I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

Discrimination in the Formation of Academic Networks:
A Field Experiment on *#EconTwitter*

Nicolás Ajzenman
Bruno Ferman
Pedro C. Sant'Anna

DISCUSSION PAPER SERIES

Discrimination in the Formation of Academic Networks:
A Field Experiment on *#EconTwitter*

Nicolás Ajzenman
McGill University, São Paulo School of Economics and IZA

Bruno Ferman
São Paulo School of Economics - FGV

Pedro C. Sant'Anna
São Paulo School of Economics - FGV

ABSTRACT

Discrimination in the Formation of Academic Networks: A Field Experiment on *#EconTwitter**

This paper assesses the results of an experiment designed to identify discrimination in users' following behavior on Twitter. Specifically, we created fictitious bot accounts that resembled humans and claimed to be PhD students in economics. The accounts differed in three characteristics: gender (male or female), race (Black or White), and university affiliation (top- or lower-ranked). The bot accounts randomly followed Twitter users who form part of the *#EconTwitter* academic community. We measured how many follow-backs each account obtained after a given period. Twitter users from this community were 12% more likely to follow accounts of White students compared to those of Black students; 21% more likely to follow accounts of students from top-ranked, prestigious universities compared to accounts of lower-ranked institutions; and 25% more likely to follow female compared to male students. The racial gap persisted even among students from top-ranked institutions, suggesting that Twitter users racially discriminate even in the presence of a signal that could be interpreted as indicative of high academic potential. Notably, we find that Black male students from top-ranked universities receive no more follow-backs than White male students from relatively lower-ranked institutions.

JEL Classification:	J15, J16, A11, C93, I23
Keywords:	discrimination, economics profession, gender, race, social media

Corresponding author:
Nicolas Ajzenman
FGV EESP - Escola de Economia de São Paulo
474 Rua Itapeva, Bela Vista
01332-000 São Paulo
Brazil
E-mail: nicolas.ajzenman@fgv.br

1 Introduction

Social and professional networks are important determinants of labor market outcomes, especially in academia, where collaboration is central (Jackson et al., 2017; Goyal et al., 2006). In academia, such networks help explain researchers' productivity (Rose and Georg, 2021; Ductor et al., 2014; Azoulay et al., 2010) and can increase the likelihood of publications (Ductor and Visser, 2022) or promotions (Zinovyeva and Bagues, 2015). However, individuals' access to formal and informal networks within academia is far from homogeneous. This access may depend on group-based characteristics such as gender, race, or university affiliation (Beaman et al., 2018). If individuals with some of those characteristics are discriminated against when forming their networks, they may have worse professional outcomes, and this may negatively impact diversity within the profession. This lack of diversity has been extensively documented in academia —- particularly in economics — with many arguing that it may lead to less innovative ideas and talent misallocation (Bayer and Rouse, 2016; Lundberg and Stearns, 2019; Bayer et al., 2020; Schultz and Stansbury, 2022). To help remedy this situation, professional associations such as the American Economic Association (AEA) have sought to promote diversity and create a more inclusive environment for underrepresented groups in economics. Nevertheless, intense debate persists over the causes of such disparities.

In this paper, we investigate one plausible cause: discrimination in the formation of professional networks. We focus in particular on discrimination among academic economists in a social media setting: Twitter. Social media in general (and Twitter specifically) has become an important networking tool for academics in recent years, providing a means to debate ideas, connect with other researchers, publicize work and meet potential collaborators. A micro-blogging and social networking service on which people communicate in short messages, Twitter has become a particularly popular platform among academics. The platform is often recommended to researchers, especially to those in the early stages of their careers (e.g., Cheplygina et al., 2020). There is also evidence that Twitter offers tangible benefits to scientists; tweeting about a paper may, for instance, increase its citation count (Luc et al., 2021). Among academic economists, Twitter seems to be especially favored. For instance, between January and February 2022, over 14,000 accounts used the hashtag *#EconTwitter*, which this group typically uses. Moreover, conferences such as the AEA Annual Meeting have featured panels on how to use Twitter (e.g., Wolfers, 2015), suggesting that many economists consider this tool to be valuable for networking. In order to obtain the greatest benefit from Twitter, it is essential to constitute a network of followers — users who see one's posts on their timeline and are more likely to interact with those posts. The increasing importance of Twitter in forming academic networks highlights the need to understand whether, and to what extent, users discriminate in their decisions of whom to follow on the platform.

To study this question, we conducted a pre-registered experiment to identify discrimination in Twitter users' following behavior, focusing on the academic economics community on the platform.[1] We created fictitious accounts ('bots') on Twitter that resembled humans

---

[1] AEA RCT Registry ID AEARCTR-0009507. The experiment was approved by the Ethical Compliance Committee on Research Involving Human Beings at Fundação Getulio Vargas (CEPH/FGV, IRB approval

and claimed to be PhD students in economics. The accounts differed in terms of three characteristics: gender (male or female), race (Black or white), and university affiliation (top- or lower-ranked university). These are the most salient dimensions in the debate about the lack of diversity in economics, as demonstrated by the results of a recent climate survey conducted by the AEA among current and former members of the association (Allgood et al., 2019).[2] The accounts randomly followed Twitter users who are part of the academic economics community, and we measured how many follow-backs each account obtained. The use of follow-backs as an outcome is standard in the literature (e.g., Mosleh et al., 2021) and is economically interesting because this action has a relatively low cost, suggesting that the disparities we find could be even larger in other (costlier) situations.

We find significant differences in the rate of follow-backs across all three of the dimensions we experimentally varied. First, members of the *#EconTwitter* community are 2.1 percentage points more likely to follow White than Black PhD students. Second, they are 3.5 pp more likely to follow students of top-ranked schools than those from lower-ranked universities. Third, they are 4.3 pp more likely to follow female than male students. In relative terms, these results mean that members of this community are 12% more likely to follow White students compared to Black students, 21% more likely to follow accounts of students affiliated with top-ranked universities compared to students of lower-ranked ones, and 25% more likely to follow female compared to male students.

Interestingly, we also show that the racial gap in follow-backs is maintained among students claiming to be affiliated with top-ranked universities. This suggests that racial discrimination persists even in the presence of a signal that could be interpreted by some as indicative of higher academic potential. Indeed, a Black man from a top-ranked institution receives roughly the same rate of follow-backs as a Wite man from a relatively lower-ranked university (16.7% against 16.4%). In other words, the premium (in terms of follow-backs) that a Black male PhD student receives from being affiliated with a top-ranked university would be just enough for him to garner as many follow-backs as a White male student affiliated with a lower-ranked institution. For comparison, a White male student affiliated with a top-ranked university is followed back 18.7% of the time, which is 2 pp more than a Black student from the same type of institution.

We also gathered rich data on our subjects and performed a set of pre-registered heterogeneity analyses. We created a binary measure of concern about the lack of diversity in the field of economics, which takes a value of one if the user follows the Twitter account of an organization that addresses this topic (approximately 15% of our sample follow one or more such accounts). We show that, on average, users who signal a concern about the lack of diversity in economics do not seem to exhibit racial discrimination in their following behavior. However, they are relatively more biased in following bots from top-ranked universities than users who are not part of this group. This result suggests that Twitter users who

---

n. 034/2022).

[2]There are several other extremely important dimensions when it comes to disparities in economics and beyond that would have been interesting to study. For instance, recent work explores discrimination against Asian Americans in the context of college admissions (Arcidiacono et al., 2022). Biases against LGBTQ+ individuals or relative to country of origin are also significant. We focused here on the three above-mentioned dimensions given their salience and considerations of the experiment's power and feasibility.

signal concern about the lack of diversity in economics do not discriminate based on race, but do still have some blind spots in their behavior, namely relating to university affiliation and academic elitism. In addition, we explore other dimensions of heterogeneity, including users' number of followers, which we interpret as a measure of social media reach. We find no difference in the follow-back behavior of subjects with a number of followers above and below the median number among our user sample. These findings suggest that the behavior we encounter is not restricted to a subset of *#EconTwitter* accounts with relatively limited reach.

Overall, our results indicate that the formation of networks within the economic community on Twitter is unequal. We interpret this as evidence of discrimination across the three dimensions we studied. Indeed, apart from the group-based characteristics we experimentally manipulated, all accounts are identical, indicating that any differential treatment they receive is due to discrimination. The findings related to race and university affiliation align with economists' perceptions as captured by AEA's climate survey (Allgood et al., 2019), and with the relatively scarce literature on disparities in these dimensions within academia. We therefore fill a gap in the literature by documenting discrimination on the basis of race and university affiliation in an academic setting.

However, the result for gender — subjects were more likely to follow-back female than male accounts — runs counter to the overwhelming evidence, both within economics and in other contexts, of discrimination against women. In particular, we find that the differential in following-back male and female bots is larger among men than women, suggesting that the result we obtained is mostly driven by male Twitter users disproportionately following-back female bot accounts. We note that different mechanisms may be at play to explain this result. Some users, conscious of the barriers faced by women in the profession, could be attempting to engage more with women to correct for those barriers, in a type of affirmative action or positive discrimination. It is, however, also possible that some were using Twitter with the objective of establishing social rather than professional connections, and disproportionately wish to establish such connections with women. These two motives might have different implications in terms of the consequences of having more Twitter followers on women's professional outcomes. In particular, a larger number of followers would not necessarily be as positive if follow-backs were driven by social instead of professional reasons. Recently, positive discrimination towards women has been documented in the context of elections to academic associations (Card et al., 2022a,b), and the hiring of software engineers (Finley, 2022). Thus, a result such as ours is not unprecedented in academic or professional contexts. Unfortunately, we cannot elicit the motives behind each follow, and therefore cannot determine which type of mechanism dominates in our setting.

