Maastricht University

Geographical Concentration and Editorial Favoritism within the Field of Laboratory Experimental Economics (RM/19/029-revised-)

Janis Cloos, Matthias Greiff,
Hannes Rusch

Geographical Concentration and Editorial Favoritism within the Field of Laboratory Experimental Economics

GSBE

Geographical Concentration and Editorial Favoritism
within the Field of Laboratory Experimental Economics

Janis Cloos, Matthias Greiff and Hannes Rusch[✉]

This version: May 04, 2020

Abstract: We examine geographical concentration, scientific quality, and editorial favoritism in the field of experimental economics. We use a novel data set containing all original research papers ($N = 596$) that exclusively used laboratory experiments for data generation and were published in the *American Economic Review*, *Experimental Economics,* or the *Journal of the European Economic Association* between 1998 and 2018. The development of geographical concentration is examined using data on authors' affiliations at the time of the respective publication. Results show that research output produced by US-affiliated economists increased slower than overall research output, leading to a decrease in geographical concentration. Several proxies for scientific quality indicate that experiments conducted in Europe are of higher quality than experiments conducted in North America: European experiments rely on a larger total number of participants as well as participants per treatment, and receive more citations compared to experiments conducted in North America. Examining laboratory experiments published in the *AER* more closely, we find that papers authored by economists with US-affiliations receive significantly fewer citations in the first 5 and 10 years after publication compared to papers by authors from the rest of the world.

JEL classification: A11, A14, C90, I23

Keywords: laboratory experiments, favoritism, geographical concentration, methodological standards, network effects, gender effects

1 Introduction

Several studies on publications in academic top journals show a dominant position of US research. An important question is whether this dominance is the product of the sheer size of the US, of discrimination, or of other barriers to entry for researchers from the rest of the world. In the last decades, research on geographical and institutional concentration has received a lot of attention. For economic top journals, several studies show that a large but declining share of research is produced by US-based economists (Ek and Henrekson 2019, Glötzl and Aigner 2019, Kocher and Sutter 2001, Kalaitzidakis et al. 1999, Hodgson and Rothman 1999, Elliott et al. 1998, Frey and Pommerehne 1988).

In this paper, we focus on one particular field: experimental economics. Our analyses cover all laboratory experiments published between 1998 and 2018 in three important journals: the *American Economic Review* (AER), arguably one of the top journals for general economics worldwide, *Experimental Economics* (Exp Econ), the top field journal, and the *Journal of the European Economic Association* (JEEA), arguably the European top journal for general economics.

Our first contribution is an analysis of geographical concentration. We examine its development using data on authors' affiliations at the time of publication. Similar to previous work, our analysis reveals that although the US still has a dominant position, its dominance has decreased over time.

As a second contribution, we examine the gender composition of authors. We find that less than 20% of the total number of authors in our data set are women. More than 60% of the papers in our data set were written solely by male authors. For the field of experimental economics we do not observe an increase in the relative share of female authors. These observations are very similar to the results of a recent study by Hengel and Moon (2020) who consider a much larger number of papers from economic top journals.

Our third contribution is an analysis of several proxies capturing the scientific quality of the published experiments. Drawing on the literature on experimental methods, we identify four proxies that are indicative of the quality of laboratory experiments and which can be determined by the experimenters: the total number of participants, the number of participants per treatment, the number of treatments, and the strength of monetary incentives (measured as real 2015-USD

per average participant-minute). To ensure comparability between experiments, we restrict our analyses to laboratory experiments and exclude all other types of experiments.

Our results show substantial differences between Europe and North America, the two regions where more than 80% of the experiments were conducted. For all three journals, experiments conducted in Europe have a larger number of participants per treatment. In the AER and in Exp Econ, experiments conducted in Europe also have a larger total number of participants. We find no differences in the strength of monetary incentives. Hence, the (inflation-adjusted) total costs of experiments are higher for experiments conducted in Europe.

Our fourth contribution is to test whether the dominance of US-affiliated economists on the editorial board of the AER gives rise to editorial favoritism. Editorial favoritism can be a serious problem for scientific progress (Brogaard et al. 2014, Medoff 2003, Laband and Piette 1994a). It arises when editorial decisions are not exclusively based on scientific criteria but also on non-scientific criteria, such as authors' institutional connections: if non-scientific criteria influence editorial decisions, colleagues with social ties to the members of a small elite of editors will find it "easier" to get their papers accepted. Such favoritism is a problem when the quality of published research suffers from it, i.e., when the papers from editors' buddies are published while better papers by other scholars are rejected.

Unlike the existing literature (Brogaard et al. 2014, Medoff 2003, Laband and Piette 1994a), our analysis of editorial favoritism clusters groups by geographical regions (US and Europe) and not by individual institutions (e.g., universities). As we lack data on submissions, we focus on published papers. In the presence of editorial favoritism, we expect papers from authors without social ties to the editors to be of higher average quality than papers from authors with social ties to the editors.

In our analyses of possible editorial favoritism, we also control for the gender composition of author teams. In recent papers, Card et al. (2020) and Hengel and Moon (2020) found that female-authored papers receive more citations than male-authored papers. Here, we are able to analyze whether such a gender citation premium also exists in the field of experimental economics. We confirm this finding for the AER and the JEEA, albeit the effect not reaching statistical significance in all specifications.

The number of participants and treatments, and the strength of monetary incentives constitute a set of *ex ante* proxies for quality that are available to editors when deciding about publication.

However, originality, timeliness, and practical relevance are also strongly related to a paper's quality, but are not captured by our proxies for quality. Therefore, we use the number of citations that a paper received in the 5 and 10 years after publication as an *ex post* proxy for quality. Assuming that editors exclusively aim to maximize their journal's impact factor (which is debatable, but see Card and DellaVigna 2020), there should be no systematic difference in the quality of papers authored by economists with and without affiliations on the editorial board's continent. However, this is not always the case. We find that AER papers (co-)authored by economists with a US affiliation receive, on average, fewer citations in the 5 and 10 years following publication compared to papers by European authors published in the same outlet. This pattern was not found for papers published in Exp Econ and the JEEA, suggesting that the AER's selection criteria for research based on laboratory experiments fail to maximize the journal's impact factor, or that editors pursue other goals[1].

To the best of our knowledge, this is the first study in which the quality of laboratory experiments is compared across journals and geographical regions. By incorporating objectively measurable quality proxies into our analysis, we shed some light on the question of to what extent (differences in) citations are driven by (differences in) quality characteristics, social ties, and gender. In this respect, our analysis goes beyond the existing work on editorial favoritism. Surprisingly, and despite the large differences in the total number of participants and participants per treatment, we find no robust evidence for any effect of these quality proxies on the number of citations. Yet, we are confident that our results provide insights for editors on how to improve their journals' selection criteria and to increase their journals' impact factors.

The remainder of this paper is organized as follows. We start by reviewing the related literature in section 2. In section 3, we describe our data set and our selection criteria for journals and papers before we examine geographical concentration and the gender composition of authors in section 4. In section 5, we compare laboratory experiments conducted in Europe and North America using our proxies for quality. We describe our tests for editorial favoritism in section 6 and conclude in section 7.

---

[1] We decided to focus on editors and not on referees for three reasons. First, editors have the power to overrule referees' suggestions. Second, editors tend to select referees to whom they have easy access, implying that social ties between author and editor and between author and referee are correlated (Hamermesh 1994). And third, data on referees is hard to obtain (an exception is Card et al. (2020)).

2 Related Literature

2.1 Geographical Concentration

Kalaitzidakis et al. (1999) look at ten high-quality economics journals from 1991 to 1996 and show that between 69% and 90% of all authors have an affiliation in North America. Based on all papers in 15 top economics journals between 1977 and 1997, Kocher und Sutter (2001) find that 72.2% of these papers were authored by economists working in the US. They also find that US dominance reduced over time. Ek and Henrekson (2019) look at the top-five economics journals and reveal that between 82% (in 1994) and 65% (in 2017) of all authors have an affiliation in North America, and 95% of the North American authors have their affiliation in the US. Other studies find similar results regarding the decline in US dominance (Glötzl and Aigner 2019, Neary et al. 2003)[2]. For the field of laboratory experiments, we expect a similar pattern.

2.2 Social Ties and Editorial Favoritism

Editorial favoritism occurs when it is easier for an economist with a social tie to one or more editors to get her paper accepted, compared to a paper of the same quality submitted by an economist without such a social tie. Editorial favoritism might occur if, for example, editors favor their former PhD students, or authors employed at the same university or the same region[3].

A relatively high share of papers by authors who are connected to one of the journal's editors is often taken as an indication for editorial favoritism. However, this is neither necessary nor sufficient. As pointed out by Brogaard et al. (2014, 252), the decisive question is "…whether editors use information advantages to improve selection decisions, or whether they bow to conflicts of interest."

---

[2] US dominance can be observed not only among authors but also among editors. Hodgson and Rothman (1999, 168 - 170) look at the top 30 economics journals and show that more than 83% of all editorial board members are from the US. In section 6.1, we describe the composition of the editorial boards for the journals we examined.

[3] Discrimination implies that papers with a lower quality are accepted for publication, while papers with a higher quality are rejected. If the paper's true quality is somehow revealed after publication (possibly, by the number of citations), this might damage a journal's reputation (a loss in the journal's impact factor). So how can discrimination persist? It can persist as long as the true quality is not revealed, or if there is a form of collusion between editors of different journals, who all discriminate in the same way. In experimental oligopoly markets, it has been found that a shrinking market facilitates collusion (Abbink and Brandts 2009). In economics, we have a similar situation. As more and more journals are being published in English but outside the US (e.g., the JEEA), the US-market is shrinking. However, rather than investigating discrimination over time, we focus on the possible existence of discrimination in one particular field.

The earliest study on editorial favoritism in economics by Laband and Piette (1994a) analyzes all papers published in 28 top economics journals in the year 1984. Laband and Piette define a social tie as one of the authors and one member of the editorial board having received their PhD from the same university or one of the authors and one member of the editorial board being affiliated with the same university in 1984. Controlling for papers' length, journal quality, author's age, gender, and reputation, they find a positive correlation between the number of citations in the 5 years after publication and the existence of a social tie. They note that few well-cited high-quality papers mainly drive this positive correlation. However, their results also show that economists who had social ties to the editors authored more than two thirds of the papers with remarkably few citations. Thus, Laband and Piette find no clear evidence in favor of or against editorial favoritism. Based on their results, they conclude that "on balance, journal editors use their professional connections to search out good papers, rather than print substandard material written by their colleagues or buddies" (Laband and Piette 1994a, 200). Medoff (2003) analyzes all papers published in six top economics journals in 1990 and also finds that papers by authors with a social tie to the editors or co-editors receive a higher number of citations in the years following publication. He examines the citations received in the 5 and 10 years after publication and observes that the positive effect of social ties is stable over time.

More recently, several papers have examined further effects of social ties. Based on an analysis of 50,000 articles in 30 major economics and finance journals since 1955, Brogaard et al. (2014) find that the publication rates of authors in these journals increase by 100% as soon as one of their current colleagues is the editor in charge of the corresponding journal. Their results show that articles by authors with social ties to editors receive significantly more citations than articles by authors without such connections. Colussi (2018) analyzes all 1,620 papers published in four top-journals (AER, *Econometrica*, the *Journal of Political Economy*, and the *Quarterly Journal of Economics*) between 2000 and 2006. Colussi's results show that editors are more likely to accept papers from colleagues (an 8% increase at the baseline) and former PhD students (a 14% increase at the baseline). Further, he shows that about 43% of all papers were written by an author connected to at least one editor in charge of the journal at the time of publication. However, since the base-rate of submitting authors is unknown, this value is difficult to interpret. Heckman (2017) finds a similar pattern in two house journals, the University of Chicago's *Journal of Political Economy* and Harvard's *Quarterly Journal of Economics*. Using data on publications from 2000 to 2016, he shows that 14.3% of all papers in the *Journal of Political Economy* come from authors affiliated with the University of

Chicago, and 24.7% of all publications in the *Quarterly Journal of Economics* stem from authors affiliated with Harvard.

Taken together, the results from previous studies suggest that (1) social ties can improve the chances of getting published because they reduce editors' search costs for high-quality papers, and (2) that conditional on being accepted for publication, on average, papers from authors with social ties receive more citations. None of the studies finds clear evidence for editorial favoritism (but see Shepherd 1995, for anecdotal evidence).

The studies discussed above focus on editorial favoritism and the effects of social ties. They discuss social ties exclusively at the institutional level but not at the regional level. Frey and Pommerehne (1988, 106-107) point out that it may also generally be easier for American economists to publish in American-dominated journals as opposed to non-American economists.

Studies on editorial favoritism (Medoff 2003, Laband and Piette 1994a) use data on individual authors (e.g., affiliations, age, gender, reputation) and characteristics of the papers (e.g., length, JEL code, position within the journal issue). These characteristics, however, are not related to any methodological standards. In section 5.1, we argue that there are four proxies that capture different methodological standards for laboratory experiments. Our analysis of editorial favoritism for laboratory experiments includes these four proxies, and thus extends the existing literature on editorial favoritism.

2.3 Gender Effects

The primary interest of our paper is to investigate the geographical concentration of authors and possible signs of editorial favoritism in the field of laboratory experimental economics. In light of a growing literature addressing the underrepresentation of women in economics (i.e., Lundberg and Stearns 2019, Bayer and Rouse 2016) and finding quality-relevant gender effects (Card et al. 2020, Hengel and Moon 2020, Hengel 2019), we investigate the effect of an authors' gender on an articles' quality in this article as well. Below, we briefly address the main findings of recent studies that examine the influence of the authors' gender on the *ex ante* and *ex post* quality of articles[4].

---

[4] For more detailed explanations of these gender effects, we recommend the studies by Card et al. (2020) and Hengel (2019).

Analyzing the readability scores of abstracts of more than 9,000 articles published in four of the top-5 economics journals, Hengel (2019) finds that female authored articles have a higher writing quality than male authored articles. Her results show that the gender gap in readability scores increases from the working paper to the published paper. This suggests that stricter requirements are imposed on female authors during the review process. In line with this result, for *Econometrica* papers, Hengel observes that peer review for female-authored papers takes on average six months longer. She therefore concludes that editors demand higher quality standards from female authors. In a follow-up paper, Hengel and Moon (2020) use citation data of more than 11,000 papers published in the top-5 economics journals[5] to examine whether the authors' gender affects the number of citations a paper has received in the years after publication. Controlling for authors' reputation, number of authors, JEL code, and journal-year fixed effects, their results show that female authored papers receive more citations than male authored papers. Hengel's and Moon's results also illustrate that male authors' citations increase when they co-author with female authors while female authors' citations decrease when they co-author with male authors.

Card et al. (2020) examine the relationships between the gender of authors and referees, the referee recommendations, the decisions of editors, and the citations a paper receives. Using data from nearly 30,000 papers submitted to the JEEA, the *Quarterly Journal of Economics*, the *Review of Economics and Statistics*, and the *Review of Economic Studies*, and controlling for the respective field and authors' reputation they find that female-authored papers receive 25% more citations than male-authored papers. They further find evidence that editors, irrespective of referees' and authors' gender, largely follow the recommendations of their referees. Card et al. argue that editors are gender-blind and that both female and male referees are equally likely to set higher standards for female authors.

3 Data on Laboratory Experiments
3.1 Selection Criteria

Although we are interested in the broad field of experimental economics, we decided to focus exclusively on papers that report results from laboratory experiments. Precisely, we focus on laboratory experiments that generate data in a controlled process using student participants who

---

[5] The journals included in Hengel (2019) are the AER, *Econometrica*, the *Journal of Political Economy*, and the *Quarterly Journal of Economics*. Hengel and Moon (2020) additionally include articles from the *Review of Economic Studies*.

interact in an artificial environment[6]. In order to ensure the greatest possible comparability, we do not consider papers that contain other types of experiments[7].

For laboratory experiments, objectively measurable proxies for an experiment's quality exist. This is because, compared to field experiments, laboratory experiments have the exclusive advantage that the experimenter has a high degree of control. Proxies for an experiment's quality are the total number of participants, the number of treatments, the number of participants per treatment, and the strength of monetary incentives. We describe these proxies in more detail in section 5.1.

There is no doubt that field experiments play an important role in experimental economics (for a detailed discussion see Czibor et al. 2019, Carpenter et al. 2005, Harrison and List 2004). However, in field experiments, at least one of our quality proxies is out of the experimenter's control. For example, an artefactual field (or lab-in-the-field) experiment relies on a population of participants specifically selected for a given research question (e.g., chess players in Bühren et al. 2012 and Levitt et al. 2011) that cannot be fully controlled by the experimenter. Similarly, framed field experiments take place in a natural environment (e.g., the National Rural Support Programme Offices in Afzal et al. 2017) which also cannot be fully controlled by the experimenter. In natural field experiments, moreover, both the participant population and the environment cannot be controlled exogenously.

3.2   Our Data Set

Our data set contains 596 papers published between 1998 and 2018 (410 from Exp Econ, 145 from the AER (including papers from AER's Papers and Proceedings[8]), and 41 from the JEEA)[9]. We chose 1998 as the starting year for our analysis because Exp Econ was founded that year. Exp Econ is the specialized field journal of the *Economic Science Association*. It publishes laboratory and field experiments as well as related theoretical papers and reviews. We chose the AER and the JEEA because they are the official journals of the *American Economic Association* and the *European Economic Association*, respectively. Both are general

---

[6] Our data contain papers where the majority of the participants were students. Papers where a large share of participants belonged to special groups (such as job professionals or caste members) are not included in the data set.

[7] For example, the AER paper by Imas (2016) is not included in our data set because the paper contains two laboratory experiments and a robustness check conducted online via Amazon M-Turk.

[8] Since 2018, the *AEA Papers and Proceedings* are available as an independent journal. Our data set contains only those papers from Papers and Proceedings that were published in the May issue of the AER in the years before 2018.

