients. In particular, I include a quadratic term for the number of potential teammates. One can see that there are indeed non-monotonicities in teaming up, as the quadratic term is negative and significant. This additional result corroborates the idea that researchers suffer from increasing coordination costs associated with larger teams.

G Robustness Analysis

In this Section, I present a number of robustness analyses in support of the empirical findings of Section 4. Note that in this Section I present only the results for

	Stage 2 Two-Step Pseudo-Likelihood
Outcome Stage	
No. of Physics Seniors	-0.121 (0.004)
No. of Physics Juniors	-0.123 (0.004)
No. of Engineer Seniors	-0.112 (0.150)
No. of Engineer Juniors	-0.088 (0.006)
No. of Others	-0.091 (0.004)
Participation Stage	
Dummy for Physics Senior	0.098 (0.003)
Dummy for Physics Junior	0.021 (0.002)
Dummy for Engineer Senior	-0.276 (0.078)
Dummy for Engineer Junior	-0.042 (0.003)
No. of Potential Teammates	-0.133 (0.007)
No. of Potential Teammates^2	-0.012 (0.005)

Notes: The Table reports the marginal effects at the mean from Step 2 of the structural model without heterogeneity. The number of observations is 3,045. The number of potential entrants is 10. All regressions include macro-project and time controls, and a constant. Bootstrapped standard errors in parenthesis.

Table 11: Results Non-Monotonicity Potential Teammates

Step 2 of the analysis.

The decision to participate in one project might potentially affect the decision to participate in subsequent projects because of, for instance, some accumulated expertise. Potential spillovers across projects are a concern for the validity of the estimates as the model would not be able to capture them, given that the decisions to join projects are treated as independent. This could bias the results if the spillovers are due to unobserved factors. If these factors are common to all projects belonging to the same macro-project, adding macro-projects eliminates the concern. Any other unobserved component not captured by macro-project controls can create correlation in unobservables across projects.[43]

43. It is computationally challenging to de ne a correlation structure in unobservables in this

To alleviate this concern, I re-estimate the model using the sub-sample of projects that are far apart in time. It is plausible to think that a project in the early development of the experiment (say, in 2012) is not directly connected to a project at the end of the time period. To do so, I list the projects in chronological order and divide the set of projects into three chunks. Then, I make random draws from the first and last chunks to determine the set of projects in the sub-sample. Results are presented in Table 12.

Notice that the effects remain similar to those found in the main specification of the structural model. In particular, i) an additional project participant of any researcher type decreases the probability of project completion and ii) an additional potential teammate decreases the probability of participation, though the effect is slightly smaller than the one found in Table 5. Hence, correlation across projects does not seem to be a concern for the results.

A second concern is that project participation might be the result of coordinated decisions among teammates. The model of project participation assumes that the participation decision is non-cooperative. Although projects are short-lived and very frequent, and hence there is presumably little time for coordination, I cannot exclude that some coordination might happen offline outside the Logobook, hence generating bias in the results of the participation stage.[44] If anything, this should be stronger for teammates that work together more frequently, as the cost of coordination is much lower relative to less frequent teams.

Hence, to partially rule out this concern, I re-estimate the model excluding from the sample those couples that collaborate together on more than 50 projects.[45] Results are shown in Table 13. Similarly to the other robustness results, the effects are in line with those found in the main specification of the full structural model. This is reassuring as one can partially rule out coordination as a potential factor confounding the results.

Table 14 reports robustness results conducted on a sub-sample of projects carried

framework.

44. Anecdotal evidence shows that this seems not to be the case, as con rmed by researchers working at Virgo.

45. As a criterion, I use couples and not teams because the costs of coordination are likely lower for couples than for larger teams.

	Step 2 Two-Step Pseudo-Likelihood
Outcome Stage	
No. of Physics Seniors	-0.069 (0.003)
No. of Physics Juniors	-0.068 (0.002)
No. of Engineer Seniors	-0.041 (0.070)
No. of Engineer Juniors	-0.069 (0.007)
No. of Others	-0.081 (0.004)
Dummy for External Firm or Group	-0.055 (0.010)
Dummy for Parent Project	-0.075 (0.009)
Dummy for Comment	-0.078 (0.008)
Participation Stage	
Dummy for Physics Senior	0.041 (0.004)
Dummy for Physics Junior	0.012 (0.004)
Dummy for Engineer Senior	-0.176 (0.032)
Dummy for Engineer Junior	-0.018 (0.001)
Dummy for External Firm or Group	-0.065 (0.006)
Dummy for Parent Project	-0.005 (0.000)
Dummy for Comment	-0.105 (0.002)
No. of Potential Teammates	-0.099 (0.008)

Notes: The Table reports the marginal effects at the mean from Step 2 of the structural model without heterogeneity. The number of observations is 1,780. The number of potential entrants is 10. All regressions include macro-project and time controls, and a constant. Bootstrapped standard errors in parenthesis.

Table 12: Results Robustness 1: Projects Far in Time

out exclusively by juniors. The analysis addresses the potential concern that juniors might be forced by seniors to participate in projects with them, undermining the non-cooperative assumption in the participation stage, as discussed before. Also for this case, the estimated effects remain similar in spirit to those found in the main specification (Table 5).