Our findings have important implications for the debate over discrimination and representation in academia. First, we highlight that Twitter is an increasingly relevant platform for networking within academia. It is also generally perceived as being able to create a horizontal and democratic environment, eliminating barriers to networking that would be faced in other contexts. However, we show that some groups may experience discrimination when forming networks on the platform. Thus, though many perceive Twitter as egalitarian, broader disparities may, in fact, persist. Given the potential academic benefits of Twitter

(such as increasing paper citations or fostering collaborations), the type of discrimination we document could lead to worse professional outcomes for researchers who experience discrimination on the platform.[3] Second, our experiment contributes to the growing debate about diversity and inclusion within the economics profession (e.g., Bayer and Rouse, 2016; Bayer et al., 2020; Allgood et al., 2019) by providing experimental evidence of discrimination in a setting that is relevant to individuals' career success. Moreover, since academic economists are an active community on Twitter, a considerable portion of this debate has taken place on Twitter itself, with many economists voicing their concern about the profession's lack of diversity through posts on the platform. Our experiment highlights the extent to which these concerns are translated into action in a context in which action has a low cost: following a PhD student from an underrepresented group or a relatively low-ranked university.

This paper contributes to three strands of the literature. First, since we aim to measure discrimination experimentally, our work relates to the audit and correspondence studies literatures (Pager, 2007). Most experiments in this tradition are designed to assess employment discrimination, either by using actors in real-life situations (Pager et al., 2009), or mail-in resumes that include group-specific names or other textual information to generate experimental variation — a strategy used in the seminal work of Bertrand and Mullainathan (2004) and, more recently, by Kline et al. (2021).[4] A few audit studies have been conducted in the context of academia, though their focus has never been academic networks — specifically, Milkman et al. (2012, 2015) conduct an experiment by sending e-mails to university professors to analyze pre-admission discrimination of prospective graduate students, while Baker et al. (2022) document discrimination on the part of instructors in responses to students' forum posts in online classes. We contribute to the audit literature in several ways. Ours is among the first experiments to study discrimination in the formation of networks in academia, and we do so in a real-world, natural environment. Moreover, we study discrimination in a context considerably different from traditional labor markets, namely the formation of networks. Third, while most audit studies are restricted to studying low-wage markets, which generally involve discrimination against relatively low-skilled workers, our setting involves discrimination against PhD students. Therefore, we can verify whether discrimination is also pervasive within this high-skilled group. Furthermore, we explore the intersectionality between affiliation, gender, and race.

Second, this study relates to the growing literature on disparities within the economics

---

[3]This does not imply that the aggregate effect of Twitter on diversity is negative. It could be that researchers who experience discrimination on the platform would face even larger barriers in the profession in the absence of Twitter. Our results only imply that, even in this social media setting, discrimination may still be present and some people face greater barriers when forming their networks.

[4]There are audit studies of discrimination in several other contexts. For a review, see Bertrand and Duflo (2017). In particular, some scholars have examined the impact of college credentials on labor market outcomes, which is related to our analysis of the effect of university affiliation. For instance, Deming et al. (2016) assesses employers' perception of the value of a post-secondary degree obtained online or in person, while Gaddis (2015) documents the interaction between race and undergraduate university credentials in employment decisions. Our study is also related to experiments that use photos (AI-generated or not) instead of text to signal a dimension that is being manipulated. This strategy, while less common, has been used by researchers seeking to impart information that cannot easily be conveyed textually, such as physical appearance (Rooth, 2009), ethnicity (Arceo-Gomez and Campos-Vazquez, 2014; Lancee, 2021), or religiosity, the latter in the case of veiled and unveiled Muslim women (Fernández-Reino et al., 2022).

profession. An extensive body of work documents the ways in which women in academia are disproportionately affected by different contingencies, such as the COVID-19 pandemic (Deryugina et al., 2021), or treated worse than men in a variety of contexts within the profession, such as the peer-review process (Hengel, Forthcoming; Card et al., 2020), elections to professional associations (Card et al., 2022a,b), representation in conferences (Chari and Goldsmith-Pinkham, 2018), academic seminars (Dupas et al., 2021), recommendation letters (Eberhardt et al., 2022), allocation to undesired tasks such as serving on committees (Babcock et al., 2017), recognition for co-authored work (Sarsons et al., 2021), and in anonymous web-forums (Wu, 2020, 2018). Among these papers, that most directly related to ours is Wu (2020, 2018), who shows that discussions about women in the anonymous internet forum Economic Job Market Rumors tend to emphasize personal characteristics rather than professional accomplishments. Twitter is an interesting comparison point to this forum in that interactions are mostly not anonymous (although it is possible to create anonymous accounts, most users disclose their identity) and the set of economists who use Twitter may be different from those who use the forum. More generally, we contribute to this literature by studying gender discrimination in a novel context.

Compared to the evidence on gender disparities, that on racial disparities and elitism in economics are scarcer. That said, the under-representation of non-whites in the profession has been clearly documented. For instance, less than 10% of PhD degrees in economics awarded to US residents are to Black, Native-American, or people of Latin American origin (Bayer et al., 2020), a number lower than in other disciplines (Bayer and Rouse, 2016). Nevertheless, little work has looked at racial disparities in other aspects of the profession and the mechanisms that explain the latter — such as discrimination. Our experiment provides evidence in both of these areas. Furthermore, few papers have studied the effects of credentials or university affiliation in academia. While some scholars document that economists from top-ranked universities have historically been favored to receive distinguished prizes or election to associations (Cherrier and Svorenčík, 2020; Hoover and Svorenčík, 2020; Card et al., 2022a), it is unclear whether this happens due to academic merit or privilege related to better networks (Hoover and Svorenčík, 2020). Few papers have tried to disentangle these two explanations. One exception is Huber et al. (2022), who show that reviewers who see that a Nobel-laureate researcher authored a paper are more likely to recommend publication than those who see that an early-career researcher authored the same paper. We contribute to this discussion by documenting discrimination on the basis of credentials in the formation of academic networks, shedding light on elitism within academia.

Finally, our experiment is related to the growing literature on experiments on social media and other online platforms.[5] One approach in this literature has been to perform a treatment (for example, sending private or public messages) and see how this affects the subsequent behavior of treated units. This methodology has been used to study news consumption and misinformation (Levy, 2021; Pennycook et al., 2021), reactions to sanctions (Munger, 2017), or content moderation (Jiménez Durán, 2022). Bursztyn et al. (2022) exploit the willingness to send a tweet to study what motivates people to express dissenting opinions. A strategy closer to ours consists of creating fictitious accounts and examining users' interactions with

---

[5]For comprehensive reviews, see Mosleh et al. (2022) and Guess (2021).

these accounts. Mosleh et al. (2021) use this method to show that Twitter users are likelier to form social ties with accounts that share their political partisanship. Similarly, Bohren et al. (2019), Edelman et al. (2017) and Doleac and Stein (2013) test for discrimination on an online Q&A forum, an online house rental platform and an online advertisement website, respectively. Our experiment thus builds on the literature on field experiments on social media and other online platforms to analyze discrimination in the formation of networks. We contribute by using the methodology put forth by Mosleh et al. (2021) to study, for the first time, discrimination in the formation of networks in a social media environment.

The rest of this paper is organized as follows. In the next section, we provide background on Twitter and diversity within the economics profession, as well as some observational evidence from our sample of accounts on *#EconTwitter* to contextualize the disparities faced by economists in this environment. In Section 3, we describe the experimental design and then, in Section 4, discuss our main results. In Section 5, we present the heterogeneity results and finally, in Section 6, we consider potential concerns about experimental validity.

2   Background

2.1   Twitter

The setting of this experiment is Twitter, one of the most popular social media platforms among academics (Lupton, 2014). Twitter is a micro-blogging platform where users can share content in brief posts (*tweets*) of up to 280 characters. On Twitter, it is common to use hashtags — short expressions beginning with the symbol $\#$ — to signal a post's topic. These hashtags allow users to easily find others tweeting about their topics of interest. Users can also *retweet* or *like* posts from others, amplifying the content by making it visible to their followers.

On Twitter, most users have public profiles, which means that their posts are publicly visible. Although it is also possible to have *protected* — i.e., private — accounts, the default configuration is for an account to be public. Users can connect via follows, which do not need to be reciprocated, unlike other social media platforms such as Facebook. Indeed, to follow a public account, a user merely needs to click "follow" on the account's profile page. Once an account has been followed, the user who has been followed receives a *follow notification* on their account, informing them that a new account is following their profile.[6] This notification shows the follower's profile, at which point the user who has been followed may decide to follow that account back, do nothing, or block it. Once someone follows another account, that account's new tweets, retweets, and likes may appear on the follower's *timeline* (Twitter's main page).

Finally, each user with a public account has a profile page that is visible to all other users, which generally includes a profile picture, a background picture, and a short description

---

[6]Twitter sends this follow notification in most cases, though occasionally this does not happen. Twitter may deem an account to be acting suspiciously and *shadow-ban* it by making it invisible to other users. In this case, a followed user would not receive a follow notification.

(called a *bio*) provided by the user. The profile page also shows the tweets that the user has published and metrics on their use of the platform such as the number of tweets, followers, and friends (the profiles the user follows).

Possibly due to Twitter's structure of short and public posts, making it easy to dynamically share ideas with a large audience, it has been widely adopted by academics as their preferred social media platform for professional use. This is noticeable in at least three ways. First, in an international survey of academics, Lupton (2014) finds that 90% of respondents report using Twitter in their current professional work (the next most popular social media platform, LinkedIn, was used by only 60% of respondents). Although this survey is slightly outdated and has a sample biased towards social media users, its results suggest that Twitter is indeed extremely popular among academics. Second, many academics advise others to join Twitter. Cheplygina et al. (2020), for instance, write that "using Twitter appropriately can be more than just a social media activity; it can be a real career incubator in which researchers can develop their professional circles, launch new research projects and get helped by the community at various stages of the projects." Lee (2019), Rust (2019), Cherrier (2018), and many others give similar advice. Moreover, conferences such as the AEA Annual Meeting have featured panels on how to use Twitter (e.g., Wolfers, 2015), further suggesting that academics perceive the platform as being relevant for networking. Third, focusing on the academic economics community on Twitter, we found that over 14,000 different accounts tweeted or retweeted at least one post containing the hashtag *#EconTwitter* — which academic economists typically use — between January and February 2022. Beyond Twitter's popularity among academics, there is experimental evidence that researchers might actually benefit from a Twitter presence; for instance, Luc et al. (2021) show that tweeting about a paper considerably increases its citation count one year after the tweet. All of this points to the increasing importance of Twitter to the academic profession.