[9] A full list of the included papers can be found in the supplementary material.

interest journals and publish only a small share of laboratory experiments[10]. We have decided to collect data from these relatively high ranked journals to ensure that the individual papers have received a certain number of citations and because we assume that these journals receive a high number of submissions from different geographical regions.

For each paper in our data set, we collected data on:
- the authors and their affiliations at the time of publication,
- the number of female authors,
- the number of citations received by the individual authors in the 5 years before publication,
- the total number of pages,
- the total number of references,
- the journal's impact factors in the year the paper was published, and
- the paper's number of citations in the 5 and 10 years after publication.

In addition, we extracted data on the characteristics of the laboratory experiments reported in the papers. If available, for each laboratory experiment we collected data on:
- the total number of participants,
- the number of treatments,
- the duration of the experiment,
- average earnings per participant, and
- the year in which the experiment was conducted.

Citation data was obtained from *Web of Science*. Data on impact factors were taken from *CitEc*, a *RePEc* service that provides citation data for economics[11]. In line with Hengel and Moon (2020), we determined an author's gender either by clearly gendered first names, by photographs on university or personal websites, or by personal pronouns used in texts about the author. All other data were obtained directly from the published papers. Since not all papers report data on all characteristics, our data set contains some papers for which some or all of the

---

[10] On the IDEAS/RePEc list of Simple Impact Factors for Journals, the AER is ranked 10th, the JEEA is ranked 19th and Exp Econ is ranked 35th (see https://ideas.repec.org/top/top.journals.simple.html, accessed August 29 2019). On the latest version of the Journal Quality List of Anne-Wil Harzing, which provides an overview of several different ratings, the AER is rated A+ and the JEEA and Exp Econ are both rated A (see https://harzing.com/download/jql_title_2019-02.pdf, accessed August 29 2019).

[11] For the AER see, http://citec.repec.org/s/2018/aeaaecrev.html, for Exp Econ see, http://citec.repec.org/s/2018/kapexpeco.html and for the JEEA see, http://citec.repec.org/s/2018/blajeurec.html (accessed July 9, 2019). Data for Exp Econ's impact factor for 1998 and for the JEEA's impact factor before 2011 was not available.

experiment's characteristics are missing. In case of missing values, we tried to obtain the data from working paper versions or directly from the authors, which was successful in some cases.

4 Geographical Concentration and Gender Composition in Experimental Economics

In this section, we examine the geographical concentration of published laboratory experiments at the country-level and the gender composition of author teams. In section 4.1, we look at country shares for each journal for the longest available period. In section 4.2, we examine how geographical concentration has changed over time. In section 4.3, we scrutinize the gender composition of author teams in our dataset and compare it between the three journals.

We measure geographical concentration as follows (see also Combes and Linnemer 2003, Kocher und Sutter 2001). Let IF_i be the 2-year impact factor of the journal in which paper i was published in the respective publishing year. Let n_i be the number of authors of paper i. An author's weight is then given by IF_i/n_i. The weighted score for a single country and a given time period is calculated by aggregating the values of all authors whose affiliated institution was located in this country (at the time of publication). Given the weighted scores of all countries, we compute each country's share.

4.1 Pooled Over Time

Table 1 shows the shares for selected countries[12]. The first three columns depict the results by journal while the last column presents the aggregate results for laboratory experiments published in all three journals. *Table 15* (appendix A) lists all countries' scores and shares by journal. Our results remain qualitatively similar when individual papers are not weighted by the journal's impact factor but rather assigning each author a weight of $1/n_i$.

---

[12] We have assigned Israel, Russia and Turkey to Europe and Guatemala and Mexico to North America throughout the whole paper.

	Shares			
	AER	**Exp Econ**	**JEEA**	**all three**
Austria	2.63	4.86	0.00	3.60
France	0.65	3.03	2.94	2.20
Germany	4.41	12.08	13.29	9.57
Israel	1.48	0.84	0.00	0.98
Italy	0.39	3.50	1.55	2.24
Netherlands	3.30	4.68	7.35	4.47
Norway	0.54	0.68	6.31	1.20
Spain	2.68	5.88	2.04	4.40
Sweden	0.71	1.25	0.00	0.94
Switzerland	5.01	2.24	20.05	4.98
UK	6.08	8.80	9.50	7.94
Other European countries	0.78	1.89	0.00	1.36
Europe	**28.65**	**49.73**	**63.03**	**43.82**
US	63.33	37.84	29.58	45.77
Other North American countries	1.79	3.06	1.53	2.47
North America	**65.11**	**40.90**	**31.11**	**48.24**
Australia and NZ	2.41	4.99	2.07	3.81
Asia	3.32	4.09	2.65	3.68
South America	0.50	0.29	1.15	0.45
Other continents	**6.23**	**9.37**	**5.87**	**7.94**
Total	**100.00**	**100.00**	**100.00**	**100.00**

Table 1: Geographical concentration, countries' relative shares for selected countries, pooled data from 1998 to 2018 for the AER, from 1999 to 2018 for Exp Econ and from 2011 to 2018 for the JEEA. Continents are marked in bold.

We see that US economists are dominant in all journals, with this dominance being most pronounced for the AER. When looking at the European countries, Germany, the UK, Switzerland, Spain, the Netherlands, and Austria stand out because they have the highest outputs within Europe.

4.2 Development Over Time

Looking at the development over time, we divide the 20-year period from 1999 to 2018 into four five-year periods[13]. *Table 2* presents the scores and shares of all countries with a share larger than 1% (for the whole period from 1999 to 2018). Scores and shares of all countries can be found in appendix A *Table 16*. The analyses in *Table 2* and *Table 16* only contain data on papers from the AER and Exp Econ, as CitEc impact factors for the JEEA are only available for 2011-2018.

The first result that sticks out is the sharp increase in the field's research output. Looking at the total output by aggregating scores over all countries (*Table 2* and appendix A, *Table 16*), we see an increase from 108.50 (1999-2003) to 282.86 (2004-2008) to 608.36 (2009-2013)

---

[13] We exclude observations from 1998, as CitEc impact factors for Exp Econ are only available from 1999 onward.

followed by a decrease to 422.38 (2014-2018). The decrease in the most recent period could be due to four reasons: first, field and online experiments have developed into an increasingly established research method in recent years (see e.g., Stewart et al. 2015), partially replacing lab experiments[14]. Second, a larger fraction of papers use multiple methods for data collection which are not included in our data set. Third, other journals may have become more open toward laboratory experiments or new journals that publish laboratory experiments have been established. And fourth, editors' tastes may have changed so that other research methods are preferred to laboratory experiments.

Based on the scores (appendix A, *Table 16*), we computed compound growth rates. Aggregated over all countries, the field's research output increased by 57.31% from each five-year period to the next. As expected, the compound growth rate for US-economists is much smaller (29.51%). High growth rates are observed for Australia (211%), Austria (192%), Italy (157%), China (133%), Germany (118%), the UK (122%), the Netherlands (110%), and Japan (100%). Of course, these growth rates are partly due to the fact that some countries started from very low levels of research output in the period 1999-2003.

---

[14] This is similar to the results in Nikiforakis and Slonim (2019). They show that the share of published field experiments in the top-5 journals increased steadily until 2015, but has been decreasing since then. For the periods 2009-2010 to 2017-2018, they report a 66% drop in published laboratory experiments in the top-5 journals.

	Shares				
	Scores				
	1999-2003	**2004-2008**	**2009-2013**	**2014-2018**	**1999-2018**
Austria	0.61	0.12	6.53	3.98	3.99
	0.67	0.34	39.71	16.63	57.34
France	2.93	0.19	2.51	2.68	2.11
	3.18	0.55	15.27	11.33	30.33
Germany	4.66	5.46	9.50	12.47	9.13
	5.06	15.43	57.80	52.66	131.28
Israel	1.05	0.77	1.72	0.43	1.15
	1.13	2.18	10.44	1.82	16.58
Italy	0.60	1.22	2.94	2.63	2.34
	0.65	3.44	17.91	11.09	33.60
Netherlands	2.27	4.15	3.67	5.42	4.13
	2.46	11.73	20.30	22.88	59.38
Spain	1.66	2.77	8.13	1.80	4.66
	1.80	7.85	49.46	7.59	67.02
Sweden	0.99	1.68	0.23	1.84	1.04
	1.07	4.74	1.37	7.76	14.95
Switzerland	8.96	1.29	4.31	1.81	3.28
	9.72	3.64	26.24	7.64	47.24
UK	3.00	6.45	8.86	8.46	7.83
	3.25	18.24	53.93	35.72	112.64
Other Euro. countries	1.74	0.73	1.14	2.37	1.45
	1.89	2.07	6.95	10.01	20.92
Europe	**28.47**	**25.86**	**49.54**	**45.25**	**41.73**
	30.89	**73.16**	**301.38**	**191.13**	**600.22**
US	65.92	68.88	41.17	36.79	47.55
	71.52	194.84	250.48	155.38	683.91
Other NA countries	1.55	0.87	3.05	3.35	2.59
	1.68	2.46	18.57	14.14	37.18
North America	**67.46**	**69.75**	**44.23**	**40.13**	**50.14**
	73.20	**197.3**	**269.05**	**169.52**	**721.09**
Australia and NZ	2.28	1.55	4.41	5.69	3.99
	2.48	4.39	26.84	23.63	57.33
Asia	1.79	2.83	1.72	7.92	3.78
	1.94	8.01	10.45	33.47	54.33
South America	0.00	0.00	0.11	1.10	0.37
	0.00	0.00	0.64	4.63	5.27
Other continents	**4.07**	**4.38**	**6.24**	**14.61**	**8.14**
	4.42	**12.40**	**37.93**	**61.73**	**116.93**
Total	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**
	108.50	**282.86**	**608.36**	**422.38**	**1438.26**

Table 2: Shares and scores for the geographical concentration of AER and Exp Econ authors for selected countries. Data pooled over different periods. For each country, the first row reports the country's share and the second row reports the country's score.

Looking at countries' relative shares (the upper row of each country in *Table 2* and *Table 16*), we find that the US-dominance decreased from about 65.92% (1999-2003) to 36.79% (2014-2018) while several countries increased their relative shares. The countries that gained most are Germany (7.81%), the UK (5.46%), Australia (4.38%), Austria (3.32%), and the Netherlands (3.15%). These results imply that geographical concentration has decreased. Indeed, this is

confirmed if we look at the Herfindahl index[15] which first increased from 0.337 (1999-2003) to 0.485 (2004-2008) but then decreased to 0.181 (2009-2013) and to 0.171 (2014-2018). In sum, our data show that, for laboratory experimental economics, the geographical concentration of AER and Exp Econ authors has substantially decreased.

The discussion above focused on the development of impact-weighted research output. Note that we looked at output (i.e., published papers in only two journals) without considering inputs (e.g., manpower, financial resources). Hence, we cannot say anything about a country's productivity. Because input data comparable across countries is not available, we follow Kocher and Sutter (2001) in using a country's population as proxy and compute each country's output per million inhabitants in *Table 17* and *Table 18* (appendix A)[16].

Considering the first time period (1999-2003, see *Table 17*), we see that output per million inhabitants was below 0.2, except for Switzerland (1.36), New Zealand (0.46), and the US (0.26). Considering only the last period (2014-2018, see appendix A, *Table 18*), we see that output was above 0.40 for several countries (Australia 0.90, Germany 0.65, UK 0.55, US 0.49) and that productivity was very high in some small countries (Austria 1.95, Singapore 1.40, the Netherlands 1.36, Norway 1.17, Switzerland 0.93, Denmark 0.83, Sweden 0.80, New Zealand 0.53).

Interestingly, there are some differences between our results and the results reported in Kocher and Sutter (2001, 414). According to their results, Israel has the highest and the US have the second-highest productivity. UK ranks third, Switzerland ranks 9th, Germany ranks 17th, and Spain ranks 19th. Recall that while Kocher and Sutter consider papers from all areas, we focus on laboratory experimental economics. In addition, Kocher and Sutter include data until 1997 while our data on geographical concentration starts from 1999. Because of the different time periods and our specialization on one particular field, the results are not directly comparable.

---

[15] The Herfindahl index is the sum of all countries' squared market shares. For a monopoly, the index is equal to one. Smaller numbers indicate less concentration.

[16] This approach only gives a crude picture and has several shortcomings (c.f. Kocher and Sutter 2001, 414). It implicitly assumes that resources devoted to experimental economics is highly correlated with a country's number of inhabitants which is problematic especially in large, less-developed countries. Data for population sizes was taken from the World Bank (see, https://databank.worldbank.org/data/indicator/SP.POP.TOTL/1ff4a498/Popular-Indicators, accessed May 22, 2019) and for Taiwan from the UN (see, https://population.un.org/wpp/Download/Standard/Population/, accessed May 22, 2019). To compute scores per million inhabitants for a specific time period we used the population size for the period's first year (e.g., 1999 population for the period 1999-2003).

Yet, a tentative explanation for the differences is that several European countries (Switzerland, Austria, Germany, Spain, the Netherlands) have specialized in experimental economics.

4.3 Gender Composition of Author Teams

The 596 papers in our data set were written by a total of 1498 authors. 282 (18.83%) of these authors are female. *Table 3* depicts the gender composition by the number of authors per paper. For all numbers of authors, the proportion of papers written solely by male authors is more than 50%.

number of authors	1	2	3	>3	all
papers	$N = 71$	$N = 246$	$N = 197$	$N = 82$	$N = 596$
solely female authored papers	$N = 21$ 29.58%	$N = 11$ 4.47%	$N = 2$ 1.02%	$N = 0$ 0.00%	$N = 34$ 5.70%
solely male authored papers	$N = 50$ 70.42%	$N = 151$ 61.38%	$N = 121$ 61.42%	$N = 41$ 50.00%	$N = 363$ 60.91%
mixed gender authored papers	-	$N = 84$ 34.15%	$N = 74$ 37.56%	$N = 41$ 50.00%	$N = 199$ 33.39%

Table 3: Descriptive statistics on the gender composition of author teams by number of authors per paper.

Table 4 shows descriptive statistics on the gender composition of authors for each of the three journals. Across all journals, the share of papers written solely by female authors is only 5.70%. With 33.39%, the share of papers with at least one female author is much higher. Across all papers and journals, the average of the relative share of female authors is 20.16%. From 1998 to 2018, no clear trend can be identified for the relative share of female authors (see appendix A, *Figure 3*). Our results are comparable to those of Hengel and Moon (2020) who observe only a marginal increase in the percentage share of female authors for papers published since 1990 (11%) and 2010 (14%). In our data set, the relative shares of female authors do not differ between the journals (two-sample Mann-Whitney tests, all p-values > 0.1).

	AER	Exp Econ	JEEA	all three
authors	$N = 370$	$N = 1013$	$N = 115$	$N = 1498$
female authors	$N = 60$ 16.22%	$N = 204$ 20.14%	$N = 18$ 15.65%	$N = 282$ 18.83%
male authors	$N = 310$ 83.78%	$N = 809$ 79.86%	$N = 97$ 84.35%	$N = 1216$ 81.17%
papers	$N = 145$	$N = 410$	$N = 41$	$N = 596$
solely female authored papers	$N = 6$ 4.14%	$N = 27$ 6.59%	$N = 1$ 2.44%	$N = 34$ 5.70%
solely male authored papers	$N = 95$ 65.52%	$N = 240$ 58.54%	$N = 28$ 68.29%	$N = 363$ 60.91%
mixed gender authored papers	$N = 44$ 30.34%	$N = 143$ 34.88%	$N = 12$ 29.27%	$N = 199$ 33.39%

Table 4: Descriptive statistics on the gender composition of author teams for each journal.

5 Quality Characteristics

In the preceding section, we described the development of geographical concentration by looking at the number of publications weighted by the journals' impact factor. Such an analysis neglects any differences between individual papers published within the same journal in a given year. If there are significant quality differences between different papers published in the same journal in a given year, the results from the preceding section could be biased. By focusing on the characteristics of individual experiments and papers, we try to uncover such differences in order to provide a more detailed picture.

In subsection 5.1, we describe four proxies for the *quality of laboratory experiments*. We compare these proxies in subsection 5.2. This comparison reveals that there are differences regarding methodological standards between experiments conducted in North America and Europe[17]. In subsection 5.3, we describe three proxies for the *quality of papers* that report experimental results. All seven proxies are available prior to publication and might influence editors' decisions. This is not the case for the number of citations a paper receives after publication, which is arguably the most important *ex post* quality proxy as discussed in subsection 5.4. Finally, we compare the number of citations between experiments conducted in North America and Europe in subsection 5.5.

5.1 Proxies for the Quality of Laboratory Experiments

Although our proxies for the quality of an experiment (shorthand: P1-P4) are objectively measurable, they are only proxies. Thus, it is possible that they only weakly correlate with the true quality of an experiment, which is of course unknown.

(P1 and P2) The first two quality proxies are the *total number of participants* and the *number of participants per treatment*. Both are related to the experiment's statistical power. For a more detailed discussion of power analysis in experimental economics see Czibor et al. (2019), Ioannidis et al. (2017), Bellemare et al. (2016, 2014), and Zhang and Ortmann (2013).

---

[17] Since there are only a few papers in our data set where the majority of authors or all authors are female, we do not examine gender differences for the quality proxies. However, we control for the effect of gender in our regression analyses in section 6.

Experiments with higher power tend to generate fewer false positives and, hence, results are more likely to be replicated. This in turn could increase confidence in the experiment's results[18].

For a given number of treatments, a larger *total number of participants* yields more observations which increases statistical power if statistical testing is carried out at the participant-level. A similar proxy for quality is the *number of participants per treatment*. If an experiment contains a large number of treatments, and if the statistical analysis involves a pairwise comparison of treatments, the *number of participants per treatment* is a more suitable proxy for the experiment's power.