	Step 2 Two-Step Pseudo-Likelihood
Outcome Stage	
No. of Physics Seniors	-0.076 (0.003)
No. of Physics Juniors	-0.076 (0.002)
No. of Engineer Seniors	-0.083 (0.034)
No. of Engineer Juniors	-0.057 (0.008)
No. of Others	-0.064 (0.004)
Dummy for External Firm or Group	-0.047 (0.011)
Dummy for Parent Project	-0.080 (0.008)
Dummy for Comment	-0.069 (0.008)
Participation Stage	
Dummy for Physics Senior	0.040 (0.004)
Dummy for Physics Junior	0.014 (0.004)
Dummy for Engineer Senior	-0.143 (0.030)
Dummy for Engineer Junior	-0.017 (0.001)
Dummy for External Firm or Group	-0.038 (0.005)
Dummy for Parent Project	-0.001 (0.000)
Dummy for Comment	-0.101 (0.003)
No. of Potential Teammates	-0.136 (0.008)

Notes: The Table reports the marginal effects at the mean from Step 2 of the structural model without heterogeneity. The number of observations is 2,057. The number of potential entrants is 10. All regressions include macro-project and time controls, and a constant. Bootstrapped standard errors in parenthesis.

Table 13: Results Robustness 2: No Frequent Couples

Recall from Section 3 that to define for each project the set of potential entrants I make random draws from the pool of available researchers. Based on descriptive evidence, the availability threshold is defined as being involved in four or more projects in the month a project is about to start. In an alternative specification, I lower this threshold to three or more projects. Table 15 reports robustness results from an

	Step 2 Two-Step Pseudo-Likelihood
Outcome Stage	
No. of Project Participants	-0.050 (0.005)
Dummy for External Firm or Group	-0.067 (0.005)
Dummy for Parent Project	-0.098 (0.006)
Dummy for Comment	-0.088 (0.003)
Participation Stage	
Dummy for Physics Junior	-0.110 (0.005)
Dummy for Engineer Junior	-0.069 (0.004)
Dummy for External Firm or Group	-0.039 (0.003)
Dummy for Parent Project	-0.078 (0.004)
Dummy for Comment	-0.115 (0.005)
No. of Potential Teammates	-0.188 (0.006)

Notes: The Table reports the marginal effects at the mean from Step 2 of the structural model without heterogeneity. The number of observations is 1,100. The number of potential entrants is 10. All regressions include macro-project and time controls, and a constant. Bootstrapped standard errors in parenthesis.

Table 14: Results Robustness 3: Only Juniors

alternative specification where the set of potential entrants is defined over a different threshold for availability. As one can see, the effects are very similar to those found in the main specification (Table 5).

In the final robustness, I address the potential concern that the results might be influenced by the size of the pool of potential entrants, as discussed in Section 3.4. Table 15 reports robustness results from an alternative specification where the number of potential entrants is 9 and not 10. As one can see, also for this set of results, the effects are similar in spirit to those found in the main specification (Table 5).

	Stage 2 Two-Step Pseudo-Likelihood
Outcome Stage	
No. of Project Participants	-0.101 (0.001)
Dummy for External Firm or Group	-0.069 (0.005)
Dummy for Parent Project	-0.114 (0.006)
Dummy for Comment	-0.101 (0.004)
Participation Stage	
Dummy for Physics Senior	0.089 (0.002)
Dummy for Physics Junior	0.019 (0.002)
Dummy for Engineer Senior	-0.240 (0.004)
Dummy for Engineer Junior	-0.040 (0.004)
Dummy for External Firm or Group	-0.039 (0.003)
Dummy for Parent Project	-0.042 (0.004)
Dummy for Comment	-0.100 (0.004)
No. of Potential Teammates	-0.127 (0.003)

Notes: The Table reports the marginal effects at the mean from Step 2 of the structural model without heterogeneity. The number of observations is 3,045. The number of potential entrants is 10. All regressions include macro-project and time controls, and a constant. Bootstrapped standard errors in parenthesis.

Table 15: Results Robustness 4: Different Threshold Availability

	Stage 2 Two-Step Pseudo-Likelihood
Outcome Stage	
No. of Project Participants	-0.100 (0.001)
Dummy for External Firm or Group	-0.071 (0.005)
Dummy for Parent Project	-0.112 (0.005)
Dummy for Comment	-0.099 (0.004)
Participation Stage	
Dummy for Physics Senior	0.098 (0.002)
Dummy for Physics Junior	0.016 (0.002)
Dummy for Engineer Senior	-0.257 (0.003)
Dummy for Engineer Junior	-0.039 (0.004)
Dummy for External Firm or Group	-0.039 (0.003)
Dummy for Parent Project	-0.046 (0.004)
Dummy for Comment	-0.098 (0.004)
No. of Potential Teammates	-0.133 (0.003)

Notes: The Table reports the marginal effects at the mean from Step 2 of the structural model without heterogeneity. The number of observations is 3,045. The number of potential entrants is 9. All regressions include macro-project and time controls, and a constant. Bootstrapped standard errors in parenthesis.

Table 16: Results Robustness 5: 9 Potential Entrants