2.2 Diversity in Economics

Given the growing importance of Twitter to the formation of academic networks, one important question is to what extent Twitter users from the academic community exhibit discrimination in their decision of whom to follow on this social media. We focus our analysis on the academic economics community, a profession that has recently seen increasingly active debate over its lack of diversity, much of which has taken place on Twitter itself.

The American Economic Association recently conducted a Climate Survey among current and former members of the association to assess the *status quo* in the discipline, with particular attention to aspects that limit inclusiveness in the profession (Allgood et al., 2019). The survey's results suggest that experiences differ depending on an individual's gender or race. With respect to gender, men were two times more likely than women to report being satisfied with the overall climate in economics; furthermore, women reported facing considerably more discrimination or unfair treatment in academia than men. In terms of race, Black and non-Black economists declared widely divergent experiences: nearly half (47%) of Black economists reported being discriminated against or treated unfairly in the profession based on their race, against only 4% of their White counterparts. Finally, while the survey

did not cover the issue of elitism in the profession, this issue came up frequently in responses to open-ended questions as an additional dimension of the climate problem in economics. In assessing these answers, Allgood et al. (2019) point out that at least 250 comments addressed the topic of elitism, mostly claiming that "the profession is controlled by economists from the top institutions."[7]

The perceptions expressed in the survey, particularly those regarding gender discrimination, are echoed in an extensive body of literature investigating disparities within economics. However, the evidence on discrimination based on race or university affiliation in the profession is considerably scarcer, a gap that our experiment aims to address.

Interestingly, some of the gender, racial, and institutional disparities noted by the survey respondents are reproduced in the observational data from *#EconTwitter*. Using our sample of *#EconTwitter* users,[8] we are able to analyze the observational distribution of followers across different groups of Twitter users from this community. Among the users whose gender we could classify (approximately 60% of the sample), 27% are women. Among academics, 9% claim to be affiliated with a top-ten US university. Finally, among users whose perceived race or ethnicity we classified, 80.45% are White, while 7.94% are Black and 5% are Asian.

Figure 1 shows the distribution of the number of followers per account for men and women and for users affiliated with top-ten universities or not (conditional on being an academic). There seem to be no significant differences between men and women in terms of the distribution of the number of followers. However, when it comes to users' race or ethnicity, we observe a significant difference in the number of followers in favor of those classified as White. Figure 1b show that the distribution of followers (in logs) for White members of *#EconTwitter* is shifted to the right compared to the distribution for non-White members of this community. In particular, the median number of followers of White users is 582, against only 370.5 for non-White users. Finally, when it comes to university affiliation, the difference in favor of users from top-ten institutions is striking: the median number of followers for those affiliated with top-ten institutions is 1,224, against only 558 for academics from other institutions. These differences suggest that university affiliation and race or ethnicity matter to the formation of networks on this platform. However, many factors may explain the differences (or lack thereof) in the number of followers between these groups, such that these results alone are not evidence of discrimination. Our experiment allows us to identify whether discrimination exists in these dimensions (and in which direction), since

---

[7]To illustrate, we quote some of the responses in the Climate Survey below:

- *"Those outside the top ten tend to be discounted, dismissed, and not taken seriously."*

- *"(…) My impression is that, in economics, EVERYONE other than the top-ranked people are made to feel weak and excluded to some degree, so the remedy may be to do things that help everyone feel more included.(…)"*

- *"Discrimination in economics is based on topic of research, membership in the 'Top 5' club, and having a PhD from an exclusive set of universities.(…)"*

[8]This sample is composed of the universe of public accounts that tweeted or retweeted a post containing the term *#EconTwiter* between January and February 2022. We discuss the procedure used to create this sample and to obtain users' characteristics in Section 3 and Table B.2 in the Appendix.

the accounts created are identical in all dimensions except those we study, as explained in the following section.

3 Design and Data

3.1 Experimental Design

We conducted our pre-registered experiment (AEA RCT Registry ID AEARCTR-0009507) on Twitter between May and August 2022. Specifically, we created fictitious accounts (called 'bot' accounts) that claimed to be PhD students interested in economics-related topics. The accounts differed in their gender (male or female), race (Black or white), and university affiliation (top-ranked or lower-ranked institution). We signaled race and gender using images generated by artificial intelligence – in particular, race was signaled by skin color, although we acknowledge that race is a social construct that encompasses several dimensions. The account's university affiliation was signaled by a university's name in its Twitter bio. In the next section, we describe in detail the process of creating the bot accounts.

Since we varied three characteristics into two groups each, there are eight account types (treatment arms) in the experiment. We ran the experiment in twelve-day waves. In each wave, we activated one bot of each type. At the beginning of the wave, each bot randomly followed approximately 100 Twitter users who were identified as part of the *#EconTwitter* community — we describe the procedure for identifying this population of users below. At the end of the active period, we identified which of the treated accounts had followed the bots back, our main outcome of interest in the experiment. We also counted the total number of followers each bot had obtained, since other Twitter users may have organically decided to follow one or more of the accounts. We use this data to validate our findings in Section 6.

3.1.1 Bot Accounts

One of this experiment's main challenges lies in creating credible Twitter profiles. In principle, we wanted our profiles to be a good representation of a student at the beginning of their PhD program (since students in the final years of a PhD tend to be relatively well-known in the economics community and generally have their own website, which would make it more difficult to construct a believable profile). Table B.1 in the Appendix shows the typical characteristics of real Twitter profiles of first and second-year PhD students in economics. The bot accounts we created are not far from the median account, though they were more recent and had fewer tweets. For instance, most real accounts we found do not include a website or a public location. Moreover, any concerns relating to the plausibility of the experimental accounts are eased by the fact that the experiment had a take-up rate of 18.7% (take-up was fairly constant over the experimental waves we conducted, ranging from 21% to 14.4% and with no perceivable trend, as shown in Appendix Figure B.3).

Given these facts, we believe that the profiles built for the bot accounts are credible. Figure 2 provides examples of experimental accounts, while Table 1 describes the elements

of the fictitious profiles we created.

First, in terms of profile pictures, we used AI-generated images from Generated Media Inc. This tool makes it possible to create human-like portraits by controlling the image's attributes (specifically, in our case, gender, head pose, age, emotion, skin tone, hair color and length, and whether or not the avatar was wearing glasses or makeup). One challenge in using photos is that they may convey some other information, omitted to the researcher, and therefore not fully allow isolating the effect of the characteristics that were intentionally manipulated (Rich, 2018). Using AI-generated pictures means that we can minimize this concern by keeping all image attributes constant and only varying the desired ones. When creating the experimental profile pictures, we started from a random image in the correct age cohort. We then varied only the gender or skin tone attributes while keeping all other attributes constant, allowing us to construct sets of four images with the same "base". This procedure significantly reduces the concerns related to differences in dimensions other than those we are interested in here (race and gender).

Overall, we generated ten sets of four images. In each experimental wave, we randomly chose two sets (one for bots affiliated with top-ranked universities and another for ones affiliated with lower-ranked universities).[9] All images, as well as the waves in which they were used, are displayed in Appendix Figure A.1.

Apart from the picture, all profiles had a background image (set to the landscape of the city in which they claimed to be doing their PhD). The profiles did not have a website and did not include a location. As discussed above and shown in Appendix Table B.1, this is similar to the average profile of a first- or second-year PhD student. We also asked a group of approximately 30 economists and students with Twitter accounts to follow the profiles so that all of them had a certain number of followers from the start. All this helped make the profiles more credible.

Regarding bots' names, we created a list of common names and surnames based on those most prevalent in the 2000 US Census, excluding those that are race or ethnicity-specific (i.e., Hispanic names and first names that are disproportionately more likely to be used by White or Black people) and gender-neutral names.[10] We did not use racialized names to avoid concerns related to heterogeneity in the perceptions of race and because some names could be correlated with other characteristics such as socio-economic status (Gaddis, 2017; Fryer and Levitt, 2004). The bot's name was randomly generated by matching a first and last name from this list. The bot's bio informs its university affiliation and field of interest, held constant across all bots from the same wave. Finally, at the beginning of the experimental waves (before treatment), all active accounts randomly retweeted posts from the accounts of academic journals in economics. The objective of these two strategies was to increase the salience of the signal that the accounts belonged to academic economists and to make them

---

[9]At first, this randomization was planned in order to balance the number of times each set of images was used to represent a lower-ranked and a top-ranked university. However, given that we stopped the experiment before running all initially planned waves (see discussion in Section 3.1.3), this balancing was imperfect.

[10]Specifically, we used the NamSor tool to predict the gender of the names on our list and excluded those with less than 90% accuracy in gender prediction. To define race-specific names, we used data from Tzioumis (2018).

more realistic (it would be unusual for an account not to have any tweets or information on interests).

The fictional bot accounts differed in three dimensions: gender (male or female), race (signaled by black or white skin color), and university affiliation (highly ranked — or "elite" — versus lower ranked). The artificially generated profile image signaled the bot's race and gender, while the account's bio indicated university affiliation. To select the universities used in our experiment, we first considered the ten highest-ranked universities in the 2017 USNews Ranking of universities' economics graduate programs, along with the universities ranked between positions 79 and 100[11] that make their list of students publicly available.[12] To avoid concerns related to exposing specific universities, we randomly selected five high-ranked and five lower-ranked universities from this set of 20 universities for use in the experiment. Then, for each experimental wave, we randomly selected one of the five highly-ranked and one of the five lower-ranked universities.

3.1.2   Sample Selection and Assignment into Treatment

The study focuses on the academic economics community on Twitter, which commonly uses *#EconTwitter* to communicate. Therefore, accounts belonging to the community of users that employ this hashtag represent a natural subject pool for the experiment. Using Twitter's API, we obtained a dataset of all accounts that tweeted or retweeted messages containing the term *#EconTwitter* in January and February 2022. We restricted our sample to unprotected ("public") accounts. Moreover, since we wanted to maximize the chances that the subject accounts would interact with our bot accounts, our pre-analysis protocol also excluded all accounts with a follows/friends ratio above 15 (the approximate ratio of the 95th percentile of our subject pool) from the sample, along with profiles with fewer than ten followers. This step was intended to exclude institutional accounts, which generally have many followers but follow only a few users, as well as profiles that are overly selective in their choice of whom to follow. We also manually identified institutional accounts and bots and removed them from the sample, ultimately ending up with a sample of 10,226 subjects. See Appendix A.1 for a complete description of how we constructed our subject pool.