Arguably, another possibility would be to consider the number of independent observations as a proxy for quality. However, the definition of what constitutes an independent observation is not clear-cut. Thus, we decided not to follow this approach (see chapter 3 in Svorenčík and Maas 2016). A further related aspect concerns the possibility to econometrically control for dependencies (see Moffatt 2015, 5-6). If dependencies are controlled for, the number of independent observations will underestimate the quality of the experiment[19].

(P3) The third proxy is the *number of treatments*. Ceteris paribus, a larger *number of treatments* increases quality because it allows to test for more alternative explanations. Testing for a higher number of alternative explanations, increases the scientific value of a paper. Linking the interpretation of results to a larger number of existing papers and ruling out more alternative hypothesis, yields more credibility to the authors claims.

(P4) The fourth proxy is the *strength of monetary incentives*. According to the methodological literature on laboratory experiments, monetary incentives are related to the quality of an experiment. In fact, the use of monetary incentives is an established methodological standard in economics (see Hertwig and Ortmann 2001, 390) and all experiments in our data set use monetary incentives. This is possibly the case because experimental economists expect that they cannot publish their paper if they would use hypothetical rather than monetary incentives

---

[18] Camerer at al. (2016) were able to successfully replicate the results of 11 out of 18 experiments. With the goal of making replication studies more attractive, Drazen et al. (2019) suggest a journal-based replication in which the journal contracts a replication after having accepted but before publishing a paper.

[19] For example, consider a T-times repeated public good game with stranger matching. Assume the experiment consists of two treatments, S sessions per treatment, and N participants per treatment. Due to the stranger matching protocol, all observations within the same sessions are dependent. The most conservative way would be to take a session's average contribution (pooled across participants and time) as one independent observation and compare session averages across treatments. Proceeding this way, would yield S independent observations per treatment. Exploiting the panel structure of the data and applying a random effects model, we could use each single decision as an observation resulting in a total of $2STN$ observations.

(Camerer and Hogarth 1999, 31)[20]. We operationalized the *strength of monetary incentives* by dividing a participant's average earnings (including the show-up fee) by the duration of the experiment (in minutes) and converting the result into real 2015 US-Dollars[21].

Regarding the *strength of monetary incentives*, Davis and Holt (1993, 24-25) argue that incentives should be high enough to cover opportunity costs. While this rule defines a lower limit for incentives, it does not postulate a clear relation between the strength of incentives and behavior. Camerer and Hogarth (1999, 31) show that the *strength of monetary incentives* affects behavior and increases data quality. With stronger incentives, subjects are less likely to show thoughtless behavior or make errors, leading to lower variance and less noise in the data. This, in turn, increases (statistical) power, allows for more precise statistical testing, and increases the likelihood of successful replication. According to this line of reasoning, experimental economists could increase the quality of their experiments by paying higher rewards[22].

Table 5 gives an overview of mean values for the quality proxies for the three journals. At first glance, we see that *total number of participants*, *number of participants per treatment,* and *number of treatments* are higher for the AER and the JEEA than for Exp Econ. As *Table 5* illustrates, the variable *strength of monetary incentives* could only be computed for 54% of AER papers, 72% of Exp Econ papers and 66% of JEEA papers. This is because, even in these prestigious journals, many papers do not report the number of participants, average earnings per participant, or the duration of the experiment.

---

[20] For a more detailed discussion of monetary incentives see Bardsley et al. (chapter 6, 2010), Ortmann (2009), Read (2005), Guala (chapter 11, 2005), and Rydval and Ortmann (2004).

[21] Data for exchange rates is taken from fxtop.com. Data for CPI is taken from https://data.oecd.org. In some papers it was not stated in which year the experiment was conducted. In these cases we estimated the corresponding value by subtracting three years from the year of publication.

[22] A related aspect that is discussed by Camerer and Hogarth (1999) concerns participants' cognitive abilities. Together with monetary incentives and intrinsic motivation, cognitive abilities affect participants' mental effort. However, the effect from incentives on mental effort is likely to be non-monotonic, as recognized by Davis and Holt: "No amount of money can motivate participants to perform a calculation beyond their intellectual capacities, any more than generous bonuses would transform most of us into professional athletes" (Davis and Holt 1993, 24). Because no data on participants' cognitive abilities are available, we could not include cognitive abilities in our set of quality proxies. We believe that this is unproblematic because there is no reason to expect any systematic differences in participants' cognitive abilities across experiments.

	AER $N = 145$	Exp Econ $N = 410$	JEEA $N = 41$
P1: total # of participants	253.93 (198.32) $N = 137$	194.49 (128.45) $N = 407$	284.58 (177.61) $N = 40$
P2: # participants per treatment	68.76 (87.59) $N = 135$	55.33 (40.97) $N = 405$	61.82 (41.81) $N = 34$
P3: # treatments	4.48 (3.00) $N = 143$	3.90 (2.29) $N = 407$	4.74 (2.43) $N = 35$
P4: strength of monetary incentives	0.42 (0.25) $N = 79$	0.31 (0.14) $N = 294$	0.35 (0.16) $N = 27$

Table 5: Mean values, standard deviations (in parentheses), and number of observations for experiments' proxies.

5.2   Are there Differences in Methodological Standards?

In this subsection, we examine whether there are clearly defined methodological standards with regard to the *ex ante* proxies for quality as described above. If this were the case, one would expect to find no differences regarding these proxies between experiments conducted in different regions.

In the following, we look only at laboratory experiments conducted in North America or Europe. More than 83% of the laboratory experiments in the papers within our dataset were conducted in these two regions.

P1: total # participants	North America	Europe	both regions	*p*-value
AER	220.33 (165.06) $N = 76$	361.79 (254.22) $N = 39$	268.30 (209.74) $N = 115$	0.0004
Exp Econ	168.56 (95.55) $N = 166$	204.93 (126.44) $N = 169$	186.91 (113.50) $N = 335$	0.0114
JEEA	243.57 (117.08) $N = 14$	295.83 (212.04) $N = 23$	276.05 (181.90) $N = 37$	0.9568
all three	188.03 (123.88) $N = 256$	240.47 (174.32) $N = 231$	212.90 (152.05) $N = 487$	0.0005

Table 6: Mean values, standard deviations (in parentheses), and number of observations for total number of participants by region and journal. The last column shows the exact p-value from a two-sided Mann-Whitney test.

P2: # participants per treatment	North America	Europe	both regions	p-value
AER	55.00 (44.95) N = 75	109.12 (143.31) N = 39	73.51 (94.31) N = 114	0.0001
Exp Econ	49.20 (33.94) N = 165	59.97 (40.76) N = 168	54.63 (37.87) N = 333	0.0033
JEEA	37.90 (18.25) N = 11	73.53 (47.16) N = 21	61.29 (42.88) N = 32	0.0122
all three	50.43 (37.17) N = 251	69.63 (72.08) N = 228	59.57 (57.29) N = 479	<0.0001

Table 7: Mean values, standard deviations (in parentheses), and number of observations for participants per treatment by region and journal. The last column shows the exact p-value from a two-sided Mann-Whitney test.

P3: # treatments	North America	Europe	both regions	p-value
AER	4.44 (2.98) 3 N = 81	4.78 (3.39) 4 & 6 N = 40	4.55 (3.11) 3 N = 121	0.7009
Exp Econ	3.94 (2.30) 4 N = 167	3.79 (2.34) 4 N = 168	3.87 (2.32) 4 N = 335	0.3810
JEEA	5.00 (3.02) 2 N = 12	4.48 (2.09) 3 N = 21	4.67 (2.43) 3 N = 33	0.8438
all three	4.15 (2.57) 4 N = 260	4.03 (2.55) 4 N = 229	4.09 (2.56) 4 N = 489	0.4945

Table 8: Mean values, standard deviations (in parentheses), modes, and number of observations for number of treatments by region and journal. The last column shows the exact p-value from a two-sided Mann-Whitney test.

P4: strength of monetary incentives	North America	Europe	both regions	p-value
AER	0.39 (0.17) N = 41	0.48 (0.34) N = 30	0.43 (0.26) N = 71	0.1264
Exp Econ	0.32 (0.14) N = 118	0.35 (0.17) N = 131	0.34 (0.15) N = 249	0.0425
JEEA	0.38 (0.16) N = 7	0.33 (0.17) N = 18	0.35 (0.17) N = 25	0.7451
all three	0.34 (0.15) N = 166	0.37 (0.21) N = 179	0.36 (0.18) N = 345	0.0491

Table 9: Mean values, standard deviations (in parentheses), and number of observations for strength of monetary incentives by region and journal. The last column shows the exact p-value from a two-sided Mann-Whitney test.

Table 6 to *Table 9* show the mean values of the *ex ante* proxies for quality by journal and region. With respect to the *total number of participants* (P1) or the *number of participants per treatment* (P2), there seems to be no universal methodological standard. Rather, there appear to be systematic differences between experiments conducted in both regions.

Except for the JEEA, experiments conducted in Europe involve a significantly larger *total number of participants*. For all journals, European experiments have a significantly larger *number of participants per treatment*. These discrepancies possibly emerge from differences with regard to (the definition of) what constitutes an independent observation, as emphasized by Frans van Winden:

> "I can remember at the Amsterdam meetings that there were some heated discussions, and they were related to three topics. First of all, what is an independent observation? What I remember is that people from the United States were more liberal in the sense that they applied parametric statistics whereas the Germans, especially of course Reinhard Selten and the group he influenced, were stricter on that."
> (Frans van Winden, cited in Svorenčík and Maas 2016, 188-198).

Another possible explanation for the differences in the *total number of participants* and the *number of participants per treatment* (P1 and P2) is that there are disparities in the types of experiments conducted in the two regions (e.g., individual decision making, public good games, market experiments...) that require different numbers of participants. If there are systematic differences in the distribution of types, our measures P1 and P2 will be biased. A cursory look into our data reveals that for almost all experiment types, the number of participants is higher for experiments conducted in Europe, indicating that it is unlikely that our results are driven by differences in types. However, while some experiments can be easily classified as certain types, there are many experiments that are hard to classify.

With regard to the *number of treatments*, there are no significant differences between North-American and European experiments (see *Table 8*). Across journals and continents, the mode of treatments per experiment is four.

The bottom row of *Table 9* indicates that the *strength of monetary incentives* is significantly higher for experiments conducted in Europe. However, the difference is small and, at the journal level, we only observe a small significant difference for Exp Econ. We therefore conclude that

this observation is compatible with the existence of a common standard for how to incentivize participants in Europe and North America.

Our results for the *total number of participants* and the *strength of monetary incentives* could suggest that experimental economists from Europe have to pay a higher price for publishing their papers, compared to experimental economists from North America. Another explanation might be that economists in Europe have easier access to research funding compared to their colleagues in the US[23].

The differences in the *total number of participants* (P1) could be driven by a cohort effect. Section 4.2 showed that the US's research output was very high in the first two periods before it then decreased. If P1 increases over time, and if this increase has the same magnitude for experiments conducted in North America and Europe, the differences in P1 in *Table 6* would only reflect the fact that the share of experiments conducted in the US decreased. To rule out such cohort effects, we look at P1 for four periods. *Table 19* (appendix B) shows that there are almost no significant differences in P1 for the first and second period. However, from the periods 2009-2013 and 2014-2018, P1 for European experiments for both journals is always higher and in most cases, the difference is statistically significant. Moreover, the difference between North America and Europe increases over time. This suggests that the differences in P1 for the pooled data (1999-2018) are not driven by the decreasing share of US output. Rather, they are driven by differences in P1 in the two most recent periods.

5.3   *Ex Ante* Proxies for a Paper's Quality

Ex ante proxies represent information that is available before the paper is accepted for publication. Thus, such proxies can be used by referees and editors when deciding about whether to accept or reject a submitted paper. Our four proxies, P1 to P4, represent such *ex ante* proxies. However, aspects such as novelty, originality, and practical relevance are also available *ex ante* and are important for a paper's quality. Attempting to account for such additional *ex ante* proxies, we include four additional variables: the number of pages, the number of

---

[23] We thank Gary Charness for drawing our attention to this aspect. Unfortunately, there are no sources for the amount of research funding for experimental economists in different regions. However, we consider it unlikely that there is a causal relationship between the availability of research funding and the average *total number of participants*. Even if research funds were easier available in Europe, they could be used, for example, to conduct more (single) projects.

references, authors' reputation, and the number of authors. When testing for editorial favoritism (in section 6), we will control for these variables.

(P5) In line with the results of Card and DellaVigna (2014, 2013), Medoff (2003), and Laband and Piette (1994a), we expect a positive correlation between the *number of pages* and quality. We assume that editors are willing to allocate more journal space to high quality papers. Since the journals have different formats and therefore different numbers of words per page, whenever necessary, we converted the number of pages into Exp Econ-equivalent pages.

(P6) We expect a positive correlation between the *number of references* and quality. We assume that authors who cite a larger number of references have studied a larger amount of related literature before conducting their own research. As we explain in subsection 5.4 below, we use citations received in the 5 and 10 years after the publication of a paper as a proxy for the *ex post* quality. It seems plausible that a paper with a higher *number of references* also receives more attention from the scientific community than a paper with only a smaller *number of references*.

(P7) As shown in Card et al. (2020), Hengel and Moon (2020), Brogaard et al. (2014), Medoff (2003), and Laband and Piette (1994a), an *author's reputation* is likely to have a positive effect on the paper's quality. We proxy authors' reputation by their stock of citations. In case of a single author, we take the total number of citations received by the author during the 5 years prior to the publication of the paper. In case of more than one author, we take the average of all authors' citations during the 5 years prior to the publication of the paper[24]. Editors could view a scholar's stock of prior citations as a signal of the expected scientific contribution of her paper (see Medoff 2003, 428-429)[25]. It is also conceivable that, regardless of the paper's quality, papers authored by well-known economists tend to attract more citations (Merton's Matthew effect, see Merton 1968).

(P8) In line with the related literature (Card et al. 2020, Hengel and Moon 2020, Brogaard et al. 2014), we control for the *number of authors* for each paper. While Hengel and Moon (2020) find a positive effect for the number of authors on the number of citations, Brogaard et al.

---

[24] The average number of authors increased significantly from the first period (1999-2003, mean: 2.31) to each of the three subsequent periods (2004-2008, mean: 2.54, $p = 0.032$; 2009-2013, mean: 2.65, $p < 0.01$; 2014-2018, mean: 2.53, $p = 0.025$; p-values from two-sided Mann-Whitney tests). However, the increase is small and there are no significant differences in the average number of authors in the last three periods. We therefore assume that our measure of reputation is not distorted by changes in the average number of authors.

[25] This requires that editors can identify authors (identities) despite the referee process being double blind. In a field experiment, Blank (1991) found that in AER's double blind referee process, about 50% of referees could correctly identify the authors. Given that many authors post preliminary versions of their papers online, we assume that, in many cases, editors can also identify the authors.

(2014) find negative or null effects, depending on the statistical models[26]. Therefore, we do not expect to find a clear effect for this proxy.

5.4 Citations as Proxy for *Ex Post* Quality

Ex post proxies measure the quality of a paper after its publication. We focus on only one proxy, namely citations[27]. This does not imply that citations reflect a paper's true quality. Rather, we decided to use the *number of citations* because data on citations is widely available and is heavily used (at least as a yardstick) to allocate positions and resources (Card et al. 2020, Card and DellaVigna 2020, Hamermesh 2018, Moed 2006, Laband and Piette 1994b). Moreover, assuming that editors seek to maximize their journals' impact factor, they have an incentive to accept papers that they expect to receive a high number of citations.

Figure 1 and *Figure 2* show the distributions of citations received in the 5 and 10 years following the publication by journal. Due to the small number of observations for the JEEA, we plot the distributions only for the AER and Exp Econ (but present some descriptive statistics in *Table 10*). The distributions are heavily skewed to the right. As one would expect, laboratory experiments published in the AER receive significantly more citations than laboratory experiments published in Exp Econ.
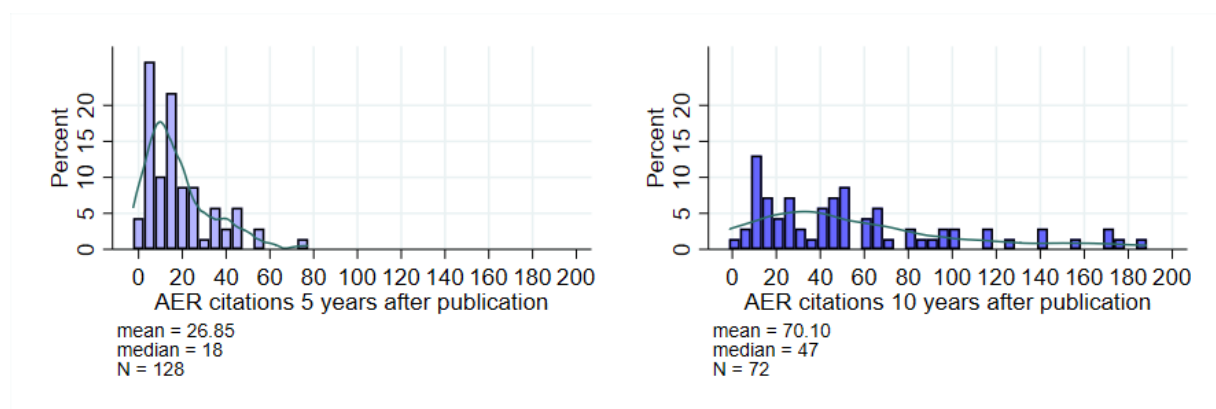
Figure 1: Distribution of citations for laboratory experiments published in the AER, 5 and 10 years after publication.

---

[26] Brogaard et al. (2014) note that this result is not very meaningful, as their measure of the authors' reputation is strongly correlated with the number of authors.

[27] Other *ex post* proxies are being reprinted in anthologies or the amount of media-coverage.
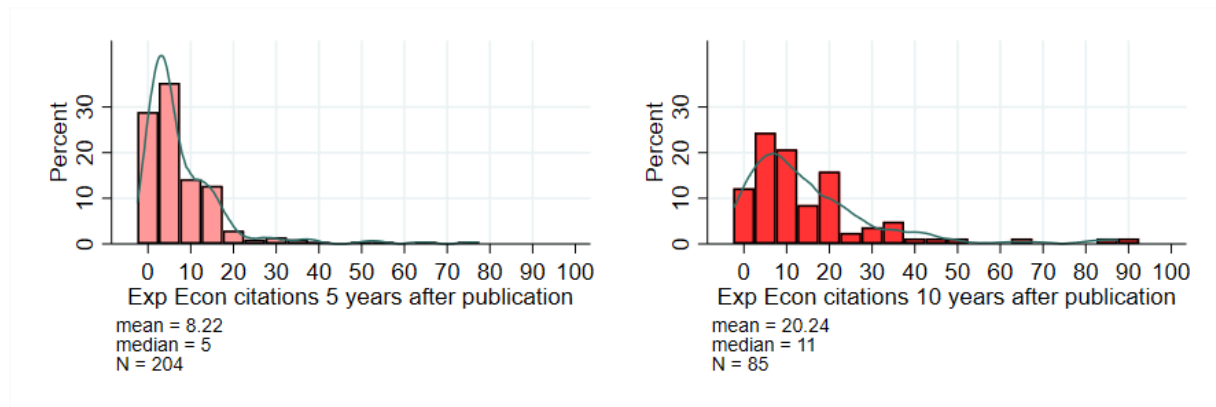
Figure 2: Distribution of citations for laboratory experiments published in Exp Econ, 5 and 10 years after publication.

5.5 Are there Differences in Citations?

Subsection 5.2 showed that North American and European experiments differ with respect to the *total number of participants* (P1) and the *number of participants per treatment* (P2). This raises the question whether this difference is associated with a difference in *ex post* quality, i.e., the *number of citations*. If there is a positive correlation between the *total number of participants* or the *number of participants per treatment* and the experiment's quality, one would expect that European experiments receive a higher *number of citations* on average.

Table 10 and *Table 11* show that laboratory experiments conducted in Europe receive more citations compared to laboratory experiments conducted in North America. For the 5-year period after publication, this holds for all three journals, while for the 10-year period after publication this only applies to papers published in the AER and Exp Econ. For the JEEA, there are too few observations because the JEEA only launched in 2003.

Most differences for mean values are not only statistically significant but also, more importantly, large in magnitude. For example, looking at citations in the 10 years after publication for experiments published in the AER, we see that European experiments received, on average, 124 citations while North American experiments only received an average of 61 citations. Since citations are heavily skewed to the right we also perform a test for differences in medians in *Table 10* and *Table 11*. There are only significant differences in medians for citations received in the 5 years after publication for laboratory experiments published in the AER and Exp Econ.

26

citations five years after publication		North America	Europe	both regions	p-value
AER	mean	23.32	40.29	28.81	0.0021
	median	17.00	32.00	20.00	0.029
	90%-p.	58.00	97.00	65.00	
	sd	(22.22)	(35.69)	(28.30)	
	N	73	35	108	
Exp Econ	mean	6.20	10.62	8.50	0.0037
	median	4.00	6.00	4.00	0.007
	90%-p.	15.50	23.00	17.00	
	sd	(7.87)	(13.03)	(11.06)	
	N	80	87	167	
JEEA	mean	20.89	27.20	24.21	0.6748
	median	12.00	10.00	12.00	0.788
	90%-p.	83.00	97.00	83.00	
	sd	(24.61)	(47.58)	(37.57)	
	N	9	10	19	
all three	mean	14.73	19.74	16.98	0.1091
	median	7.00	11.00	9.00	0.140
	90%-p.	36.00	46.00	42.00	
	sd	(18.79)	(27.75)	(23.33)	
	N	162	132	294	

Table 10: Mean values, medians, 90% percentiles, standard deviations (in parentheses), and number of observations for citations in the 5 years after publication ('c5') by region and journal. The upper (lower) row of the last column contains the p-values of a two-sample Mann-Whitney test (continuity corrected Pearson Chi-square test), comparing the distribution of c5 between North America and Europe.

citations ten years after publication		North America	Europe	both regions	p-value
AER	mean	61.34	123.62	75.74	0.0178
	median	46.50	101.00	40.00	0.482
	90%-p.	128.00	275.00	171.00	
	sd	(74.56)	(106.33)	(85.94)	
	N	44	13	57	
Exp Econ	mean	13.55	33.57	20.63	0.0511
	median	9.00	14.00	11.00	0.541
	90%-p.	27.00	88.00	43.00	
	sd	(15.07)	(45.69)	(30.92)	
	N	42	23	65	
JEEA	mean	8.00	31.50	23.67	*not enough*
	median	8.00	31.50	17.00	*observ.*
	90%-p.	8.00	46.00	46.00	
	sd	-	(20.51)	(19.86)	
	N	1	2	3	
both	mean	37.66	64.26	45.74	0.0762
	median	19.00	35.00	21.00	0.318
	90%-p.	86.00	176.00	118.00	
	sd	(58.90)	(82.47)	(67.72)	
	N	87	38	125	

Table 11: Mean values, medians, 90% percentiles, standard deviations (in parentheses), and number of observations for citations in the 10 years after publication ('c10') by region and journal. The upper (lower) row of the last column contains the p-values of a two-sample Mann-Whitney test (continuity corrected Pearson Chi-square test), comparing the distribution of c10 between North America and Europe.

6 Citations, Quality Proxies, and Editorial Favoritism

In this section, we examine whether the differences in citations are driven by differences in the quality proxies P1 to P8 or if they are indicative of editorial favoritism.

6.1 Editorial Favoritism

With regard to the quality proxies discussed in subsection 5.1, there is no reason to assume that clearly defined standards exist. In absence of clearly defined standards, editors have a lot of leeway. Moreover, the proxies do not capture a paper's originality and relevance. Originality and relevance are highly subjective and might be unknown to editors when deciding about publication. This makes it possible that non-scientific criteria can have an influence on editors' decisions, thus facilitating discrimination. It is even possible that the US dominance results from the behavior of editors, whose decisions are influenced by non-scientific criteria[28].

In order to assess whether editorial favoritism is present in experimental economics, we use an analysis similar to Medoff (2003) and Laband and Piette (1994a). Unlike Medoff (2003) and Laband and Piette (1994a), we do not focus on a single year of publication, but a period of several years. In the literature (Colussi 2018, Brogaard et al. 2014, Medoff 2003, Laband and Piette 1994a), there are multiple methods for measuring social ties. For our analysis we do not consider social ties between authors and editorial board members at the institutional level. Instead, we specify our measure of social ties as the share of authors with a US affiliation[29]. More specifically, we test the hypothesis that for AER papers, the share of US-affiliated authors is negatively correlated with the number of citations.

If confirmed, this would indicate that there is editorial favoritism in the AER. We hypothesize editorial favoritism for the AER, because almost all members of the AER's editorial board are affiliated with a North American (mostly US) institution[30]. We do not hypothesize editorial favoritism for Exp Econ and the JEEA, because only half of their editorial boards consists of North American-affiliated economists[31].

---

[28] Other reasons for the US dominance are mentioned in Frey and Pommerehne (1998). Firstly, the social and political conditions for research are very good in the US, leading to immigration of skilled researchers into the US. Secondly, the incentives for good research (in terms of publications in top journals) are very strong in the US. And thirdly, language: papers published in a language other than English will rarely receive the same amount of attention as papers published in English because English is the lingua franca of economics. For economists whose native language is not English, publishing in English is associated with higher costs because one has to adopt the language, style, and format common to the US. We believe that these reasons have weakend during the last years.

[29] Laband and Piette (1994a: 201) note that a broad range of social ties between editors and authors are not included in their analysis. Due to this fact and due to our long observation period, we are confident that our measure of social ties is suitable to show possible effects of editorial favoritism.

[30] AER's editorial board consists of 77 individuals, serving as editor, coeditor, or board member. 72 individuals (93.50%) are affiliated with a North American-institution. Out of the five remaining individuals, three received their PhDs from US universities (https://www.aeaweb.org/journals/aer/about-aer/editors, accessed May 22, 2019).

[31] Exp Econ's editorial board consists of 59 individuals, serving as editor, advisory editor, or member of the editorial board. 28 individuals (47.46%) are affiliated with North American-institutions. The rest of the editorial board includes 23 individuals (38.98%) with a European affiliation, 6 individuals with an affiliation in Australia

In order to test our hypothesis, we apply the following regression models.

Models 1 and 4:

$$c_i(5,10) = \beta_0 + \beta_1 ST_i + \beta_2 Fem_i + \gamma_5 P5_i + \gamma_6 P6_i + \gamma_7 P7_i + \gamma_8 P8_i + \delta_i P\&P_i + u_i \quad (1)$$

Models 2 and 5:

$$c_i(5,10) = \beta_0 + \beta_1 ST_i + \beta_2 Fem_i + \gamma_1 P1_i + \gamma_4 P4_i + \gamma_5 P5_i + \gamma_6 P6_i + \gamma_7 P7_i +$$
$$\gamma_8 P8_i + \delta_i P\&P_i + u_i \quad (2)$$

Models 3 and 6:

$$c_i(5,10) = \beta_0 + \beta_1 ST_i + \beta_2 Fem_i + \gamma_2 P2_i + \gamma_3 P3_i + \gamma_4 P4_i + \gamma_5 P5_i + \gamma_6 P6_i + \gamma_7 P7_i +$$
$$\gamma_8 P8_i + \delta_i P\&P_i + u_i \quad (3)$$

The dependent variables $c_i(5)$ and $c_i(10)$ represent the total number of citations paper i received during the 5 or 10 years after publication. Because the number of citations is left-censored by 0, we used Tobit regressions. For our robustness checks (appendix B) we followed the literature (Hengel and Moon 2020, Card et al. 2020) and used the inverse hyperbolic sine (asinh) of citations and the log of "1 + citations" as dependent variables. Note that data on $c_i(5)$ ($c_i(10)$) is available only for papers published in 2013 (2008) or earlier.

Our measure of social ties, ST_i, is the relative share of US-affiliated authors for paper i[32]. The coefficient β_1 estimates the difference in quality (measured by the number of citations) between AER papers authored by scholars with and without a US-affiliation. In the presence of editorial favoritism, we expect $\beta_1 < 0$.

---

or New Zealand, and one individual each with an affiliation in Asia or South America (https://www.springer.com/economics/economic+theory/journal/10683?detailsPage=editorialBoard, accessed May 22, 2019). The JEEA's editorial board consists of 72 individuals, serving as editor, associate editor, or member of the advisory board. 37 individuals are affiliated with a European institution, 34 individuals are affiliated with a North-American institution, and one individual is affiliated with an institution in Asia (https://www.eeassoc.org/index.php?site=JEEA&page=175&trsz=45, accessed October 1, 2019).

[32] This definition differs from e.g., Laband and Piette (1994a) and Medoff (2003) who focus on authors who received their PhDs from the same university the editor was affiliated with at the time of the publication. We focus on US versus non-US-affiliations because the AER has several rules that prevent editorial favoritism on the level of institutions. For example, "[co-]editors are recused from papers involving current colleagues at the same institution (regardless of department), as well as graduate students at the same institution" (for all rules see: https://www.aeaweb.org/journals/aer/about-aer/editorial-policy). We excluded papers ($N = 7$) reporting laboratory experiments conducted in the US by authors that do not have an affiliation in the US, because for these papers it is unclear whether a social tie exists.

The variable Fem_i measures the relative share of female authors for paper i. If higher quality requirements for female authors lead to a substantially higher quality of the respective papers, we expect $\beta_2 > 0$.

The paper's quality proxies, *number of pages* (P5), *number of references* (P6), *authors' reputation* (P7), and *number of authors* (P8) are included in each model specification. Models 1 and 4 only include these proxies. In models 2 and 5, the *total number of participants* (P1) and the *strength of incentives* (P4) are added. In models 3 and 6, we further include the *number of participants per treatment* (P2), the *number of treatments* (P3), and the *strength of incentives* (P4). Based on sections 5.1 and 5.3, except for the *number of authors* (P8), we expect all proxies' coefficients to be positive. For the AER regressions, $P\&P_i$ is an indicator variable for papers published in AER's Papers and Proceedings, which we include in models 1 to 6.

	(1) c5	(2) c5	(3) c5	(4) c10	(5) c10	(6) c10
Constant	24.61** (10.77)	40.23** (19.86)	43.68** (21.57)	162.39*** (46.70)	283.48*** (84.80)	350.73*** (85.21)
ST	-11.82** (5.23)	-24.90*** (8.78)	-26.02*** (8.99)	-53.62** (25.67)	-118.71*** (38.98)	-152.56*** (36.19)
Fem	13.82* (8.09)	20.92 (13.03)	21.21 (13.31)	29.90 (36.89)	14.30 (81.78)	-34.58 (82.83)
Experiment Quality						
P1 (total nr. of part.)		-0.02 (0.02)			-0.06 (0.08)	
P2 (part. per treatm.)			-0.04 (0.03)			-0.20** (0.09)
P3 (treatments)			-1.10 (1.39)			-2.24 (7.86)
P4 (incentives)		-3.00 (13.19)	-2.83 (13.54)		-6.76 (42.68)	-18.62 (41.18)
Paper Quality						
P5 (pages)	-0.30 (0.27)	-0.08 (0.44)	0.01 (0.45)	-2.04 (1.29)	-1.79 (2.48)	-2.46 (2.96)
P6 (references)	0.63*** (0.16)	0.53** (0.24)	0.51** (0.24)	0.58 (0.80)	-1.05 (1.53)	-1.94 (1.62)
P7 (reputation)	0.02*** (0.01)	0.02* (0.01)	0.02 (0.01)	0.06 (0.09)	0.31 (0.29)	0.42 (0.30)
P8 (authors)	-3.06 (2.51)	-5.49 (3.79)	-5.83 (3.94)	-10.13 (10.61)	-25.24 (19.52)	-22.86 (15.51)
Papers&Proceed.	-11.10 (7.92)	-0.16 (16.42)	5.25 (18.66)	-70.00** (28.93)	-36.80 (57.68)	-68.85 (61.55)
N	127	68	66	71	26	25
Pseudo R^2	0.03	0.03	0.04	0.01	0.04	0.06

Table 12: Tobit regressions of c(5) and c(10) on social ties (ST), rel. share of female authors (Fem) and papers' and experiments' quality proxies. AER papers. Standard errors in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 12 shows the regression results for papers published in the AER. Taking the *number of citations* in the 5 years after publication as a proxy for quality, we find a significant negative effect of social ties for models 1 to 3 ($\beta_1 < 0$). This suggests that papers from authors with stronger social ties (i.e., a larger share of US-affiliated authors) have a lower quality compared to papers from authors with affiliations outside the US. More precisely, β_1 measures the difference in citations between a paper with solely US-affiliated authors and a paper without any US-affiliated authors. Based on models 1 to 3, papers with solely US-affiliated authors receive between 12 and 26 citations less in the 5 years after publication. The results are even stronger when we take the *number of citations* in the 10 years after publication as a proxy for

quality (models 4 to 6). On average, papers with solely US-affiliated authors receive between 54 and 153 citations less compared to papers without any US-affiliated authors. For alternative specifications of our dependent variables (appendix B, *Table 20* and *Table 21*) we find qualitatively similar results[33].

This difference can be interpreted as indirect evidence for discrimination, indicating that a paper from a US-based author is accepted for publication despite a higher-quality paper from an author with an affiliation outside the US being available. Of course, this only holds true if high-quality papers from authors without social ties to the editorial board were submitted and rejected. However, given the AER's rejection rate, we think that this is highly plausible[34].

This, however, does not imply that editors consciously discriminate against authors from outside the US. Rather, it is possible that editors systematically fail to predict the number of citations a paper will receive in the years following its publication (i.e., editors wrongly predict that papers authored by US-based economists receive more citations than they actually do). Nonetheless, if citations can be taken as a proxy for quality, this will slow down scientific progress, and editors will fail to maximize their journal's impact factor.

6.2   How do the Quality Proxies of AER Papers Affect Citations?

Regarding the proxies for experimental quality (P1 to P4), only the *number of participants per treatment* (P2) affects citations. For each model, the *total number of participants* (P1), the *number of treatments* (P3), and the *strength of monetary incentives* (P4) have no effect on the *number of citations* in the 5 or 10 years after publication. This does not imply that editors do not care about the experiments' quality proxies. Rather, it could be that there are minimum standards that have to be fulfilled in order to get a paper published, but going beyond those standards does not increase experimental quality any further. In other words, the perceived

---

[33] As further robustness checks, we test how our results from subsections 6.1 to 6.3 change if we use another proxy for social ties (see appendix B, *Table 22* to *Table 27*). Instead of ST_i we introduce an indicator variable that takes a value of 1 if all experiments included in a paper were carried out exclusively in North America (US and Canada). For 4 to 5 models in *Table 22* to *Table 24* (AER) and for 4 to 6 models in *Table 25* to *Table 27* (Exp Econ) we see that experiments conducted in North America receive significantly fewer citations in the 5 and 10 years after publication. The results for the quality characteristics (P1 to P8) are qualitatively similar to the results in section 6.

[34] In 2018, for example, only 5.85% of all submissions were accepted at the AER (Duflo 2019). Medoff (2003, 427) discusses several reasons why it is problematic to infer a possible editorial favoritism from the acceptance rates of a journal.

quality of an experiment that has a sufficient number of treatments and sufficiently strong incentives might not increase by adding additional treatments or by paying higher incentives.

In model 6 in *Table 12* (see also appendix B, *Table 20* and *Table 21*), the *number of participants per treatment* (P2) is statistically significant but negative – contrary to what we expected. Why should a paper's quality decrease when the *number of participants per treatment* increases? Finding a smaller effect requires an experiment with a larger *number of participants per treatment*. Thus, if there is an inverse relation between quality and effect size, possibly because smaller effects are perceived as less interesting, this could explain the negative effect.