We obtained a set of variables for each subject using Twitter's API. Specifically, we have information on the number of tweets, followers, and friends. We also have information on location for the accounts that choose to make this information public, which we recode to the regional level. Furthermore, we know whether the account is verified, the number of likes ("favorites") it performed, and its date of creation. From the subjects' Twitter bio, we can also infer more specific information: we created a dummy variable equal to one if the bio

---

[11]This ranking is highly correlated with both the IDEAS/RePEc and the Tilburg university rankings, but has some advantages: first, it is a ranking of exclusively US institutions, and focuses on universities without differentiating between specific departments. Second, the methodology is based on a survey of academics in peer institutions, so more accurately represents the perceptions of the universities held by academics themselves (by contrast, the other two rankings are based on citations and publications).

[12]We restrict our analysis to universities that make their list of PhD students publicly available because the practice is common to all highly ranked universities. Therefore, using lower-ranked universities that do not maintain a public list of students could bias our results.

contains the name of a highly ranked university; we also created indicator variables for the user's occupation. Finally, we used the user's first name to predict their gender. The details of how these variables were defined are provided in Appendix Table B.2.

Tables 2 and 3 display descriptive statistics for the subjects. The data is interesting in and of itself as it paints a picture of who belongs to the *#EconTwitter* community. Overall, the sample is mainly comprised of men (73% of the accounts we could classify), and economists from Europe and North America (US or Canada). We were able to identify the profession of roughly 60% of the sample; of these, approximately one-third are professors, while 13.61% are PhD Students. On average, an account from our sample follows 1,245 accounts (with a median figure of 469), is followed by almost 4,000 accounts (median = 644), and has tweeted 22,000 times (median = 2,559). Figure B.1 shows the year each subject created their account. A large share of accounts was created between 2009 and 2011, and the number of new accounts started growing again in 2020.

Treatment assignment was performed for each wave. Following the suggestion of Athey and Imbens (2017), we performed block randomization as a way to improve balance, using the following variables: gender (male, female, missing); profession (professor, graduate student, other, missing); number of followers (above or below median). This gave us 24 strata. We sampled randomly from within each stratum, assigning the same proportion of users in each stratum to each bot account. Specifically, each bot account was assigned approximately 100 accounts to follow.[13] Misfits were reassigned globally by creating a "misfits stratum" and sampling from there (see Carril, 2017). The bot accounts always followed the designated accounts on a Thursday (see the timeline in Appendix table A.1). We followed subjects manually to minimize the chance that Twitter would consider the accounts' behavior to be suspicious. In each wave, we also randomized the order in which we created the bots and followed the subjects, thus eliminating the concern that a specific treatment type has a timing advantage.

Once a treated user is followed by a bot account, they receive a follow notification on Twitter. Figure 3 illustrates such a notification. The way a user sees the notification depends on whether he or she is using Twitter from the mobile or iPad applications or from a desktop computer. In all cases, the user immediately sees the bot's photo. In the mobile or iPad apps, the user also sees the description (which indicates the bot's university and research interests). The user only sees this description on a computer if they click on (or hover the mouse cursor over) the profile. However, to follow back the account, every desktop user will inevitably need to either click on the profile or hover their cursor over it, thus seeing the bot's description and, therefore, its university affiliation. To get a sense of how our subjects use Twitter, we live-streamed their tweets over the course of October 2022. During this period, we collected most of the tweets and retweets sent from these accounts, as well as the source of the tweet (e.g., Twitter's mobile or desktop app). The median number of tweets each user sent in this period was 16. For each user, we computed the share of the tweets that were sent via the mobile or iPad app. On average, users in our sample sent 63% of their tweets and retweets via the mobile/iPad app (median = 81%); 38% of users sent *all* of their tweets from their cellphone/iPad, and 78% sent *at least one* tweet from the mobile or iPad

---

[13]There will be some variation in this number to reduce the number of misfits.

app. Hence, the users in our sample seem, on average, to use the mobile or iPad apps more frequently than the desktop app or other sources (such as third-party apps or automated tweets). We note that this measure is probably a lower bound of Twitter use via mobile, since we can only capture the source of tweets and retweets, not overall time spent on the platform.

Apart from following the experimentally assigned accounts, each bot account also followed one account of someone who knew about the experiment. This person then informed us whether or not they received a follow notification. The objective being to guarantee that the users who are followed are being notified of that fact.[14] As we determined in the pre-analysis plan, if an account was "shadow-banned", we dropped it from the analysis. This happened with just a single account out of the 80 experimental accounts we activated.

3.1.3 Timing

For each experimental wave, we activated eight accounts (one of each type). We describe the procedure used to activate the accounts in detail in Appendix C. Within each wave, we used the following timeline (illustrated in Appendix Table A.1):

(i) **Day 0:** Creation of accounts according to the procedures described in Table 1. The account retweets two posts from the accounts of academic journals in economics. These posts are chosen randomly among recent posts from these accounts that already have more than three retweets.

(ii) **Day 1:** Each bot account follows the users assigned to it.

(iii) **Day 13:** After twelve days of being active, we count the number of followers gained by each account and delete all information on the account (see Appendix C for an explanation of the specific procedure used).

Therefore, the experimental waves had a twelve-day span. Appendix Figure B.4 shows that this is enough time to capture all responses to a follow by a bot. Indeed, over 83% of follow-backs in the experiment took place within one day of the treatment (i.e., the experimental account following the subject), with over 90% and 95% of follow-backs occurring within 2 and 5 days of the treatment, respectively. Therefore, the twelve-day period is sufficient to observe all relevant behavioral responses to the treatment.

We originally planned to run 30 experimental waves between May 23rd 2022 and December 20th 2022. Under this approach, we would have followed over 20,000 accounts (with some subjects treated more than once), which would have provided many more observations than necessary for our main analysis to have sufficient power. We had decided to plan more rounds

---

[14]On Twitter, so-called "shadow-banning" is a relevant concern. This is a type of punishment that Twitter can deploy against users whose behavior on the platform seems suspicious. In practice, all activity from a shadow-banned user is "hidden" to other users, including notifications of follows. Therefore, this allowed us to verify that none of our bot accounts had been shadow-banned before using their results.

than needed to address any potential practical problems that might arise, such as Twitter blocking some of the accounts, and to increase power to conduct the heterogeneity analysis.

We stopped earlier because a Twitter user saw some of the accounts posted about the experiment on Twitter during the eleventh wave. This post compromised the continuity of the experiment because, from that moment on, many subjects would potentially know that the accounts were inauthentic. We were aware of the possibility of something like this happening and do not believe it threatens the results of the previous waves. In fact, we find it reassuring that such a post only happened after the experiment had been running for three months and we had already followed over eight thousand subjects. Moreover, while the Twitter user who posted about the experiment was in our subject pool, he or she had not yet been assigned to a treatment, indicating that suspicion regarding the experimental accounts did not come from treated subjects. In Section 6.1, we discuss why this and other potential threats to experimental validity are not of great concern.

3.2 Empirical Strategy

To estimate discrimination based on gender, race, and university affiliation in follow-back behavior, we restrict our sample to the experimentally assigned pairs of subjects and bots. Our outcome of interest is Y_{ijst}, an indicator equal to one if subject i from stratum s followed-back bot account j during wave t. Given that the treatment (i.e., the follows from bot accounts) was randomly assigned, the causal effects of the bot accounts' gender, race, and university ranking can be identified and estimated by comparing the probability of follow-backs between each group of bot accounts. Specifically, we estimate the following equation by OLS:

$$Y_{ijst} = \beta_1 \times GENDER_j + \beta_2 \times RACE_j + \beta_3 \times RANK_j + X_i^{'}\lambda + \delta_t + \theta_s + \phi_{st} + \varepsilon_{ijst} \quad (1)$$

where $RACE_j$ is a dummy variable equal to one if the bot account represents a Black PhD student; $GENDER_j$ is equal to one if the bot account represents a woman, and $RANK_j$ is equal to one if the bot account claims to be affiliated with a top-ranked university. We include wave, stratum, and wave × stratum fixed effects.[15] Finally, X_i is a vector of pre-treatment characteristics of subject i, which includes the following variables: continent (dummies for Europe, US/Canada, and other); profession (dummies for professor, grad student, industry/tech, and other); gender; year of account creation; indicators for affiliation with a top-10 university, having a background picture, following accounts addressing the lack of diversity in academia, and having a verified account; and the number of Twitter followers and friends. In all cases, if a variable is missing, we create an indicator for missing.

We are interested in β_1, β_2, and β_3, which we interpret as the average difference (in percentage points) in the follow-back rate due to gender, race, and university affiliation, respectively. We also report regression results that include the interactions between the

---

[15]We include stratum fixed effects following the suggestion of Bruhn and McKenzie (2009). We also include stratum × wave fixed effects to account for possible differences in the behavior of subjects from different strata at different moments in time. Moreover, it should be noted that the stratum fixed effects include a misfit dummy.

14

variables $GENDER_j$, $RACE_j$ and $RANK_j$. This additional specification allows us to study how the group-based characteristics interact in the decision to follow back the bots.

In all cases, the parameter we identify is a measure of "total discrimination" against a given group (Bohren et al., 2022). Indeed, in our context, a Twitter user i may make a different follow-back decision for different groups of bots for two reasons. On the one hand, this user may have a bias against a certain group (for instance, due to preferences or implicit biases). On the other hand, the user may be unbiased at first and willing to follow any type of account, but their decision depends on whether other users decided to follow bot j. In the latter case, if other users were biased in favor of a given group, user i may end up only following bots from that group even if *ex-ante* they would be indifferent between groups (conditional on followers). Bohren et al. (2022) call these two types of discrimination "direct" and "systemic," respectively. In our setting, we only identify the aggregate, or total, discrimination.[16] This is a policy-relevant parameter, as it represents the true wedge between groups, taking into account the systemic effect.