Next, we turn to the paper's quality proxies (P5 to P8). The *number of pages* (P5) and the *number of authors* (P8) have no effect on citations. Taking *the number of citations* in the 5 years after publication as a proxy for quality (models 1 to 3), the *number of references* (P6) has a significant effect. Approximately 1.5 to 2 additional references yield one additional citation. The *authors' reputation* (P7) has a small effect for models 1 and 2. For 50 citations a single author (or an average co-author) has accumulated in the 5 years prior publication, the expected number of citations increases by 1[35]. However, if we take citations in the 10 years after publication as dependent variable, the *number of references* and the *authors' reputation* have no effect.

It seems that if the paper's quality proxies have any effects on citations, these effects are positive for citations during the 5 years after publication, but disappear for citations during the 10 years after publication. Perhaps the *number of references* (P6) and the *authors' reputation* (P7) attract some attention to the paper, leading to more citations, but this effect vanishes as the paper grows older.

6.3 How do the Quality Proxies of Exp Econ and JEEA Papers Affect Citations?

In *Table 13* and *Table 14*, we report the main results for papers published in Exp Econ and the JEEA respectively.

---

[35] Note that our proxy for reputation (the number of citations an author has received in the 5 years before the paper's publication) is right-skewed (see appendix B, *Figure 4* and *Figure 5*). For AER-papers from North America (Europe), the median is 74 (88) and the mean is 147 (247). This reflects Merton's Matthew effect. Those who already have large number of citations receive more citations.

	(1) c5	(2) c5	(3) c5	(4) c10	(5) c10	(6) c10
Constant	9.36***	6.50*	6.19	34.31**	5.28	-5.50
	(3.05)	(3.87)	(4.17)	(13.96)	(18.46)	(18.73)
ST	-2.45	-1.82	-1.78	-10.64	-3.75	-10.04
	(1.59)	(1.75)	(1.74)	(7.33)	(9.77)	(9.40)
Fem	-1.38	-0.79	-1.79	-2.63	9.03	-1.18
	(2.64)	(2.87)	(2.93)	(12.05)	(15.51)	(14.83)
Experiment Quality						
P1 (total nr. of part.)		0.01			0.08*	
		(0.01)			(0.05)	
P2 (part. per treatm.)			0.00			0.07
			(0.02)			(0.15)
P3 (treatments)			0.62			7.28***
			(0.38)			(2.26)
P4 (incentives)		3.91	2.63		14.34	20.19
		(4.19)	(4.23)		(22.63)	(21.26)
Paper Quality						
P5 (pages)	-0.43***	-0.43***	-0.45***	-1.19*	-1.38*	-1.07
	(0.13)	(0.13)	(0.13)	(0.61)	(0.72)	(0.68)
P6 (references)	0.28***	0.21***	0.22***	0.59**	0.53	0.49
	(0.06)	(0.07)	(0.07)	(0.29)	(0.36)	(0.34)
P7 (reputation)	0.02***	0.03***	0.03***	0.03	0.05	0.03
	(0.01)	(0.01)	(0.01)	(0.03)	(0.07)	(0.07)
P8 (authors)	-0.55	-0.01	-0.25	-1.57	3.68	1.74
	(0.79)	(0.85)	(0.84)	(3.40)	(4.40)	(4.12)
N	201	138	138	84	50	50
Pseudo R^2	0.02	0.03	0.03	0.01	0.02	0.04

Table 13: Tobit regressions of c(5) and c(10) on social ties (ST), rel. share of female authors (Fem), and experiments' and papers' quality proxies. Exp Econ papers. Standard errors in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

As can be seen from *Table 13*, for Exp Econ papers the relative share of US-based authors (*ST*) has no significant effect on citations. The *number of treatments* (P3) has a positive effect on citations in the 10 years after publication. For each additional treatment, the expected *number of citations* increases by 7.28. The *total number of participants* (P1) has no effect on citations in the 5 years, and a small positive effect on citations in the 10 years after publication. The *number of participants per treatment* (P2) and the *strength of monetary incentives* (P4) are not significant.

The *number of references* (P6) has a small but positive effect on citations in models 1 to 4. The *authors' reputation* (P7) has a small positive effect, but only on citations in the 5 years after publication. Finally, papers with a higher *number of pages* (P5) tend to receive fewer citations.

The non-significant effects of P1 to P4 possibly stem from the fact that the distribution of citations in the 5 years after publication is much narrower for Exp Econ, compared to the AER and the JEEA (see *Table 10*). For alternative specifications of the dependent variables (appendix B, *Table 28* and *Table 29*) we find qualitatively similar results. Only in model 1, the relative share of US-based authors (*ST*) takes a significantly negative value. In model 6, the effect on the *number of treatments* (P3) is no longer significant.

Regarding the *author's reputation* (P7) for Exp Econ, our results are in line with the findings of previous literature (Medoff 2003, Laband and Piette 1994a)[36]. However, our results differ regarding the *number of pages* (P5) since Card and DellaVigna (2014, 2013), Medoff (2003), and Laband and Piette (1994a) find that longer papers are associated with a higher number of citations.

---

[36] Card et al. (2020) and Hengel and Moon (2020) use alternative measures for authors' reputation (previous top publications) and are therefore only partially comparable with our results. Nevertheless, they also find that the reputation of authors has a positive effect on the number of citations.

	(1) c5	(2) c5	(3) c5
Constant	21.43 (34.78)	-10.76 (28.05)	-33.76 (45.81)
ST	-18.85 (20.22)	7.26 (13.77)	10.40 (22.96)
Fem	57.73** (23.38)	37.30** (14.82)	57.82** (21.44)
Experiment Quality			
P1 (total nr. of part.)		0.17*** (0.03)	
P2 (part. per treatm.)			0.54** (0.20)
P3 (treatments)			5.62 (3.52)
P4 (incentives)		-75.22 (43.29)	-95.55 (63.91)
Paper Quality			
P5 (pages)	-0.28 (0.79)	-1.58** (0.53)	-1.25 (0.93)
P6 (references)	0.62 (0.53)	0.57 (0.36)	0.79 (0.54)
P7 (reputation)	0.01 (0.01)	0.01 (0.01)	0.02 (0.01)
P8 (authors)	-6.17 (6.62)	7.77 (4.63)	6.31 (7.14)
N	21	15	15
Pseudo R^2	0.06	0.19	0.12

Table 14: Tobit regressions of c(5) on social ties (ST), rel. share of female authors (Fem), and experiments' and papers' quality proxies. JEEA papers. Standard errors in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 14 shows that, for papers published in the JEEA, the relative share of US-affiliated authors (*ST*) also has no significant effect on citations in the 5 years after publication. Note that *Table 14* only reports the results for models 1 to 3 because there are only four laboratory experiments for which we have data on the citations in the 10 years after publication. The *total number of participants* (P1) and the *number of participants per treatment* (P2) each have a positive effect. The *number of treatments* (P3) and the *strength of monetary incentives* (P4) are not significant.

For the JEEA, the results for a papers' quality (P5 to P7) show almost no effects. Only in model 2, papers with a higher *number of pages* (P5) tend to receive fewer citations.

When we change the specification of the dependent variable (see appendix B, *Table 30* and *Table 31*), for the JEEA, the positive effects for the *total number of participants* (P1) and *number of participants per treatment* (P2) persist. For both alternative specifications, the *number of references* (P6) has a positive effect in model 1.

6.4   How does the Authors' Gender Affect Citations?

For the AER (*Table 12* and appendix B, *Table 21* to *Table 24*), we observe that the relative share of female authors (*Fem*) has a significant positive effect on the number of citations a paper receives in the 5 years following publication for some specifications. For the specifications with our alternative proxy for social ties (experiment conducted in North America), the effect of *Fem* is strongest in models 2 and 3, where we control for the impact of the experimental quality proxies. For the JEEA (*Table 14* and appendix B, *Table 30* and *Table 31*), the relative share of female authors (*Fem*) has a strong significant positive effect on the number of citations in all specifications. For Exp Econ, no significant effect of the relative share of female authors (*Fem*) can be observed in any specification.

With regard to possible gender effects, our results for the AER and the JEEA are partly in line with the results of Card et al. (2020) and Hengel and Moon (2020). However, their results rely on many more observations than our results which is why we refrain from a deeper interpretation of our results in this aspect. Surprisingly, we observe no gender effects for Exp Econ. This could indicate that gender effects are mainly present in highly ranked journals.

6.5   Network Effects as an Alternative Explanation

Differences in citation practices and network effects could be an alternative explanation for the significantly lower number of citations received by US-affiliated authors in the AER. The majority of experimental economists (about 58 %) are affiliated with Europe[37]. If they are more familiar with each other's work, this could lead to a higher number of citations from European authors to other European authors. Put bluntly, most experimental economists are in Europe, and if they cite each other, this increases the number of citations received by European economist but not by non-European economists. This is in line with Frey and Pommerehne

---

[37] To get an idea of the total number of experimental economists in different regions, we checked the affiliations of all authors of the current RePEc list of authors in experimental economics ($N = 1831$). With 57.84%, more than half of all experimental economists on this list have a European affiliation, followed by economists with an affiliation in North America (30.20%), Asia (5.90%), Australia or New Zealand (3.82%), South America (1.86%), and Africa (0.38%). See, https://ideas.repec.org/i/eexp.html (accessed October 5, 2019).

(1988, 107) who argue that economists may tend to cite economists from the same country more often for personal or professional reasons, or simply because they know them best. If this is true, citations are a biased measure of quality. Moreover, the results from *Table 12* might not be driven by the composition of AER's editorial board, but simply by network effects resulting from differences in citation practices and differences in the population of experimental economists with a US- or European-affiliation.

To test this alternative explanation, we assess where citations originate from. We selected the most cited papers from our dataset[38] that were authored solely by economists with a US- (16 papers) or European-affiliation (11 papers), and analyzed where their received citations came from. The results show that the US-affiliated authors received, on average, 38.72% of their citations from North America, 50.42% of their citations from Europe, and 10.86% of their citations from the rest of the world. The respective numbers for the European-affiliated authors are 63.17% from Europe, 27.53% from North America and 9.31% from the rest of the world.

The numbers for the European-affiliated authors thus almost match the geographical distribution of experimental economists on the RePEc list. Compared to the European-affiliated authors, the US-affiliated authors received a considerably lower share of their citations from Europe. This suggests that differences in citation practices and network effects that increase the citations of European-affiliated authors could at least partly drive our results. However, if differences in citation practices and network effects were solely responsible for our findings regarding *ex post* quality, we would expect a negative effect of social ties across all journals. As shown in *Table 13* and *Table 14*, this is not or, in the case of alternative model specifications, only partially the case. Moreover, the assumption that social ties between experimental economists in Europe are as strong as those between experimental economists in the US does not seem plausible. While the majority of US-affiliated authors are likely to be native English speakers and all work within one country, our dataset for the European region includes authors with affiliations in 20 different countries. Only a few of the European-affiliated authors share the same mother tongue, and it is also likely that the European-affiliated authors will, most frequently, attend conferences and seminars at national level. For our dataset, we therefore believe that social ties between US-affiliated authors are more pronounced than between European-affiliated authors.

---

[38] The criteria were that these papers received at least 40 (100) citations in the 5 (10) years after publication.

6.6 Limitations

There are a number of limitations to our approach. First, we only collected data on laboratory experiments published in three journals. It is not clear whether a different journal selection would yield similar patterns on geographical concentration and the experiments' quality proxies. Despite the focus on three journals, we think that our results are noteworthy considering the *ex ante* and *ex post* quality differences between experiments conducted in North America and Europe.

Second, for the collection of citation data we used *Web of Science*. Since *Web of Science* has not indexed all journals and literature sources, the citations of *Web of Science* deviate from citation counts by other platforms, e.g. *Google Scholar*[39]. For our data set, we can therefore not claim to have recorded all citations received by a paper in the years after its publication, or to have recorded all citations received by a single author in the 5 years prior publication of the respective paper. However, we do not regard this incomplete recording of citations as problematic. While it can be expected to introduce a level effect, it should not yield any systematic misrepresentations between authors.

Third, when testing for editorial favoritism, our data set contains only a small number of laboratory experiments published in the AER. This implies that our results must be interpreted with caution. Though our findings are compatible with the existence of editorial favoritism, they do not prove its existence (see section 6.5). Furthermore, there is no reason to believe that our results, which are based exclusively on papers reporting results from laboratory experiments, can be transferred to other subfields of economics. In addition, in our test for editorial favoritism, we assume that editors have the primary goal of maximizing the impact factor of their journal and, thus, the number of citations that the published papers receive. However, we can neither be sure that this assumption is correct nor that other quality characteristics, such as relevance for policy-making or expected media coverage, do not play a stronger role in the selection of papers.

Fourth, our regressions in section 6 only include papers published before 2014. Thus, even if editorial favoritism and gender effects actually existed up to that year, we cannot say if these problems still persist in more recent years.

---

[39] See Hamermesh (2018) for a detailed analysis of the relation between *Web of Science* and *Google Scholar* citations.

Finally, it should be noted that we only recorded authors' affiliations at the time of publication. Hence, we could not capture any social ties that existed due to previous employment, research, or study stays.

7   Conclusion

In this paper, we examined geographical concentration, compared objectively measurable quality proxies, and tested for editorial favoritism in experimental economics.

We found that geographical concentration decreased. More precisely, the US's share in research output decreased from 66% to 37% while the shares of several European countries (Germany, the Netherlands, Switzerland, and the UK) and Australia increased. This is in line with other studies (Ek and Henrekson 2019, Kocher and Sutter 2001) that, although not focusing on any particular field, also report a decreasing US dominance across fields.

Looking more closely into the papers, we examine if four *ex ante* proxies for the quality of the experiments (*total number of participants*, *participants per treatment*, *number of treatments,* and *strength of monetary incentives*) differ between journals.

Comparing experiments conducted in North America and Europe, there are no differences regarding the *number of treatments* and no or only small differences regarding the *strength of monetary incentives*. However, European experiments rely on a significantly larger *total number of participants* and a significantly larger *number of participants per treatment*. This holds for all journals, but the difference is most pronounced for the AER where, on average, experiments conducted in North America have 55 participants per treatment while experiments conducted in Europe have 109 participants per treatment.

The differences in the *total number of participants* and *participants per treatment* could reflect different methodological standards. However, they could also indicate substantial barriers to entry for European economists, especially for the AER, where the North American dominance is more pronounced in the composition of the editorial board.

Analyzing the *number of citations* papers receive in the 5 and 10 years after publication, we find that experiments conducted in Europe receive more citations compared to experiments conducted in North America. The differences are statistically significant and large, except for papers published in the JEEA, possibly because of the small number of observations for this journal. In the 10 years after publication, North American AER papers receive, on average, 61

citations while European AER papers receive, on average, 124 citations. For Exp Econ, the corresponding numbers are 14 and 34.

Finally, we test for editorial favoritism. Editorial favoritism means that social ties (here: a vast majority of members of the editorial board and a positive share of authors having a US affiliation) have a negative effect on the quality of the published papers. We focus on the AER where almost all members of the editorial board are affiliated with US universities.

Our data provide some indications of editorial favoritism for the AER: papers authored by economists with a social tie to the editorial board receive between 12 (54) and 26 (153) fewer citations in the 5 (10) years after publication compared to papers authored by economists without a social tie to the editorial board. Due to the higher number of experimental economists in Europe, network effects may, at least partly, be driving these results. Nevertheless, network effects do not seem to provide the whole explanation, since a positive share of US-affiliated authors of a paper has no negative effect on the citation counts of Exp Econ and JEEA papers.

For AER papers, we only find small or null effects of the experiment's quality proxies on citations. This could be due to the small number of observations. The *number of references* and the *authors' reputation* have a positive effect on citations in the 5 years but no effect on citations in the 10 subsequent years after publication. For Exp Econ papers, we find that the *number of pages* has a negative effect and that the *number of references* a positive effect on citations in the 5 and 10 years after publication. Further, for Exp Econ papers, our results show that *authors' reputation* has a positive effect on the *number of citations* only in the first 5 years after publication.

For the AER and the JEEA we find evidence that the share of female authors has a significantly positive effect on the number of citations in the 5 years after publication.

As a concluding remark: we find that about one third of the papers reported incomplete information about their experimental procedures. In view of the replication crisis in social sciences, this is of particular concern since only a well-documented experiment can be replicated. In addition, the statistical processing of individual results in meta-studies is hampered if, for example, it is not clear how many persons participated in an experiment. Future research would benefit from clearly defined standards that specify which characteristics of a laboratory experiment should be reported in a paper.

References

Abbink, K., & Brandts, J. (2009). Collusion in growing and shrinking markets: empirical evidence from experimental duopolies. *Experiments and Competition Policy*, 34-60.

Afzal, U., d'Adda, G., Fafchamps, M., Quinn, S., & Said, F. (2017). Two sides of the same rupee? Comparing demand for microcredit and microsaving in a framed field experiment in rural Pakistan. *The Economic Journal*, *128*(614), 2161-2190.

Bardsley, N., Cubitt, R., Loomes, G., Moffat, P., Starmer, C., & Sugden, R. (2010). *Experimental economics: Rethinking the rules*. Princeton University Press.

Bayer, A., & Rouse, C. E. (2016). Diversity in the economics profession: A new attack on an old problem. *Journal of Economic Perspectives*, *30*(4), 221-42.

Bellemare, C., Bissonnette, L., & Kröger, S. (2014). Statistical power of within and between-subjects designs in economic experiments. IZA Discussion Paper No. 8583.

Bellemare, C., Bissonnette, L., & Kröger, S. (2016). Simulating power of economic experiments: the powerBBK package. *Journal of the Economic Science Association*, *2*(2), 157-168.

Blank, R. M. (1991). The effects of double-blind versus single-blind reviewing: Experimental evidence from the American Economic Review. *American Economic Review*, 1041-1067.

Brogaard, J., Engelberg, J., & Parsons, C. A. (2014). Networks and productivity: Causal evidence from editor rotations. *Journal of Financial Economics*, *111*(1), 251-270.

Bühren, C., Frank, B., Krabel, S., & Werner, A. (2012). Decision-making in competitive framings—Strategic behavior of chess players in mini-ultimatum game chess puzzles. *Economics Letters*, *115*(3), 356-358.

Camerer, C. F., Dreber, A., Forsell, E., Ho, T. H., Huber, J., Johannesson, M., Kirchler, M., Almenberg, J., Altmejd, A., Chan, T., Heikensten, E., Holzmeister, F., Imai, T., Isaksson, S., Nave, G., Pfeiffer, T., Razen, M. & Wu, H. (2016). Evaluating replicability of laboratory experiments in economics. *Science*, *351*(6280), 1433-1436.

Camerer, C. F., & Hogarth, R. M. (1999). The effects of financial incentives in experiments: A review and capital-labor-production framework. *Journal of Risk and Uncertainty*, *19*(1-3), 7-42.

Card, D., & DellaVigna, S. (2013). Nine facts about top journals in economics. *Journal of Economic Literature*, *51*(1), 144-61.

Card, D., & DellaVigna, S. (2014). Page limits on economics articles: Evidence from two journals. *Journal of Economic Perspectives*, *28*(3), 149-68.

Card, D., & DellaVigna, S. (2020). What do editors maximize? Evidence from four economics journals. *Review of Economics and Statistics*, *102*(1), 195-217.

Card, D., DellaVigna, S., Funk, P., & Iriberri, N. (2020). Are referees and editors in economics gender neutral?. *The Quarterly Journal of Economics*, *135*(1), 269-327.

Carpenter, J. P., Harrison, G. W., & List, J. A. (Eds.). (2005). *Field experiments in economics*. Elsevier JAI.

Colussi, T. (2018). Social ties in academia: A friend is a treasure. *Review of Economics and Statistics*, *100*(1), 45-50.

Combes, P. P., & Linnemer, L. (2003). Where are the economists who publish? Publication concentration and rankings in Europe based on cumulative publications. *Journal of the European Economic Association*, *1*(6), 1250-1308.

Czibor, E., Jimenez-Gomez, D., & List, J. A. (2019). The dozen things experimental economists should do (more of). *Southern Economic Journal*, *86*(2), 371-432.

Davis, D. D., & Holt, C. A. (1993). *Experimental economics*. Princeton university press.

Drazen, A., Almenberg, A. D., Ozbay, E. Y., & Snowberg, E. (2019). *A journal-based replication of "being chosen to lead"* (No. w26444). National Bureau of Economic Research.

Duflo, E. (2018). Report of the editor: American Economic Review. *AEA Papers and Proceedings*, 108, 636-51.

Ek, S., & Henrekson, M. (2019). The geography and concentration of authorship in the Top Five: Implications for European economics. *Scottish Journal of Political Economy*, *66*(2), 215-245.

Elliott, C., Greenaway, D., & Sapsford, D. (1998). Who's publishing who? The national composition of contributors to some core US and European journals. *European Economic Review*, *42*(1), 201-206.

Frey, B., & Pommerehne, W. (1988). The American domination among eminent economists. *Scientometrics*, *14*(1-2), 97-110.

Guala, F. (2005). *The methodology of experimental economics*. Cambridge University Press.

Glötzl, F., & Aigner, E. (2019). Six Dimensions of Concentration in Economics: Evidence from a Large-Scale Data Set. *Science in Context*, *32*(4), 381-410.

Hamermesh, D. S. (1994). Facts and myths about refereeing. *Journal of Economic Perspectives*, *8*(1), 153-163.

Hamermesh, D. S. (2018), Citations in economics: Measurement, uses, and impacts. *Journal of Economic Literature*, *56*(1), 115-56.

Harrison, G. W., & List, J. A. (2004). Field experiments. *Journal of Economic Literature*, *42*(4), 1009-1055.

Heckman, J. (2017), Publishing and promotion in economics: The curse of the top five, panel discussion at AEA meeting 2017.
[webcast: https://www.aeaweb.org/conference/webcasts/2017,
slides: https://hceconomics.uchicago.edu/sites/default/files/file_uploads/AEA-Curse-Five-HO-SMALL-STATIC_2017-01-06d_jbb.pdf, accessed Aug 12, 2019]

Hengel, E. (2019). Publishing while female. Are women held to higher standards? Evidence from peer review. Mimeo.

Hengel, E., & Moon, E. (2020). Gender and quality at top economics journals. Working Paper, University of Liverpool, UK.

Hertwig, R., & Ortmann, A. (2001). Experimental practices in economics: A methodological challenge for psychologists?. *Behavioral and Brain Sciences*, *24*(3), 383-403.

Hodgson, G. M., & Rothman, H. (1999). The editors and authors of economics journals: a case of institutional oligopoly?. *The Economic Journal*, *109*(453), 165-186.

Holt, C. A., & Laury, S. K. (2002). Risk aversion and incentive effects. *American Economic Review*, *92*(5), 1644-1655.

Imas, A. (2016). The realization effect: Risk-taking after realized versus paper losses. *American Economic Review*, *106*(8), 2086-2109.

Ioannidis, J. P., Stanley, T. D., & Doucouliagos, H. (2017). The power of bias in economics research. *The Economic Journal 127*(605), 236-265.

Kalaitzidakis, P., Mamuneas, T. P., & Stengos, T. (1999). European economics: An analysis based on publications in the core journals. *European Economic Review*, *43*(4-6), 1150-1168.

Kocher, M. G., & Sutter, M. (2001). The institutional concentration of authors in top journals of economics during the last two decades. *The Economic Journal*, *111*(472), 405-421.

Laband, D. N., & Piette, M. J. (1994a). Favoritism versus search for good papers: Empirical evidence regarding the behavior of journal editors. *Journal of Political Economy*, *102*(1), 194-203.

Laband, D. N., & Piette, M. J. (1994b). The relative impacts of economics journals: 1970-1990. *Journal of Economic Literature*, *32*(2), 640-666.

Levitt, S. D., List, J. A., & Sadoff, S. E. (2011). Checkmate: Exploring backward induction among chess players. *American Economic Review*, *101*(2), 975-90.

Lundberg, S., & Stearns, J. (2019). Women in economics: Stalled progress. *Journal of Economic Perspectives*, *33*(1), 3-22.

Medoff, M. H. (2003). Editorial favoritism in economics?. *Southern Economic Journal*, 425-434.

Merton, R. K. (1968). The Matthew effect in science: The reward and communication systems of science are considered. *Science*, *159*(3810), 56-63.

Moed, H. F. (2006). *Citation analysis in research evaluation* (Vol. 9). Springer Science & Business Media.

Moffatt, P. G. (2015). *Experimetrics: Econometrics for experimental economics*. Macmillan International Higher Education.

Neary, J. P., Mirrlees, J. A., & Tirole, J. (2003). Evaluating economics research in Europe: An introduction. *Journal of the European Economic Association*, *1*(6), 1239-1249.

Nikiforakis, N., & Slonim, R. (2019). Editors' Preface: Trends in experimental economics (1975–2018). *Journal of the Economic Science Association, 5*(2), 143-148.

Ortmann, A. (2009). The way in which an experiment is conducted is unbelievably important: On the experimentation practices of economists and psychologists. CESifo working paper No. 2887.

Read, D. (2005). Monetary incentives, what are they good for? *Journal of Economic Methodology*, *12*(2), 265-276.

Rydval, O. & Ortmann A. (2004). How financial incentives and cognitive abilities affect task performance in laboratory settings: An illustration. *Economics Letters*, *85*(3), 315-320.

Shepherd, G. B. (1995). *Rejected: Leading economists ponder the publication process*. T. Horton and Doughters.

Stewart, N., Ungemach, C., Harris, A. J., Bartels, D. M., Newell, B. R., Paolacci, G., & Chandler, J. (2015). The average laboratory samples a population of 7,300 Amazon Mechanical Turk workers. *Judgment and Decision Making*, *10*(5), 479-491.

Svorenčík, A. & Maas H. (Eds.) (2016). *The Making of Experimental Economics - Witness Seminar on the Emergence of a Field*. Heidelberg, Springer.

Zhang, L. & Ortmann A. (2013). Exploring the meaning of significance in experimental economics. UNSW Australian School of Business Research Paper, (2013-32).

Appendix

A. Geographical Concentration and Gender Effects

| | Shares | | | |
| | Scores | | | |
	AER	**Exp Econ**	**JEEA**	**All three**
Argentinia	0.50	0.00	0.00	0.17
	2.73	0.00	0.00	2.73
Australia	2.26	3.80	2.07	3.10
	12.37	33.62	3.32	49.31
Austria	2.63	4.86	0.00	3.60
	14.35	42.99	0.00	57.34
Azerbaijan	0.00	0.15	0.00	0.08
	0.00	1.29	0.00	1.29
Belgium	0.00	0.07	0.00	0.04
	0.00	0.60	0.00	0.60
Canada	1.79	2.71	1.53	2.27
	9.76	23.98	2.44	36.19
China	2.47	1.07	0.00	1.44
	13.48	9.42	0.00	22.90
Colombia	0.00	0.29	0.00	0.16
	0.00	2.54	0.00	2.54
Czech Republic	0.30	0.12	0.00	0.17
	1.66	1.07	0.00	2.73
Denmark	0.18	0.65	0.00	0.42
	0.97	5.76	0.00	6.72
Finland	0.00	0.13	0.00	0.07
	0.00	1.13	0.00	1.13
France	0.65	3.03	2.94	2.20
	3.53	26.80	4.69	35.03
Germany	4.41	12.08	13.29	9.57
	24.09	106.85	21.24	152.19
Guatemala	0.00	0.16	0.00	0.09
	0.00	1.42	0.00	1.42
India	0.00	0.30	0.00	0.16
	0.00	2.62	0.00	2.62
Israel	1.48	0.84	0.00	0.98
	8.12	7.46	0.00	15.58
Italy	0.39	3.50	1.55	2.24
	2.15	30.95	2.47	35.57
Japan	0.27	1.21	0.00	0.77
	1.49	10.74	0.00	12.23
Latvia	0.00	0.13	0.00	0.07
	0.00	1.11	0.00	1.11
Luxembourg	0.00	0.10	0.00	0.06
	0.00	0.92	0.00	0.92
Mexico	0.00	0.19	0.00	0.11
	0.00	1.69	0.00	1.69
Netherlands	3.30	4.68	7.35	4.47
	18.01	41.36	11.75	71.12
New Zealand	0.15	1.19	0.00	0.71
	0.80	10.54	0.00	11.34
Norway	0.54	0.68	6.31	1.20
	2.94	6.00	10.09	19.03
Portugal	0.00	0.45	0.00	0.25
	0.00	4.01	0.00	4.01

Russia	0.00	0.14	0.00	0.08
	0.00	1.28	0.00	1.28
Singapore	0.50	0.80	0.00	0.62
	2.73	7.06	0.00	9.79
South Korea	0.08	0.18	2.65	0.39
	0.46	1.59	4.23	6.28
Spain	2.68	5.88	2.04	4.40
	14.64	52.05	3.26	69.95
Sweden	0.71	1.25	0.00	0.94
	3.89	11.06	0.00	14.95
Switzerland	5.01	2.24	20.05	4.98
	27.39	19.85	32.05	79.29
Taiwan	0.00	0.25	0.00	0.14
	0.00	2.18	0.00	2.18
Turkey	0.29	0.09	0.00	0.15
	1.61	0.81	0.00	2.42
UK	6.08	8.80	9.50	7.94
	33.24	77.89	15.18	126.32
US	63.33	37.84	29.58	45.77
	346.12	334.79	47.27	728.18
UAE	0.00	0.14	0.00	0.08
	0.00	1.27	0.00	1.27
Uruguay	0.00	0.00	1.15	0.12
	0.00	0.00	1.83	1.83
Total	**100.00**	**100.00**	**100.00**	**100.00**
	546.55	**884.70**	**159.82**	**1591.08**

Table 15: Countries' shares and scores, all countries, pooled data from 1998 to 2018 for the AER, from 1999 to 2018 for Exp Econ and from 2011 to 2018 for the JEEA.

	Shares				
	Scores				
	1999-2003	**2004-2008**	**2009-2013**	**2014-2018**	**1999-2018**
Argentinia	0.00	0.00	0.00	0.65	0.19
	0.00	0.00	0.00	2.73	2.73
Australia	0.65	1.39	3.31	5.03	3.20
	0.71	3.94	20.12	21.23	46.00
Austria	0.61	0.12	6.53	3.94	3.99
	0.67	0.34	39.71	16.63	57.34
Azerbaijan	0.00	0.00	0.00	0.31	0.09
	0.00	0.00	0.00	1.29	1.29
Belgium	0.00	0.00	0.00	0.14	0.04
	0.00	0.00	0.00	0.60	0.60
Canada	1.55	0.87	2.91	2.81	2.37
	1.68	2.46	17.72	11.89	34.07
China	1.19	1.04	0.37	3.89	1.59
	1.29	2.95	2.23	16.42	22.90
Colombia	0.00	0.00	0.11	0.45	0.18
	0.00	0.00	0.64	1.90	2.54
Czech Republic	0.00	0.59	0.00	0.25	0.19
	0.00	1.66	0.00	1.07	2.73
Denmark	0.00	0.14	0.26	1.11	0.47
	0.00	0.41	1.61	4.71	6.72
Finland	0.26	0.00	0.00	0.20	0.08
	0.28	0.00	0.00	0.85	1.13
France	2.93	0.19	2.51	2.68	2.11
	3.18	0.55	15.27	11.33	30.33
Germany	4.66	5.46	9.50	12.47	9.13
	5.06	15.43	57.80	52.66	131.28
Guatemala	0.00	0.00	0.00	0.34	0.10
	0.00	0.00	0.00	1.42	1.42
India	0.00	0.23	0.18	0.20	0.18
	0.00	0.65	1.11	0.86	2.62
Israel	1.05	0.77	1.72	0.43	1.15
	1.13	2.18	10.44	1.82	16.58
Italy	0.60	1.22	2.94	2.63	2.34
	0.65	3.44	17.91	11.09	33.60
Japan	0.60	1.08	0.55	1.22	0.85
	0.65	3.07	3.36	5.16	12.23
Latvia	0.00	0.00	0.18	0.00	0.08
	0.00	0.00	1.11	0.00	1.11
Luxembourg	0.00	0.00	0.15	0.00	0.06
	0.00	0.00	0.92	0.00	0.92
Mexico	0.00	0.00	0.14	0.20	0.12
	0.00	0.00	0.85	0.83	1.69
Netherlands	2.27	4.15	3.67	5.42	4.13
	2.46	11.73	22.30	22.88	59.38
New Zealand	1.63	0.16	1.10	0.57	0.79
	1.77	0.45	6.72	2.40	11.34
Norway	0.00	1.04	0.00	1.42	0.62
	0.00	2.94	0.00	6.00	8.94
Portugal	0.00	0.00	0.20	0.66	0.28
	0.00	0.00	1.23	2.78	4.01
Russia	0.00	0.00	0.21	0.00	0.09
	0.00	0.00	1.28	0.00	1.28
Singapore	0.00	0.00	0.36	1.81	0.68
	0.00	0.00	2.16	7.63	9.79
South Korea	0.00	0.00	0.26	0.00	0.14

	0.00	0.00	1.59	0.00	2.05
Spain	1.66	2.77	8.13	1.80	4.66
	1.80	7.85	49.46	7.59	67.02
Sweden	0.99	1.68	0.23	1.84	1.04
	1.07	4.74	1.37	7.76	14.95
Switzerland	8.96	1.29	4.31	1.81	3.28
	9.72	3.64	26.24	7.64	47.24
Taiwan	0.00	0.47	0.00	0.20	0.15
	0.00	1.34	0.00	0.83	2.18
Turkey	1.48	0.00	0.13	0.00	0.17
	1.61	0.00	0.81	0.00	2.42
UK	3.00	6.45	8.86	8.46	7.83
	3.25	18.24	53.93	35.72	112.64
US	65.92	68.88	41.17	36.79	47.55
	71.52	194.84	250.48	155.38	683.91
UAE	0.00	0.00	0.00	0.30	0.09
	0.00	0.00	0.00	1.27	1.27
Total	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**
	108.50	**282.86**	**608.36**	**422.38**	**1438.26**

Table 16: Development of geographical concentration of AER and Exp Econ authors for all countries, data pooled over different periods. For each country, the first row reports the country's share and the second row reports the country's score.