In terms of inference, we report standard errors clustered at the bot account level. We performed the simplest assessment proposed by Ferman (2022) to verify the reliability of our inference, given the number of clusters. We simulate our data under the null hypothesis of no treatment effects, using Bernoulli draws with a parameter equal to the average follow-back rate in the three pilots to input our outcome (follow-back). Reassuringly, we obtained a rejection rate of the null under a nominal significance level of 5% that was very close to 5%. We also report randomization inference p-values, computed under the sharp null hypothesis of no treatment effect across all potential subject-bot pairs. Formally, suppose we let $Y_i(g, r, u)$ be the potential follow-back decision of user i given that they were followed by a bot of gender $g \in \{\text{male, female}\}$, race $r \in \{\text{white, Black}\}$ and university affiliation $u \in \{\text{top-ranked, lower-ranked}\}$. In such a case, our sharp null is that $Y_i(g, r, u) = Y_i(g', r', u')$ for all i and all g, g', r, r' and u, u'. We compute this p-value by doing 1,000 permutations of treatment assignment, estimating our regressions under the null, storing the t-statistic of each permutation, and comparing it to the t-statistic of the regression using the actual data, as in Young (2019).

3.3 Balance and Attrition

Table B.3 in the Appendix shows baseline characteristics for participants in each treatment arm, i.e., that were followed by each type of bot account. The sample is balanced across all characteristics (for all variables, we cannot reject the null hypothesis of equality of means across all treatments in a joint F test).

Attrition is not a significant problem in our context. Indeed, attrition could only happen in our experiment if we were unable to find accounts in our subject sample at the moment of treatment. The account would not be treated in this case, even if it was assigned to

---

[16]We note, however, that spillovers are unlikely to play a huge part in explaining our results, since most follow-backs happen within 24 hours of the follow (see figure B.4 on the Appendix). Thus, it is unlikely that those who follow-back the experimental accounts are taking into consideration the set of subjects who have already followed it.

treatment. This could happen if users deactivated their accounts, got suspended by Twitter, or if users chose to make their profiles private.[17] All of these cases were extremely rare throughout the experiment; the overall attrition rate was 2.12%. Still, Table B.4 in the Appendix shows that there was no differential attrition across treatment arms.

4   Results: Discrimination on the Basis of Gender, Race, and University Affiliation

Figure 4 plots the follow-back rate obtained by each type of fictional account over the ten experimental waves. We observe that the follow-back rate is indeed highly unequal across groups: while White women affiliated with top-ranked institutions were followed back 23.9% of the time, the average figure for Black men affiliated with relatively lower-ranked institutions was only 14.4%, a difference of almost 10 percentage points.

While Figure 4 is an interesting illustration of the trends obtained in the experiment, it is helpful to first analyze the marginal effect of each characteristic — gender, race, and university affiliation — on follow-backs separately. To do this, Figure 5 plots the individual follow-back rate for each manipulated characteristic. Implicitly, each panel of this figure reports the average (marginal) follow-back rate for each characteristic we are interested in (say, gender), across each combination of the other two characteristics (say, race and university affiliation). For each dimension we consider, we report the p-value of a simple t-test of the difference in means on top of the respective panel of Figure 5. Apart from this simple unconditional analysis, we estimate equation (1) – our pre-registered specification – which includes wave and stratum fixed effects, and report these results in Table 4 and in the last panel of Figure 5, both with and without the (also pre-registered) additional controls listed in Section 3.2. This panel plots the coefficients and 95% confidence intervals estimated using equation (1).

Both Figure 5 and Table 4 show that the rate of follow-backs is highly unequal across types of students. First, with respect to gender — the first panel of the figure — we see that women are considerably more likely to be followed back than men in the experiment. Indeed, 20.8% of follows by female bots were reciprocated, against only 16.6% of follows by male bots. This implies that, relative to men, women are 25% more likely to receive a follow-back. This difference, of around 4.3 percentage points, is highly statistically significant, as seen in the last panel of the figure.

The second panel of Figure 5 plots the follow-back rate for White and Black students. White students are significantly more likely to be followed back than Black students. Considering the raw numbers, the difference is 1.8 percentage points. The p-value of the simple difference in means test is less than 5%. By estimating equation (1), the difference in the follow-back rate between White and Black students is calculated as 2.2 percentage points, implying that Black students are 12% less likely to be followed back than their White peers.

---

[17]In this last case, we would still be able to find the account, but did not follow it to respect the person's choice of privacy, as agreed in our IRB.

This difference is significant at the 5% level when we do not add controls and at the 1% level with additional controls.

Finally, the third panel of Figure 5 plots the follow-back rate for students affiliated with top-ranked and relatively lower-ranked universities. We observe a clear preference for students of top-ranked institutions. The raw difference in the follow-back rate between students of top and lower-ranked universities is 3.7 percentage points. Our pre-registered specification leads to an estimated difference of 3.5 pp in the follow-back rate between students affiliated with top-ranked and lower-ranked universities, implying that students of elite institutions are almost 21% more likely than students of less prestigious institutions to be followed back by members of the *#EconTwitter* community.

Given that the experimental accounts were similar across all dimensions apart from the ones we studied — within each wave, they tweeted the same papers, had the same research interests, and had the same "base" image — the differences we found in the follow-back rate across groups are due to what we call "discrimination". Following Bertrand and Duflo (2017), we define discrimination as the differential treatment faced by one group relative to another when the two groups, apart from the group-based characteristic that distinguishes them, are identical. Our results show that, indeed, there is a differential treatment — in terms of follow-backs — between men and women, White and Black students, and students affiliated with top or lower-ranked universities, who are otherwise identical.

Therefore, our interpretation is that Black students and students from lower-ranked universities are discriminated against in the formation of their professional network on *#EconTwitter*. These two findings align with the evidence of disparities in the profession and with the perceptions of economists captured by the AEA's Climate Survey (Allgood et al., 2019). Our result suggests one potential cause for disparities in these two dimensions: since individuals from these groups are discriminated against when building their network, and professional networks matter for one's relative success in the profession, these individuals may have worse professional outcomes relative to similar students who are White or affiliated with top-ranked universities.

Surprisingly, we find that women are favored relative to men when forming their networks. This result is particularly striking since it goes in the opposite direction of the evidence, both in economics and elsewhere, of gender inequality. While we cannot elicit the motives behind follow-backs, different mechanisms could arguably underlie the higher follow-back rate for female bots. Some users may be conscious of the barriers faced by women in the profession and actively attempt to engage more with this group; however, it is also possible that some users were using Twitter with the objective of establishing relationships of a social instead of professional nature. Users could disproportionately want to establish ties of this nature with women. These differences in motivation are relevant because they could have different implications in terms of the type of network developed. If the first motive dominates, women would have, on average, a larger professional network than men, suggesting a mechanism through which gender disparities in economics could be slightly offset. However, if the second motive dominates, a larger network would not necessarily translate into better professional outcomes.

Table 4 also show results for a subsample of subjects that excludes those affiliated with

the universities used in each experimental wave. Specifically, for each wave, we exclude subjects who claim, in their Twitter bios, to be affiliated with one of the two universities used in that wave (recall that, in each wave, we randomly selected one top-ranked and one lower-ranked university to determine bot's affiliation). Analysing this subsample may be relevant because subjects from the same university as a bot may decide to follow it back due to considerations related to its affinity with the bot. The last two columns of Table 4 show that results are extremely similar when these subjects are excluded, indicating that this mechanism does not drive the results.

Thus far, we have only discussed marginal results, comparing the rate of follow-backs for each of the three dimensions we experimentally manipulated without considering how they interact. However, our experimental design also allows us to study whether the interactions of group-based characteristics lead to differential treatment. While it is possible that the differences we discussed previously are constant across each possible intersection of characteristics, these characteristics could also interact less obviously in terms of a subject's decision to follow an account or not.

We illustrate some of these possible intersectional effects in Figure 6, which focuses on the intersection between a bot's university affiliation and the other two dimensions (gender and race). The Figure plots the coefficient estimates of a regression analogous to equation (1) but including interactions between the bot's affiliation indicator and the other two indicators. This gives us an estimate of the difference in follow-back rate between male and female students, or between white and Black students, conditional on university affiliation.

Our results indicate that the racial gap in follow-backs is approximately the same regardless of whether the student is affiliated with a top or relatively lower-ranked university. Indeed, as can be seen in Figure 6, we estimate that Black students affiliated with lower-ranked universities are 2.3 pp less likely to be followed back than White students from the same group of universities. Similarly, Black students affiliated with top-ranked institutions are 2 pp less likely to be followed back than White students affiliated with the same type of institutions. The difference — of 0.3 percentage points — is quantitatively very small and not significant (p-value of 0.843). This result suggests that, in our setting, racial discrimination persists even in the presence of a signal — affiliation with a top-ranked university — that could be interpreted as being indicative of the student's high potential as an academic (and, therefore, their high "quality" as a Twitter friend).

We do, however, find evidence of a university premium for women. While female bots affiliated with lower-ranked universities are 2.8 pp more likely to be followed back than their male peers from the same institutions, women are 6 pp more likely than men to be followed back conditional on affiliation with a top-ranked institution. The difference — of 3.2 pp — is statistically significant. Hence, although women generally obtained a higher rate of follow-backs than men, this effect was accentuated among women from elite universities, even relative to men with similar institutional affiliations.

Finally, we do not obtain evidence of intersectionality between students' race and gender. We find that the racial gap in follow-backs is similar for both male and female bots. The follow-back rate for different combinations of bot gender, race, and affiliation can be seen in Figure 7, which plots the follow-back rate for each type of bot sorted by their characteristics.

5 Results: Heterogeneity

We pre-registered heterogeneity analyses in four dimensions: gender, Twitter profile reach, concern about the lack of diversity in economics, and racial or ethnic classification. In this section, we discuss each of these heterogeneities. Figure 8 provides graphical representations of these analyses, while the regression results are shown in Table 5.

The plots from Figure 8 report coefficient estimates and 95% confidence intervals for regressions of the form:

$$Y_{ijst} = \beta_1 \times GENDER_j + \beta_2 \times RACE_j + \beta_3 \times RANK_j +$$
$$\beta_4 \times (GENDER_j \times GROUP_i) + \beta_5 \times (RACE_j \times GROUP_i) +$$
$$\beta_6 \times (RANK_j \times GROUP_i) + \beta_7 \times GROUP_i + \delta_t + \theta_s + \phi_{st} + \varepsilon_{ijst} \quad (2)$$

where all variables used in equation (1) have the same definition as before, and $GROUP_i$ refers to an indicator of subject i's characteristic. The meaning of this variable will depend on the kind of heterogeneity we are studying (for instance, when considering the subject's gender, $GROUP_i$ is equal to one when a subject is categorized as female and 0 when categorized as male).

Parameters β_4, β_5, and β_6 represent the difference in the follow-back rates between the two groups of subjects (male and female, those with elite profiles or not, etc.) for each of the three bot's characteristics (gender, race, and university affiliation, respectively). For instance, a positive β_4 for the $GROUP$ of female subjects would imply that the follow-back differential between female and male bots is greater among female subjects than male ones. Each plot of Figure 8 refers to one dimension of heterogeneity across subjects (gender, account quality, concern about diversity, and race). For each dimension of heterogeneity, we plot the estimated level of discrimination for each set of subjects (male and female, those with stronger or weaker accounts, etc.), as well as the estimated difference in follow-back rate for each bot characteristic (gender, race and affiliation).

5.1 Subject's Gender

We predicted each subject's gender using their full name (which users themselves provide on Twitter).[18] Our procedure allowed us to classify each user as either male or female, with a precision level given by NamSor's algorithm. Naturally, this procedure is imperfect, especially considering that some Twitter users may not identify as male or female. We can also identify some of these users by checking whether or not they indicate their preferred pronouns in their bio. In our sample of 14,055 accounts, only 72 users stated that their preferred pronoun is "they/them." As this number is too small for any meaningful analysis of the differential behavior of non-binary users, we drop them from our heterogeneity analysis.[19]

---

[18]Specifically, we employed NamSor, a widely used name-checking technology, with an algorithm that provides insights into the origin and likely gender of a name. We only considered predictions with above 90% confidence, which allowed us to classify roughly 60% of our sample. Details are available in Appendix Table B.2.

[19]The results are extremely similar if we classify those users based on the gender predicted by their names.

Thus, we only consider the subsample of subjects whose gender we could confidently predict as either male or female based on their name, and whose profiles do not include non-binary pronouns.

Figure 8a shows the results of this heterogeneity analysis. Overall, we do not find any significant difference in the follow-back behavior of male or female subjects[20] with respect to the race or university affiliation of bots. However, there is some evidence that the difference in the follow-back rate between female and male bots is greater among male subjects than among female subjects. Specifically, we estimate that men are 5 percentage points more likely to follow-back female bots than male bots, while women are only 2.5 percentage points more likely to do so. Nevertheless, this difference between male and female subjects is not statistically significant (p-value of 0.28), possibly due to low power. Still, our results suggest that, although both male and female Twitter users favor female accounts in their follow-back decision, this gender disparity is greater among men than among women. This finding could help shed light on the potential mechanisms behind female bots receiving more follow-backs than male ones, as discussed in the previous section.

The differences in the behavior of men and women on Twitter is, however, a precisely estimated zero when considering the race or institutional affiliation of the bots. This means that men and women from the *#EconTwitter* community do not seem to have different follow-back behaviors when it comes to these characteristics.

5.2 Subject Profile's Reach

One important question related to the results in Section 4 is whether the discrimination observed in terms of race, university affiliation, and gender is displayed exclusively by some profiles with lower reach or if it is widespread in the *#EconTwitter* community. In particular, it would be interesting to know whether accounts with a relatively large number of followers — and that, therefore, potentially have higher impact in the *#EconTwitter* network — also discriminate in their decision to follow back. This would be particularly concerning because follow-backs from higher-impact accounts are arguably more important than those from profiles with lower reach. Indeed, one of the main benefits of gathering followers is that those followers might engage with one's posts, amplifying them to their follower base. Follow-backs from accounts with a greater following, therefore, increase a user's potential to reach a larger audience. In addition, being followed by high-reach profiles may increase a user's reputation, since lists of followers are public on Twitter.

We test whether well-known accounts behave similarly to other accounts by assessing the follow-back behavior of accounts above and below the median number of followers in our sample.[21] We consider profiles with above the median number of followers as relatively "strong" profiles, since they have higher potential impact on social media.

---

[20]More precisely, among subjects whose gender was predicted as male than among subjects whose gender was predicted as female. To avoid using this cumbersome construction, we will refer to subjects' predicted gender as their gender.

[21]The median number of followers among treated subjects is 344 followers.

One important aspect of this analysis is that it can shed light on whether the results obtained previously are meaningful in terms of their impact on the formation of networks. For example, if we found that higher-reach accounts do not discriminate, this would suggest that the potential effects of the main results from Section 4 are less intense — since being discriminated against by low-impact accounts would not matter as much to the formation of networks. However, our findings suggest this is not the case. As shown in Figure 8b, accounts with a number of followers above the median seem to exhibit a follow-back behavior similar to those with a below-median number of followers. For all three bot characteristics, we estimate that the difference in the follow-back behavior of high and low-reach accounts is indistinguishable from zero. The results are similar for other definitions of account reach, such as those with more than 1,000 followers or those included in more than four public lists — see Appendix Figure B.2. Therefore, we have evidence that high-reach accounts exhibit discriminatory behavior comparable to relatively lower-reach accounts, i.e., that the type of discriminatory behavior we described in Section 4 is not limited to a set of potentially less impactful accounts.

5.3   Subject's Concern about the Lack of Diversity in Economics

The lack of diversity in economics has recently been much debated, particularly on Twitter. One interesting question, therefore, is whether or not those who are publicly more engaged in this debate behave differently from the rest of the accounts in the *#EconTwitter* community. In particular, a relevant concern is that people may choose to outwardly demonstrate interest in the topic as a way of signaling virtue or due to self-image concerns (Bursztyn and Jensen, 2017), while not adopting a non-discriminatory behavior in actions that are less visible to others — such as deciding whether to follow a student on Twitter.

We created an indicator of "concern about the lack of diversity in economics," equal to one if a subject follows at least one Twitter account dedicated to this topic, such as the AEA Committee on Equity, Diversity, and Professional Conduct (the full list of accounts considered for this measure is provided in Appendix Table B.2). Approximately 15% of our sample publicly signals concern with the topic in this way. Figure 8c shows the follow-back rate for each type of bot account by whether or not the subject follows one of these accounts, thereby signaling (or not) their concern about the topic of diversity.

Overall, we find that those who seem concerned with diversity do seem to behave differently from their peers in some dimensions. First, their behavior towards male and female bots is similar: both groups of subjects are more likely to follow female accounts at a similar rate. To the extent that those who follow accounts related to diversity are concerned with the topic, this might indicate that the higher rate of follow-backs received by women is at least partially motivated by individuals who are aware of the barriers women face and are actively trying to engage more with this group.

However, subjects who publicly signal concern about diversity do seem to have different follow-back behaviors when it comes to the race of the experimental account. Specifically, subjects concerned with the lack of diversity in economics do not seem to discriminate by race. They follow back White and Black students at practically the same rates — considering the

raw data, they even slightly favor Black students, with a follow-back rate of 24.9% for White students against 25% for Black students. In contrast, we estimate that those who do not signal a concern with diversity are 2.3 percentage points more likely to follow-back a White student than a Black student. As seen in Figure 8c, the difference between the estimated racial discrimination of subjects who publicly signal their concern about diversity and those who do not is above 1 pp. However, we cannot reject the hypothesis that this difference is zero, possibly due to the low power for this test. Still, the evidence suggests that subjects who seem concerned about the lack of diversity in economics do not racially discriminate in their follow-back decisions — or at least that the rate at which they discriminate is quantitatively lower.

This result suggests that those who signal concern about diversity take some measure in this regard, at least concerning race. It should be noted, however, that the action we study — following an account of a PhD student — has a low cost and may therefore not represent a sufficient movement towards an increase in diversity. Nevertheless, we find that, at least when it comes to race, publicly signaling concern about diversity is translated into less discriminatory behavior, suggesting that this public signal is not exclusively motivated by social image concerns.

Finally, in contrast to the results obtained for the bot's race, we find that subjects who signal concern about the lack of diversity in economics are those who discriminate the most in terms of university affiliation. Indeed, while both groups of subjects favor students affiliated with top-ranked institutions, the difference in the follow-back rate for top- and lower-ranked students is considerably larger among subjects who exhibit this concern (9 pp in this group, against 2.5 pp among all other subjects). Hence, the difference in behavior between the two groups is large (above 6 pp) and statistically significant (p-value of 0.018). This finding suggests that although subjects concerned about diversity appear to consciously and actively avoid race-based discrimination, discrimination on the basis of institutional affiliation is prevalent and particularly strong among this group. Therefore, the problem of elitism in academia would seem to be a blind spot among those concerned with diversity and representation in this environment.

5.4   Subject's Race or Ethnicity

Finally, we study whether subjects' behavior is heterogeneous based on their perceived race or ethnicity. We manually analyze users' profile pictures to classify subjects' race or ethnicity. Therefore, what we measure is how Twitter users' race or ethnicity are perceived by others (as opposed to self-identification). Similar procedures have been previously employed to study Twitter data, as reviewed by Golder et al. (2022). Using this method, we classified 63% of our sample into the groups "White", "Black", "Asian" and other race or ethnicity.