Country	I: Score	II: Pop. (in 1000)	III: $\frac{I}{II} * 10^6$	IV: Score	V: Pop. (in 1000)	VI: $\frac{IV}{V} * 10^6$
Period	1999–2003	1999		2004-2008	2004	
Argentinia	0.00	36,648.1	0.00	0.00	38,728.7	0.00
Australia	0.71	18,926.0	0.04	3.94	20,127.4	0.20
Austria	0.67	7,992.3	0.08	0.34	8,172.0	0.04
Azerbaijan	0.00	7,982.8	0.00	0.00	8,306.5	0.00
Belgium	0.00	10,226.4	0.00	0.00	10,421.1	0.00
Canada	1.68	30,499.2	0.05	2.46	31,995.0	0.08
China	1.29	1,252,735.0	0.00	2.95	1,296,075.0	0.00
Colombia	0.00	39,819.3	0.00	0.00	42,724.2	0.00
Czech Rep.	0.00	10,283.9	0.00	1.66	10,197.1	0.16
Denmark	0.00	5,321.8	0.00	0.41	5,404.5	0.08
Finland	0.28	5,165.5	0.05	0.00	5,228.2	0.00
France	3.18	60,496.7	0.05	0.55	62,704.9	0.01
Germany	5.06	82,100.2	0.06	15.43	82,516.3	0.19
Guatemala	0.00	11,387.2	0.00	0.00	12,796.9	0.00
India	0.00	1,034,539.2	0.00	0.65	1,126,135.8	0.00
Israel	1.13	6,125.0	0.19	2.18	6,809.0	0.32
Italy	0.65	56,916.3	0.01	3.44	57,685.3	0.06
Japan	0.65	126,631.0	0.01	3.07	127,761.0	0.02
Latvia	0.00	2,390.5	0.00	0.00	2,263.1	0.00
Luxembourg	0.00	430.5	0.00	0.00	458.1	0.00
Mexico	0.00	100,300.6	0.00	0.00	106,995.6	0.00
Netherlands	2.46	15,812.1	0.16	11.73	16,281.8	0.72
New Zealand	1.77	3,835.1	0.46	0.45	4,087.5	0.11
Norway	0.00	4,461.9	0.00	2.94	4,591.9	0.64
Portugal	0.00	10,217.8	0.00	0.00	10,483.9	0.00
Russia	0.00	147,214.4	0.00	0.00	144,067.1	0.00
Singapore	0.00	3,958.7	0.00	0.00	4,166.7	0.00
South Korea	0.00	46,616.7	0.00	0.00	48,082.5	0.00
Spain	1.80	40,386.9	0.04	7.85	42,921.9	0.18
Sweden	1.07	8,857.9	0.12	4.74	8,993.5	0.53
Switzerland	9.72	7,144.9	1.36	3.64	7,389.6	0.49
Taiwan	0.00	21,712.0	0.00	1.34	22,462.0	0.06
Turkey	1.61	62,287.3	0.03	0.00	67,007.9	0.00
UK	3.25	58,682.5	0.06	18.24	59,987.9	0.30
US	71.52	279,040.0	0.26	194.84	292,805.3	0.67
UAE	0.00	2,988.2	0.00	0.00	4,087.9	0.00
Total	**108.50**			**282.86**		

Table 17: Scores for the geographical concentration of AER and Exp Econ authors per one million inhabitants. Data pooled for periods 1999-2003 and 2004-2008.

Country	I: Score	II: Pop. (in 1000)	III: $\frac{I}{II} * 10^6$	IV: Score	V: Pop. (in 1000)	VI: $\frac{IV}{V} * 10^6$
Period	**2009-2013**	**2009**		**2014-2018**	**2014**	
Argentinia	0.00	40,799.4	0.00	2.73	42,981.5	0.06
Australia	20.12	21,691.7	0.93	21.23	23,475.7	0.90
Austria	39.71	8,343.3	4.76	16.63	8,546.4	1.95
Azerbaijan	0.00	8,947.2	0.00	1.29	9,535.1	0.14
Belgium	0.00	10,796.5	0.00	0.60	11,209.1	0.05
Canada	17.72	33,628.6	0.53	11.89	35,535.3	0.33
China	2.23	1,331,260.0	0.00	16.42	1,364,270.0	0.01
Colombia	0.64	45,416.2	0.01	1.90	47,791.9	0.04
Czech Rep.	0.00	10,443.9	0.00	1.07	10,525.3	0.10
Denmark	1.61	5,523.1	0.29	4.71	5,643.5	0.83
Finland	0.00	5,338.9	0.00	0.85	5,461.5	0.16
France	15.27	64,707.0	0.24	11.33	66,316.1	0.17
Germany	57.80	81,902.3	0.71	52.66	80,982.5	0.65
Guatemala	0.00	14,316.12	0.00	1.42	15,923.6	0.09
India	1.11	1,214,270.1	0.00	0.86	1,293,859.3	0.00
Israel	10.44	7,485.6	1.39	1.82	8,215.7	0.22
Italy	17.91	59,095.4	0.30	11.09	60,789.1	0.18
Japan	3.36	128,047.0	0.03	5.16	127,276.0	0.04
Latvia	1.11	2,141.7	0.52	0.00	1,993.8	0.00
Luxembourg	0.92	497.8	1.85	0.00	556.3	0.00
Mexico	0.85	115,505.2	0.01	0.83	124,221.6	0.01
Netherlands	22.30	16,530.4	1.35	22.88	16,865.0	1.36
New Zealand	6.72	4,302.6	1.56	2.40	4,509.7	0.53
Norway	0.00	4,828.7	0.00	6.00	5,137.2	1.17
Portugal	1.23	10,568.2	0.12	2.78	10,401.1	0.27
Russia	1.28	142,785.3	0.01	0.00	143,819.7	0.00
Singapore	2.16	4,987.6	0.43	7.63	5,469.7	1.40
South Korea	1.59	49,307.8	0.03	0.00	50,746.7	0.00
Spain	49.46	46,362.9	1.07	7.59	46,480.9	0.16
Sweden	1.37	9,298.5	0.15	7.76	9,696.1	0.80
Switzerland	26.24	7,743.8	3.39	7.64	8,188.7	0.93
Taiwan	0.00	23,017.0	0.00	0.83	23,414.0	0.04
Turkey	0.81	71,339.2	0.01	0.00	77,030.7	0.00
UK	53.93	62,276.3	0.87	35.72	64,613.2	0.55
US	250.48	306,771.5	0.82	155.38	318,386.4	0.49
UAE	0.00	7,666.4	0.00	1.27	9,070.9	0.14
Total	**608.36**			**422.38**		

Table 18: Scores for the geographical concentration of AER and Exp Econ authors per one million inhabitants. Data pooled for periods 2009-2013 and 2014-2018.
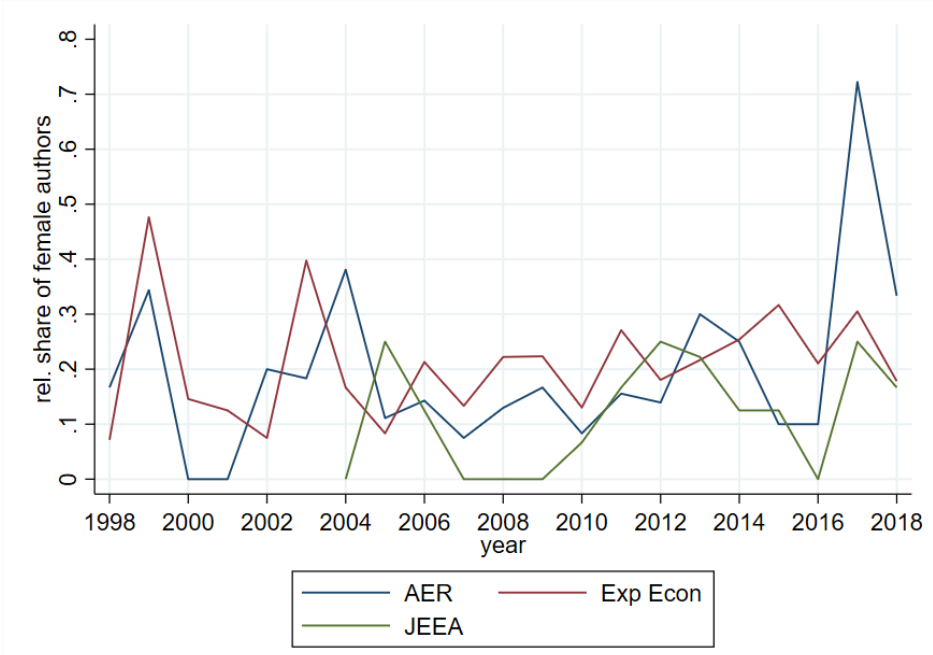
Figure 3: Development of the relative share of female authors among the three journals.

B. Citations, Quality Proxies and Editorial Favoritism

period	P1: # total participants	North America	Europe	both regions	*p*-value
1999 - 2003	AER	240.62 (159.08) *N* = 21	230.00 (124.78) *N* = 10	237.19 (146.86) *N* = 31	0.9834
	Exp Econ	134.15 (75.65) *N* = 41	154.85 (99.97) *N* = 34	143.53 (87.51) *N* = 75	0.3522
	both	170.21 (120.96) *N* = 62	171.93 (109.29) *N* = 44	170.92 (115.72) *N* = 106	0.7147
2004 - 2008	AER	201.16 (181.19) *N* = 25	432.71 (294.28) *N* = 7	251.81 (227.24) *N* = 32	0.0152
	Exp Econ	138.66 (72.10) *N* = 35	130.52 (104.34) *N* = 21	135.61 (84.79) *N* = 56	0.2332
	both	164.70 (131.59) *N* = 60	206.07 (212.29) *N* = 28	177.86 (161.57) *N* = 88	0.8814
2009 - 2013	AER	263.95 (171.46) *N* = 21	387.82 (249.52) *N* = 17	319.37 (216.13) *N* = 38	0.1020
	Exp Econ	162.86 (79.42) *N* = 28	217.24 (129.07) *N* = 49	197.47 (116.00) *N* = 77	0.0376
	both	206.18 (135.47) *N* = 49	261.18 (182.42) *N* = 66	237.75 (165.67) *N* = 115	0.0364
2014 - 2018	AER	124.44 (51.06) *N* = 9	437.60 (371.08) *N* = 5	236.29 (261.19) *N* = 14	0.1049
	Exp Econ	210.77 (110.08) *N* = 62	245.89 (126.28) *N* = 65	228.75 (119.49) *N* = 127	0.0897
	both	199.83 (108.14) *N* = 71	259.59 (158.89) *N* = 70	229.50 (138.53) *N* = 141	0.0145

Table 19: Mean values, standard deviations (in parentheses), and number of observations for number of total participants by period, region, and journal. The last column shows the p-value from a two-sided Mann-Whitney test.
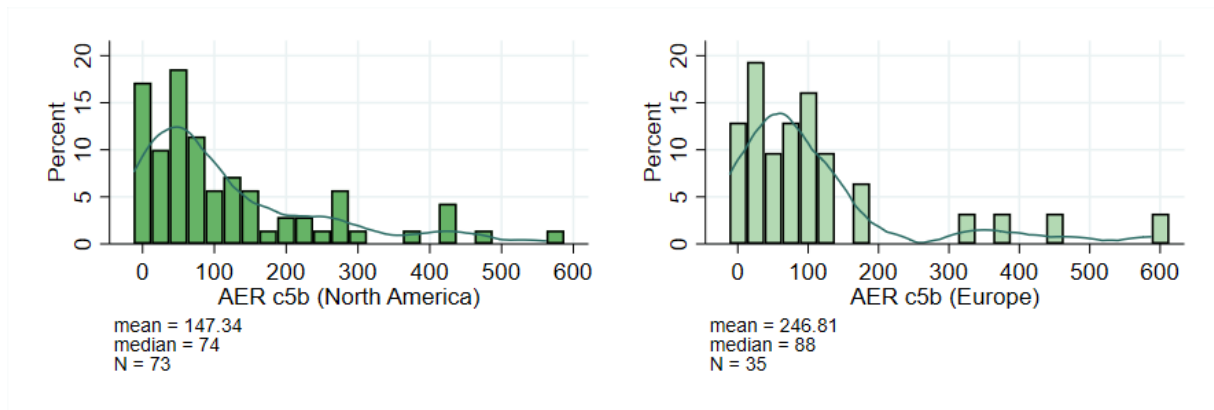
Figure 4: Distribution of citations received 5 years prior publication, for AER authors and experiments conducted in North America or Europe.
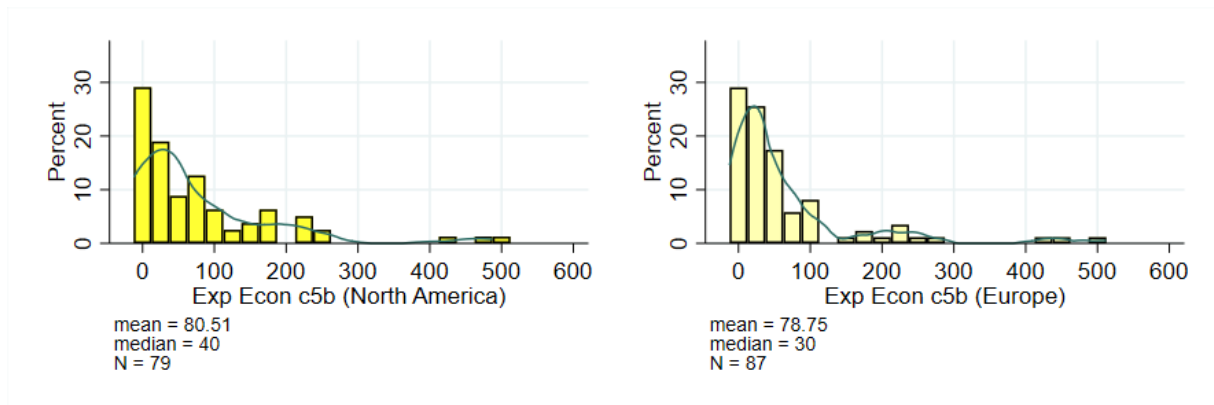
Figure 5: Distribution of citations received 5 years prior publication, for Exp Econ authors and experiments conducted in North America or Europe.

	(1) asinh(c5)	(2) asinh(c5)	(3) asinh(c5)	(4) asinh(c10)	(5) asinh(c10)	(6) asinh(c10)
Constant	3.42***	3.83***	3.85***	5.18***	6.34***	7.37***
	(0.40)	(0.60)	(0.66)	(0.59)	(1.03)	(1.05)
ST	-0.49**	-0.82***	-0.83***	-0.81**	-1.02**	-1.49***
	(0.19)	(0.27)	(0.27)	(0.32)	(0.47)	(0.45)
Fem	0.48	0.65	0.66	0.35	-0.03	-0.11
	(0.30)	(0.40)	(0.41)	(0.47)	(0.99)	(1.02)
Experiment Quality						
P1 (total nr. of part.)		-0.00 (0.00)			-0.00 (0.00)	
P2 (part. per treatm.)			-0.00 (0.00)			-0.00* (0.00)
P3 (treatments)			-0.03 (0.04)			-0.16 (0.10)
P4 (incentives)		0.28 (0.40)	0.31 (0.41)		0.33 (0.52)	-0.02 (0.51)
Paper Quality						
P5 (pages)	-0.01 (0.01)	0.01 (0.01)	0.01 (0.01)	-0.03 (0.02)	-0.03 (0.03)	-0.04 (0.04)
P6 (references)	0.02*** (0.01)	0.01 (0.01)	0.01 (0.01)	0.02* (0.01)	-0.00 (0.02)	0.01 (0.02)
P7 (reputation)	0.00** (0.00)	0.00* (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.01** (0.00)
P8 (authors)	-0.05 (0.09)	-0.16 (0.12)	-0.16 (0.12)	-0.03 (0.13)	-0.32 (0.24)	-0.35* (0.19)
Papers&Proceed.	-0.74** (0.29)	-0.35 (0.50)	-0.29 (0.57)	-0.96** (0.37)	-0.62 (0.70)	-0.64 (0.76)
N	127	68	66	71	26	25
Pseudo R^2	0.12	0.13	0.13	0.09	0.14	0.21

Table 20: Tobit regressions of asinh(c5) and asinh(c10) on social ties (ST), rel. share of female authors (Fem), and papers' and experiments' quality proxies. AER papers. Standard errors in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

	(1) log(1+c5)	(2) log(1+c5)	(3) log(1+c5)	(4) log(1+c10)	(5) log(1+c10)	(6) log(1+c10)
Constant	2.81***	3.14***	3.19***	4.52***	5.64***	6.64***
	(0.37)	(0.56)	(0.61)	(0.56)	(1.00)	(1.02)
ST	-0.45**	-0.75***	-0.77***	-0.78**	-1.00**	-1.46***
	(0.18)	(0.25)	(0.26)	(0.31)	(0.46)	(0.43)
Fem	0.46*	0.60	0.62	0.36	-0.03	-0.12
	(0.27)	(0.37)	(0.38)	(0.44)	(0.97)	(0.99)
Experiment Quality						
P1 (total nr. of part.)		-0.00			-0.00	
		(0.00)			(0.00)	
P2 (part. per treatm.)			-0.00			-0.00*
			(0.00)			(0.00)
P3 (treatments)			-0.03			-0.16
			(0.04)			(0.09)
P4 (incentives)		0.26	0.28		0.32	-0.01
		(0.37)	(0.38)		(0.51)	(0.49)
Paper Quality						
P5 (pages)	-0.01	0.01	0.01	-0.02	-0.03	-0.04
	(0.01)	(0.01)	(0.01)	(0.02)	(0.03)	(0.04)
P6 (references)	0.02***	0.01	0.01	0.02*	-0.00	0.01
	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)
P7 (reputation)	0.00**	0.00*	0.00	0.00	0.00	0.01**
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
P8 (authors)	-0.05	-0.15	-0.15	-0.03	-0.31	-0.34*
	(0.09)	(0.11)	(0.11)	(0.13)	(0.23)	(0.19)
Papers&Proceed.	-0.69**	-0.28	-0.22	-0.95***	-0.59	-0.62
	(0.27)	(0.47)	(0.53)	(0.35)	(0.68)	(0.74)
N	127	68	66	71	26	25
Pseudo R^2	0.13	0.14	0.14	0.09	0.14	0.21

Table 21: Tobit regressions of log(1+c5) and log(1+c10) on social ties (ST), rel. share of female authors (Fem), and papers' and experiments' quality proxies. AER papers. Standard errors in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