Figure 8d displays the results of this heterogeneity analysis. We divide the sample of subjects into those classified as White or non-White, pooling all other races or ethnicities into this latter group (but results are qualitatively similar if we focus exclusively on the sample of users classified as White or Black). The main objective of this analysis is to verify whether non-White subjects have different behavior than White subjects regarding bots'

race. We find suggestive evidence that this is the case. While the difference in the racial gap between White and non-White subjects is not statistically significant (possibly due to the low power to perform this comparison), we estimate that the racial gap among non-White subjects is close to zero (point-estimate of -0.002, p-value of 96.8%). This estimate supports the claim that non-White subjects do not, on average, exhibit racial discrimination in their follow-back behavior. On the other hand, White subjects are significantly more likely to follow back White than Black students. Overall, this result is interesting as it shows that Twitter users who belong to racial groups that might suffer the type of discrimination we study in this paper do not, on average, replicate such discriminatory behavior against others belonging to similar groups.

We also find differences in the behavior of White and non-White subjects regarding bots' gender and university affiliation. First, non-White subjects are 9.6 pp more likely to follow female than male bots, while White subjects are only 3.5 pp more likely to follow female than male bots. This difference, while not statistically significant, is quantitatively large. Similarly, we obtain evidence that, differently from White subjects, non-White ones do not discriminate due to university affiliation. Taken together, these two results suggest that non-White subjects are generally more aware of the barriers faced by different groups in academia, and reflect this awareness in their follow-back behavior.

6 Understanding and Validating the Results

6.1 Total Follows and Concerns with Contamination by Twitter's Algorithm

In this section, we discuss some potential concerns related to the experiment's validity and show that they are unlikely to be relevant in our setting. An important concern with the experimental design has to do with Twitter's algorithm, which can suggest accounts for users to follow. If Twitter algorithm's suggestions were correlated with some of the bots' characteristics — for instance, if the algorithm was more likely to suggest accounts of top-ranked students —, analysing total follows would lead to biased estimates of the effects we study. Therefore, we restricted our main analysis to follow-backs, enabling us to isolate our outcome of interest from the effects of the algorithm. Indeed, our analysis only considers interactions with subjects who have received a notification from one of our accounts (i.e., subjects who are followed by one of our bots), not those who might casually find one of the experimental accounts due to algorithmic suggestions.

Still, we have considerable evidence that allows us to rule out any contamination by organic follows motivated by algorithm suggestions. First, the number of such follows received by the accounts was very low. Over the ten experimental waves, the bot accounts received a total of 116 organic follows. Given that we had 71 non-shadow banned accounts, this represents a total of 1.46 organic follows per account. This is extremely few — recall that each experimental account had, on average, 50 followers (including follow-backs by experimental subjects as well as by colleagues we asked to follow all accounts). Hence, organic followers

represent less than 3% of the bots' followers. Furthermore, because the group of around 30 colleagues we asked to follow the bot accounts was largely comprised of graduate and undergraduate students from the São Paulo School of Economics, many of the users who organically followed the bot accounts were individuals who had social media connections with this group of colleagues (either because they followed or were followed by them). Since most subjects are not closely connected with these colleagues, we do not believe that Twitter's algorithm suggested the experimental accounts to the subjects to any great degree. In addition, the distribution of organic follows across groups of subjects mirrors the distribution of follow-backs we discussed previously, as seen in Table B.5. The differences are small and not statistically significant across groups. However, women did receive, on average, more organic follows than men, White students received more than Black students, and students affiliated with top-ranked institutions received more than those affiliated with lower-ranked ones. Hence, we do not expect organic follows to have affected the experiment's results.

A related concern is that when people receive a notification from one of the experimental accounts, some of the other accounts could be suggested to them by the algorithm, making the subject suspicious about the veracity of the experimental account. Indeed, when a user accesses Twitter from the desktop app (but not from the mobile app) and navigates to another user's profile — which is possible, though not necessary for a follow-back — Twitter's "who to follow?" algorithm usually suggests another three accounts for the user to follow. These suggestions are placed on the right-hand side of the screen and are not particularly prominent on the page. On the mobile or iPad app, those suggestions can only be seen if the user scrolls down on the screen. Still, if the algorithm suggests other experimental accounts, this could raise suspicion among subjects. This would only be a problem in our setting if such misgivings — if they existed — were correlated with the bots' characteristics.

We do not have evidence that such suggestions happened frequently enough to affect our results.[22] First, if such suggestions were happening, they would have become more frequent over time, as Twitter's algorithm learned about the similarity of the experimental accounts. Figure B.3 shows that the uptake rate of the experiment was fairly constant across waves, indicating that there were likely no actions taken by the platform against the accounts. We see a similar consistency for each type of bot. There is some variation across waves, which is expected, but no trend over time. This suggests that our setting was not affected by the algorithm. Second, no subject from any wave followed an experimental account other than the one that followed him or her. We would expect this to happen if the algorithm was indeed suggesting other bot accounts to treated users. If this occurred, subjects might add other bot accounts, either without noticing the similarity to the account that had followed them

---

[22]Though, this did happen at least once, leading one user to tweet about our experiment, thereby forcing us to finish the experiment earlier than planned. While this user was in our subject pool, he or she had not been "treated" in the wave in which the post was made. The most likely reason for his or her suspicion is that the algorithm suggested one of the experimental accounts as a potentially interesting user to follow, along with other similar accounts. This further reduces any concern about contamination, since the only suspicion about the experiment came from someone who was not treated. Moreover, the fact that this only happened after we had run ten complete experimental waves (plus three pilot waves) is reassuring: if more people had been suspicious about the experimental accounts, it is likely that we would have seen more tweets on the topic or would perhaps have received direct messages through the bots' accounts. This did not happen even once.

or in an effort to actively spoil our results. Thus, it is reassuring that this did not happen even a single time. Moreover, we note that algorithmic suggestions are only highlighted in the Twitter Desktop App, and our results show that most of the Twitter users in our subject pool use the mobile or iPad app more frequently. Indeed, by analyzing the source of live-streamed tweets and retweets from our subjects, we find that the median user of our sample tweets from the mobile or iPad app 81% of the time, with 38% of users exclusively using the mobile app. Note that this measure probably underestimates Twitter use via cellphones and tablets, since we can only capture the source of tweets and retweets, not overall time spent on the site. Overall, all of these evidence suggest that contamination due to Twitter's algorithm is not a major concern in our setting.

Conclusion

The lack of diversity in economics (and academia more broadly) is increasingly debated. Discrimination is often pointed out as one of the reasons for such a lack of representation of certain social groups. We conducted a field experiment on Twitter to study discrimination in the formation of social networks on a social media platform. We show that members of the *#EconTwitter* community are less likely to follow accounts of Black students and of students affiliated with relatively lower-ranked universities. They are, however, more likely to follow the accounts of female students.

This paper provides evidence that discrimination is indeed present in the formation of networks on social media, as students with otherwise identical characteristics are treated differently based on their race, gender, or university affiliation. In particular, we find that the racial gap in follow-backs occurs independently of university affiliation, i.e., racial discrimination against Black students occurs even for students affiliated with a prestigious university. Interestingly, economists who seem concerned with the lack of diversity in the profession do not exhibit racial discrimination, but do disproportionately discriminate against lower-ranked students. This suggests that elitism, possibly a less salient dimension in the diversity debate, may be a blind spot for this group.

Our results highlight that discrimination is present even on Twitter, a social media platform usually regarded as egalitarian. The findings do not, however, speak to the aggregate effects of Twitter on diversity. Counterfactually, it is possible that in a world without platforms such as Twitter, the groups that experience discrimination on the site would have even more difficulty building their networks. Still, while social media may have reduced the overall barriers faced in academia by individuals from under-represented groups or with less prestigious affiliations, our results indicate that those barriers are still present even in this setting. Further research is needed to fully understand academic Twitter's aggregate effects on welfare and the profession's diversity relative to a counterfactual world in which such platform did not exist.

This paper also has important implications for the debate on representation and discrimination in academia. We are among the first to provide evidence of discrimination in the formation of networks in academia. While there are many papers documenting disparities

in several aspects of academia, few studies have been able to elucidate the mechanisms behind those disparities. By collecting descriptive statistics on the *#EconTwitter* community, we show that academics affiliated with less prestigious universities have a lower number of followers than those affiliated with top-ranked institutions. Meanwhile, our experiment documents that discrimination is present in the context of network formation within the economics academic community, shedding light on a potential mechanism to explain some of the disparities that have been extensively documented in the profession and within *#Econ-Twitter*. Understanding that discrimination plays a role in causing these disparities may be an initial step in designing policies to reduce them. Furthermore, we note that our primary outcome — follow-backs — represents a relatively low-cost action; hence, our results may constitute a lower bound for discrimination in other contexts within academia.

Moreover, since social media — and Twitter in particular — can be a powerful networking tool for those who would not otherwise have access to solid academic networks, documenting discrimination in this setting is particularly relevant. Indeed, Twitter has the potential of having positive impacts for academics (for instance, leading to increases in a paper's citations), and is perceived by many academics as being instrumental for success in the profession (as demonstrated by the fact that many academic conferences have featured panels on using Twitter). However, we show that the barriers to networking on the platform are unequal across groups, which may reduce the platform's potential impact for those who suffer discrimination.

This paper does have certain limitations. First, we cannot elicit the motivation behind follow-backs, which would be particularly helpful for interpreting the results related to gender. In aggregate, we find that women are favored relative to men in follow-back decisions, though whether this difference in follow-backs is driven by Twitter users who are attempting to engage more with women to compensate for gender disparities, or by users seeking to establish relationships of a social rather than professional nature remains an open question. Second, by comparing follow-back rates across groups, we can only identify a measure of "total" discrimination, encompassing both follow-backs that happen due to "direct" discrimination and those that occur as spillovers of direct discrimination (also known as "systemic" discrimination). Though it would be interesting to decompose the effects we obtained between direct and systemic discrimination, "total" discrimination is nonetheless a policy-relevant parameter in this setting since it is a measure of the overall wedge due to discrimination between groups. Moreover, as most follow-backs happen less than 24 hours after treatment, spillovers are unlikely to drive the result, and it is possible that they would exacerbate the disparities if we scaled the experiment.

Finally, while we focused on the academic economics community, it would be interesting to explore how the discrimination exhibited by this group compares to that of other groups in academia and elsewhere. Another question begging further study is whether and to what extent other underrepresented groups suffer discrimination when forming their professional networks. The methodology applied in this paper could be used to study these and several other compelling and relevant issues.