	(1) c5	(2) c5	(3) c5	(4) c10	(5) c10	(6) c10
Constant	17.15*	32.39*	37.30*	123.91***	173.22**	193.58**
	(9.85)	(18.53)	(20.15)	(41.25)	(73.92)	(79.76)
North America	-6.98*	-21.44***	-23.16***	-23.63	-59.51*	-75.44**
	(4.15)	(7.01)	(7.29)	(18.53)	(28.30)	(29.13)
Fem	13.66*	27.93**	29.25**	15.73	-28.40	-25.91
	(8.16)	(13.39)	(13.64)	(37.81)	(85.14)	(97.23)
Experiment Quality						
P1		-0.01			0.01	
(total nr. of part.)		(0.02)			(0.08)	
P2			-0.03			-0.14
(part. per treatm.)			(0.03)			(0.11)
P3			-0.94			-4.56
(treatments)			(1.35)			(9.52)
P4		-2.60	-3.42		21.79	19.00
(incentives)		(13.01)	(13.38)		(43.85)	(45.52)
Paper Quality						
P5	-0.28	-0.15	-0.05	-1.88	-0.31	0.61
(pages)	(0.27)	(0.43)	(0.45)	(1.31)	(2.60)	(3.30)
P6	0.65***	0.55**	0.53**	0.60	-1.61	-1.58
(references)	(0.16)	(0.24)	(0.24)	(0.81)	(1.62)	(1.89)
P7	0.02***	0.02*	0.02*	0.04	0.02	0.11
(reputation)	(0.01)	(0.01)	(0.01)	(0.09)	(0.28)	(0.32)
P8	-2.08	-4.38	-4.97	-5.20	-9.88	-11.37
(authors)	(2.48)	(3.73)	(3.87)	(10.54)	(19.66)	(17.58)
Papers&Proceed.	-12.30	-2.17	5.31	-71.09**	-32.21	-20.16
	(7.97)	(16.27)	(18.44)	(29.56)	(62.10)	(70.17)
N	127	68	66	71	26	25
Pseudo R^2	0.03	0.04	0.04	0.01	0.03	0.03

Table 22: Tobit regressions of c(5) and c(10) on place of experimentation (North America), rel. share of female authors (Fem), and experiments' and papers' quality characteristics. AER papers. Standard errors in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

	(1) asinh(c5)	(2) asinh(c5)	(3) asinh(c5)	(4) asinh(c10)	(5) asinh(c10)	(6) asinh(c10)
Constant	3.08***	3.63***	3.73***	4.50***	5.56***	6.15***
	(0.37)	(0.55)	(0.60)	(0.53)	(0.84)	(0.86)
North America	-0.24	-0.78***	-0.82***	-0.24	-0.67*	-1.00***
	(0.16)	(0.21)	(0.22)	(0.24)	(0.32)	(0.31)
Fem	0.47	0.92**	0.96**	0.17	-0.27	0.13
	(0.30)	(0.40)	(0.41)	(0.49)	(0.97)	(1.04)
Experiment Quality						
P1		-0.00			0.00	
(total nr. of part.)		(0.00)			(0.00)	
P2			-0.00			-0.00
(part. per treatm.)			(0.00)			(0.00)
P3			-0.03			-0.22*
(treatments)			(0.04)			(0.10)
P4		0.27	0.26		0.53	0.24
(incentives)		(0.39)	(0.40)		(0.50)	(0.49)
Paper Quality						
P5	-0.01	0.01	0.01	-0.02	-0.02	-0.01
(pages)	(0.01)	(0.01)	(0.01)	(0.02)	(0.03)	(0.04)
P6	0.02***	0.01	0.01	0.02*	-0.01	0.02
(references)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)
P7	0.00**	0.00*	0.00	0.00	0.00	0.01*
(reputation)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
P8	-0.01	-0.13	-0.13	0.05	-0.21	-0.25
(authors)	(0.09)	(0.11)	(0.12)	(0.14)	(0.22)	(0.19)
Papers&Proceed.	-0.79***	-0.41	-0.26	-0.96**	-0.58	-0.17
	(0.30)	(0.48)	(0.55)	(0.38)	(0.70)	(0.75)
N	127	68	66	71	26	25
Pseudo R^2	0.11	0.15	0.16	0.06	0.13	0.20

Table 23: Tobit regressions of asinh(c5) and asinh(c10) on place of experimentation (North America), rel. share of female authors (Fem), and experiments' and papers' quality characteristics. AER papers. Standard errors in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

	(1)	(2)	(3)	(4)	(5)	(6)
	log(1+c5)	log(1+c5)	log(1+c5)	log(1+c10)	log(1+c10)	log(1+c10)
Constant	2.51***	2.96***	3.07***	3.89***	4.87***	5.44***
	(0.34)	(0.51)	(0.56)	(0.51)	(0.82)	(0.84)
North America	-0.23	-0.72***	-0.76***	-0.24	-0.65*	-0.98***
	(0.14)	(0.19)	(0.20)	(0.23)	(0.31)	(0.31)
Fem	0.45	0.85**	0.90**	0.18	-0.27	0.12
	(0.28)	(0.37)	(0.38)	(0.46)	(0.94)	(1.02)
Experiment Quality						
P1		-0.00			0.00	
(total nr. of part.)		(0.00)			(0.00)	
P2			-0.00			-0.00
(part. per treatm.)			(0.00)			(0.00)
P3			-0.03			-0.21*
(treatments)			(0.04)			(0.10)
P4		0.25	0.23		0.52	0.24
(incentives)		(0.36	(0.37)		(0.49)	(0.48)
Paper Quality						
P5	-0.01	0.01	0.01	-0.02	-0.02	-0.01
(pages)	(0.01)	(0.01)	(0.01)	(0.02)	(0.03)	(0.03)
P6	0.02***	0.01	0.01	0.02*	-0.01	0.02
(references)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)
P7	0.00**	0.00*	0.00	0.00	0.00	0.01*
(reputation)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
P8	-0.01	-0.11	-0.12	0.04	-0.21	-0.24
(authors)	(0.08)	(0.10)	(0.11)	(0.13)	(0.22)	(0.18)
Papers&Proceed.	-0.74***	-0.35	-0.19	-0.96**	-0.55	-0.15
	(0.27)	(0.45)	(0.51)	(0.36)	(0.69)	(0.74)
N	127	68	66	71	26	25
Pseudo R^2	0.12	0.16	0.16	0.07	0.14	0.20

Table 24: Tobit regressions of log(1+c5) and log(1+c10) on place of experimentation (North America), rel. share of female authors (Fem), and experiments' and papers' quality characteristics. AER papers. Standard errors in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

	(1)	(2)	(3)	(4)	(5)	(6)
	c5	c5	c5	c10	c10	c10
Constant	9.37***	7.60*	7.28*	33.43**	8.23	-3.64
	(3.02)	(3.88)	(4.19)	(13.60)	(18.35)	(18.74)
North America	-2.68*	-2.93*	-2.78*	-11.09*	-9.37	-10.17
	(1.50)	(1.62)	(1.61)	(6.29)	(8.03)	(7.51)
Fem	-1.40	-1.17	-2.09	-3.11	8.44	-3.09
	(2.63)	(2.83)	(2.89)	(11.95)	(14.86)	(14.51)
Experiment Quality						
P1		0.00			0.08*	
(total nr. of part.)		(0.01)			(0.05)	
P2			0.00			0.08
(part. per treatm.)			(0.02)			(0.15)
P3			0.56			6.77***
(treatments)			(0.38)			(2.20)
P4		4.13	2.94		18.55	22.47
(incentives)		(4.16)	(4.20)		(22.62)	(21.31)
Paper Quality						
P5	-0.43***	-0.44***	-0.46***	-1.19*	-1.31*	-1.15*
(pages)	(0.13)	(0.13)	(0.13)	(0.60)	(0.70)	(0.66)
P6	0.27***	0.20***	0.21***	0.59**	0.48	0.50
(references)	(0.06)	(0.07)	(0.07)	(0.29)	(0.35)	(0.33)
P7	0.01**	0.03***	0.03***	0.03	0.06	0.03
(reputation)	(0.01)	(0.01)	(0.01)	(0.03)	(0.07)	(0.07)
P8	-0.49	-0.12	-0.33	-1.42	2.81	1.34
(authors)	(0.78)	(0.84)	(0.84)	(3.38)	(4.41)	(4.12)
N	201	138	138	84	50	50
Pseudo R^2	0.02	0.03	0.03	0.01	0.03	0.04

Table 25: Tobit regressions of c(5) and c(10) on place of experimentation (North America), rel. share of female authors (Fem), and experiments' and papers' quality characteristics. Exp Econ papers. Standard errors in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

	(1) asinh(c5)	(2) asinh(c5)	(3) asinh(c5)	(4) asinh(c10)	(5) asinh(c10)	(6) asinh(c10)
Constant	2.19***	2.04***	1.90***	3.13***	2.32***	2.06***
	(0.32)	(0.44)	(0.48)	(0.56)	(0.68)	(0.74)
North America	-0.36**	-0.34*	-0.33*	-0.45*	-0.55*	-0.58*
	(0.16)	(0.18)	(0.18)	(0.26)	(0.30)	(0.30)
Fem	-0.11	-0.17	-0.19	0.23	0.51	0.42
	(0.28)	(0.32)	(0.33)	(0.49)	(0.55)	(0.57)
Experiment Quality						
P1		0.00			0.00	
(total nr. of part.)		(0.00)			(0.00)	
P2			0.00			0.01
(part. per treatm.)			(0.00)			(0.01)
P3			0.05			0.10
(treatments)			(0.04)			(0.09)
P4		0.53	0.48		1.09	1.19
(incentives)		(0.47)	(0.48)		(0.84)	(0.84)
Paper Quality						
P5	-0.04***	-0.04***	-0.04***	-0.05*	-0.06**	-0.06**
(pages)	(0.01)	(0.01)	(0.01)	(0.02)	(0.03)	(0.03)
P6	0.03***	0.02**	0.02**	0.03**	0.03**	0.03**
(references)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
P7	0.00***	0.00***	0.00***	0.00**	0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
P8	0.04	0.05	0.04	0.01	0.17	0.15
(authors)	(0.08)	(0.10)	(0.10)	(0.14)	(0.16)	(0.16)
N	201	138	138	84	50	50
Pseudo R^2	0.05	0.06	0.06	0.05	0.14	0.14

Table 26: Tobit regressions of asinh(c5) and asinh(c10) on place of experimentation (North America), rel. share of female authors (Fem), and experiments' and papers' quality characteristics. Exp Econ papers. Standard errors in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

61

	(1) log(1+c5)	(2) log(1+c5)	(3) log(1+c5)	(4) log(1+c10)	(5) log(1+c10)	(6) log(1+c10)
Constant	1.76***	1.62***	1.51***	2.59***	1.85***	1.61**
	(0.27)	(0.37)	(0.40)	(0.50)	(0.61)	(0.67)
North America	-0.30**	-0.29*	-0.28*	-0.40*	-0.47*	-0.49*
	(0.13)	(0.15)	(0.15)	(0.23)	(0.27)	(0.27)
Fem	-0.10	-0.14	-0.17	0.19	0.46	0.35
	(0.23)	(0.27)	(0.28)	(0.44)	(0.50)	(0.52)
Experiment Quality						
P1 (total nr. of part.)		0.00 (0.00)			0.00 (0.00)	
P2 (part. per treatm.)			0.00 (0.00)			0.01 (0.01)
P3 (treatments)			0.04 (0.04)			0.10 (0.08)
P4 (incentives)		0.44 (0.40)	0.39 (0.40)		0.94 (0.76)	1.03 (0.76)
Paper Quality						
P5 (pages)	-0.03***	-0.04***	-0.04***	-0.04*	-0.05**	-0.05**
	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)
P6 (references)	0.02***	0.02**	0.02**	0.02**	0.03**	0.02**
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
P7 (reputation)	0.00***	0.00***	0.00***	0.00**	0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
P8 (authors)	0.03	0.04	0.03	0.01	0.15	0.13
	(0.07)	(0.08)	(0.08)	(0.12)	(0.15)	(0.15)
N	201	138	138	84	50	50
Pseudo R^2	0.06	0.07	0.07	0.06	0.14	0.14

Table 27: Tobit regressions of log(1+c5) and log(1+c10) on place of experimentation (North America), rel. share of female authors (Fem), and experiments' and papers' quality characteristics. Exp Econ papers. Standard errors in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

	(1) asinh(c5)	(2) asinh(c5)	(3) asinh(c5)	(4) asinh(c10)	(5) asinh(c10)	(6) asinh(c10)
Constant	2.17***	1.90***	1.75***	3.18***	2.25***	1.97**
	(0.32)	(0.44)	(0.48)	(0.58)	(0.68)	(0.74)
ST	-0.29*	-0.20	-0.19	-0.47	-0.54	-0.62
	(0.17)	(0.20)	(0.20)	(0.30)	(0.36)	(0.37)
Fem	-0.11	-0.13	-0.15	0.26	0.68	0.55
	(0.28)	(0.33)	(0.33)	(0.50)	(0.58)	(0.59)
Experiment Quality						
P1 (total nr. of part.)		0.00			0.00	
		(0.00)			(0.00)	
P2 (part. per treatm.)			0.00			0.01
			(0.00)			(0.01)
P3 (treatments)			0.06			0.13
			(0.04)			(0.09)
P4 (incentives)		0.50	0.45		0.96	1.08
		(0.48)	(0.48)		(0.84)	(0.84)
Paper Quality						
P5 (pages)	-0.04***	-0.04***	-0.04***	-0.05*	-0.06**	-0.05*
	(0.01)	(0.01)	(0.02)	(0.02)	(0.03)	(0.03)
P6 (references)	0.03***	0.02**	0.02**	0.03**	0.03*	0.03*
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
P7 (reputation)	0.00***	0.00***	0.00***	0.00**	0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
P8 (authors)	0.04	0.06	0.05	0.00	0.20	0.17
	(0.08)	(0.10)	(0.10)	(0.14)	(0.16)	(0.16)
N	201	138	138	84	50	50
Pseudo R^2	0.05	0.06	0.06	0.05	0.13	0.13

Table 28: Tobit regressions of asinh(c5) and asinh(c10) on social ties (ST), rel. share of female authors (Fem), and papers' and experiments' quality proxies. Exp Econ papers. Standard errors in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

	(1) log(1+c5)	(2) log(1+c5)	(3) log(1+c5)	(4) log(1+c10)	(5) log(1+c10)	(6) log(1+c10)
Constant	1.74***	1.51***	1.39***	2.64***	1.78***	1.52**
	(0.27)	(0.37)	(0.40)	(0.51)	(0.62)	(0.67)
ST	-0.25*	-0.17	-0.16	-0.42	-0.44	-0.52
	(0.14)	(0.17)	(0.17)	(0.27)	(0.33)	(0.34)
Fem	-0.10	-0.11	-0.14	0.22	0.59	0.45
	(0.23)	(0.27)	(0.28)	(0.44)	(0.52)	(0.53)
Experiment Quality						
P1 (total nr. of part.)		0.00 (0.00)			0.00 (0.00)	
P2 (part. per treatm.)			0.00 (0.00)			0.00 (0.01)
P3 (treatments)			0.05 (0.04)			0.13 (0.08)
P4 (incentives)		0.42 (0.40)	0.36 (0.41)		0.82 (0.76)	0.94 (0.76)
Paper Quality						
P5 (pages)	-0.04***	-0.04***	-0.04***	-0.04*	-0.05**	-0.05*
	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)
P6 (references)	0.02***	0.02**	0.02**	0.02**	0.02**	0.02*
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
P7 (reputation)	0.00***	0.00***	0.00***	0.00**	0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
P8 (authors)	0.03	0.05	0.04	0.00	0.18	0.15
	(0.07)	(0.08)	(0.08)	(0.12)	(0.15)	(0.15)
N	201	138	138	84	50	50
Pseudo R^2	0.05	0.07	0.07	0.05	0.14	0.14

Table 29: Tobit regressions of log(1+c5) and log(1+c10) on social ties (ST), rel. share of female authors (Fem), and papers' and experiments' quality proxies. Exp Econ papers. Standard errors in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

	(1)	(2)	(3)
	asinh(c5)	asinh(c5)	asinh(c5)
Constant	3.02***	1.13	0.45
	(0.89)	(0.82)	(1.12)
ST	-0.45	-0.10	0.14
	(0.52)	(0.40)	(0.56)
Fem	1.27*	1.22**	1.70**
	(0.60)	(0.44)	(0.52)
Experiment Quality			
P1		0.00***	
(total nr. of part.)		(0.00)	
P2			0.01**
(part. per treatm.)			(0.00)
P3			0.09
(treatments)			(0.09)
P4		0.35	-0.02
(incentives)		(1.27)	(1.56)
Paper Quality			
P5	-0.01	-0.02	-0.01
(pages)	(0.02)	(0.02)	(0.02)
P6	0.03**	0.03**	0.04**
(references)	(0.01)	(0.01)	(0.01)
P7	-0.00	-0.00	-0.00
(reputation)	(0.00)	(0.00)	(0.00)
P8	-0.15	0.19	0.14
(authors)	(0.17)	(0.14)	(0.17)
N	21	15	15
Pseudo R^2	0.24	0.60	0.44

Table 30: Tobit regressions of asinh(c5) on social ties (ST), rel. share of female authors (Fem), and papers' and experiments' quality proxies. JEEA papers. Standard errors in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

	(1) log(1+c5)	(2) log(1+c5)	(3) log(1+c5)
Constant	2.45**	0.79	0.17
	(0.84)	(0.77)	(1.05)
ST	-0.44	-0.07	0.15
	(0.49)	(0.38)	(0.53)
Fem	1.22**	1.15**	1.60**
	(0.56)	(0.41)	(0.49)
Experiment Quality			
P1		0.00***	
(total nr. of part.)		(0.00)	
P2			0.01**
(part. per treatm.)			(0.00)
P3			0.08
(treatments)			(0.08)
P4		0.10	-0.25
(incentives)		(1.19)	(1.47)
Paper Quality			
P5	-0.01	-0.02	-0.01
(pages)	(0.02)	(0.01)	(0.02)
P6	0.03**	0.03**	0.03**
(references)	(0.01)	(0.01)	(0.01)
P7	-0.00	-0.00	-0.00
(reputation)	(0.00)	(0.00)	(0.00)
P8	-0.14	0.18	0.13
(authors)	(0.16)	(0.13)	(0.16)
N	21	15	15
Pseudo R^2	0.25	0.63	0.46

Table 31: Tobit regressions of log(1+c5) on social ties (ST), rel. share of female authors (Fem), and papers' and experiments' quality proxies. JEEA papers. Standard errors in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$