References

Allgood, Sam, Lee Badgett, Amanda Bayer, Marianne Bertrand, Sandra E Black, Nick Bloom, and Lisa D Cook, "AEA Professional Climate Survey: Final Report," Technical Report, American Economic Association's Committee on Equity, Diversity and Professional Conduct 2019.

Arceo-Gomez, Eva O and Raymundo M Campos-Vazquez, "Race and marriage in the labor market: A discrimination correspondence study in a developing country," *AEA Papers and Proceedings*, 2014, *104* (5), 376–80.

Arcidiacono, Peter, Josh Kinsler, and Tyler Ransom, "Asian American discrimination in Harvard admissions," *European Economic Review*, 2022, *144*, 104079.

Athey, Susan and Guido W Imbens, "The econometrics of randomized experiments," in "Handbook of economic field experiments," Vol. 1, Elsevier, 2017, pp. 73–140.

Azoulay, Pierre, Joshua S Graff Zivin, and Jialan Wang, "Superstar extinction," *The Quarterly Journal of Economics*, 2010, *125* (2), 549–589.

Babcock, Linda, Maria P Recalde, Lise Vesterlund, and Laurie Weingart, "Gender differences in accepting and receiving requests for tasks with low promotability," *American Economic Review*, 2017, *107* (3), 714–47.

Baker, Rachel, Thomas Dee, Brent Evans, and June John, "Bias in online classes: Evidence from a field experiment," *Economics of Education Review*, 2022, *88*, 102259.

Bayer, Amanda and Cecilia Elena Rouse, "Diversity in the economics profession: A new attack on an old problem," *Journal of Economic Perspectives*, 2016, *30* (4), 221–42.

_ , **Gary A Hoover, and Ebonya Washington**, "How you can work to increase the presence and improve the experience of Black, Latinx, and Native American people in the economics profession," *Journal of Economic Perspectives*, 2020, *34* (3), 193–219.

Beaman, Lori, Niall Keleher, and Jeremy Magruder, "Do job networks disadvantage women? Evidence from a recruitment experiment in Malawi," *Journal of Labor Economics*, 2018, *36* (1), 121–157.

Bertrand, Marianne and Esther Duflo, "Field experiments on discrimination," *Handbook of economic field experiments*, 2017, *1*, 309–393.

_ **and Sendhil Mullainathan**, "Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination," *American economic review*, 2004, *94* (4), 991–1013.

Bohren, J Aislinn, Alex Imas, and Michael Rosenberg, "The dynamics of discrimination: Theory and evidence," *American economic review*, 2019, *109* (10), 3395–3436.

_ , **Peter Hull, and Alex Imas**, "Systemic Discrimination: Theory and Measurement," 2022.

Bruhn, Miriam and David McKenzie, "In pursuit of balance: Randomization in practice in development field experiments," *American economic journal: applied economics*, 2009, *1* (4), 200–232.

Bursztyn, Leonardo and Robert Jensen, "Social image and economic behavior in the field: Identifying, understanding, and shaping social pressure," *Annual Review of Economics*, 2017, *9*, 131–153.

_ , **Georgy Egorov, Ingar K Haaland, Aakaash Rao, and Christopher Roth**, "Justifying dissent," Technical Report, National Bureau of Economic Research 2022.

Card, David, Stefano DellaVigna, Patricia Funk, and Nagore Iriberri, "Are referees and editors in economics gender neutral?," *The Quarterly Journal of Economics*, 2020, *135* (1), 269–327.

_ , _ , _ , **and** _ , "Gender differences in peer recognition by economists," *Econometrica*, 2022.

_ , _ , _ , **and** _ , "Gender Gaps at the Academies," Technical Report, National Bureau of Economic Research 2022.

Carril, Alvaro, "Dealing with misfits in random treatment assignment," *The Stata Journal*, 2017, *17* (3), 652–667.

Chari, Anusha and Paul Goldsmith-Pinkham, "Gender representation in economics across topics and time: Evidence from the NBER summer institute," Technical Report, National Bureau of Economic Research 2018.

Cheplygina, Veronika, Felienne Hermans, Casper Albers, Natalia Bielczyk, and Ionica Smeets, "Ten simple rules for getting started on Twitter as a scientist," 2020.

Cherrier, Beatrice, "Why Historians of Economics Should Tweet," *History of Political Economy*, 2018, *50* (3), 615–621.

_ **and Andrej Svorenčík**, "Defining excellence: seventy years of the John Bates Clark Medal," *Journal of the History of Economic Thought*, 2020, *42* (2), 153–176.

Deming, David J, Noam Yuchtman, Amira Abulafi, Claudia Goldin, and Lawrence F Katz, "The value of postsecondary credentials in the labor market: An experimental study," *American Economic Review*, 2016, *106* (3), 778–806.

Deryugina, Tatyana, Olga Shurchkov, and Jenna Stearns, "COVID-19 disruptions disproportionately affect female academics," *AEA Papers and Proceedings*, 2021, *111*, 164–68.

Doleac, Jennifer L and Luke CD Stein, "The visible hand: Race and online market outcomes," *The Economic Journal*, 2013, *123* (572), F469–F492.

Ductor, Lorenzo and Bauke Visser, "When a coauthor joins an editorial board," *Journal of Economic Behavior & Organization*, 2022, *200*, 576–595.

_ , **Marcel Fafchamps, Sanjeev Goyal, and Marco J Van der Leij**, "Social networks and research output," *Review of Economics and Statistics*, 2014, *96* (5), 936–948.

Dupas, Pascaline, Alicia Sasser Modestino, Muriel Niederle, Justin Wolfers et al., "Gender and the dynamics of economics seminars," Technical Report, National Bureau of Economic Research 2021.

Eberhardt, Markus, Giovanni Facchini, and Valeria Rueda, "Gender Differences in Reference Letters: Evidence from the Economics Job Market," 2022.

Edelman, Benjamin, Michael Luca, and Dan Svirsky, "Racial discrimination in the sharing economy: Evidence from a field experiment," *American economic journal: applied economics*, 2017, *9* (2), 1–22.

Ferman, Bruno, "Assessing inference methods," *arXiv preprint arXiv:1912.08772*, 2022.

Fernández-Reino, Mariña, Valentina Di Stasio, and Susanne Veit, "Discrimination Unveiled: A Field Experiment on the Barriers Faced by Muslim Women in Germany, the Netherlands, and Spain," *European Sociological Review*, 2022.

Finley, Marley, "Girls Who Code: A Randomized Field Experiment on Gender-Based Hiring Discrimination," 2022.

Fryer, Roland G and Steven D Levitt, "The causes and consequences of distinctively black names," *The Quarterly Journal of Economics*, 2004, *119* (3), 767–805.

Gaddis, S Michael, "Discrimination in the credential society: An audit study of race and college selectivity in the labor market," *Social Forces*, 2015, *93* (4), 1451–1479.

_ , "How black are Lakisha and Jamal? Racial perceptions from names used in correspondence audit studies," *Sociological Science*, 2017, *4*, 469–489.

Golder, Su, Robin Stevens, Karen O'Connor, Richard James, and Graciela Gonzalez-Hernandez, "Methods to Establish Race or Ethnicity of Twitter Users: Scoping Review," *Journal of medical Internet research*, 2022, *24* (4), e35788.

Goyal, Sanjeev, Marco J Van Der Leij, and José Luis Moraga-González, "Economics: An emerging small world," *Journal of political economy*, 2006, *114* (2), 403–412.

Guess, Andrew M, "Experiments using social media data," *Advances in experimental political science*, 2021, *184*.

Hengel, Erin, "Publishing while female: Are women held to higher standards? Evidence from peer review," *Economic Journal*, Forthcoming.

Hoover, Kevin D and Andrej Svorenčík, "Who runs the AEA?," *Center for the History of Political Economy at Duke University Working Paper Series*, 2020, (2020-12).

Hridoy, Syed Akib Anwar, M Tahmid Ekram, Mohammad Samiul Islam, Faysal Ahmed, and Rashedur M Rahman, "Localized twitter opinion mining using sentiment analysis," *Decision Analytics*, 2015, *2* (1), 1–19.

Huber, Juergen, Sabiou Inoua, Rudolf Kerschbamer, Christian König-Kersting, Stefan Palan, and Vernon L Smith, "Nobel and novice: Author prominence affects peer review," *University of Graz, School of Business, Economics and Social Sciences Working Paper*, 2022.

Jackson, Matthew O, Brian W Rogers, and Yves Zenou, "The economic consequences of social-network structure," *Journal of Economic Literature*, 2017, *55* (1), 49–95.

Jiménez Durán, Rafael, "The economics of content moderation: Theory and experimental evidence from hate speech on Twitter," 2022.

Kline, Patrick M, Evan K Rose, and Christopher R Walters, "Systemic discrimination among large US employers," Technical Report, National Bureau of Economic Research 2021.

Lancee, Bram, "Ethnic discrimination in hiring: comparing groups across contexts. Results from a cross-national field experiment," 2021.

Lee, Jet-Sing M, "How to use Twitter to further your research career," *Nature*, 2019, *10.*

Levy, Ro'ee, "Social media, news consumption, and polarization: Evidence from a field experiment," *American economic review*, 2021, *111* (3), 831–70.

Luc, Jessica GY, Michael A Archer, Rakesh C Arora, Edward M Bender, Arie Blitz, David T Cooke, Tamara Ni Hlci, Biniam Kidane, Maral Ouzounian, Thomas K Varghese Jr et al., "Does tweeting improve citations? One-year results from the TSSMN prospective randomized trial," *The Annals of thoracic surgery*, 2021, *111* (1), 296–300.

Lundberg, Shelly and Jenna Stearns, "Women in economics: Stalled progress," *Journal of Economic Perspectives*, 2019, *33* (1), 3–22.

Lupton, Deborah, "'Feeling better connected': Academics' use of social media," Technical Report, News and Media Research Centre, University of Canberra 2014.

Milkman, Katherine L, Modupe Akinola, and Dolly Chugh, "Temporal distance and discrimination: An audit study in academia," *Psychological science*, 2012, *23* (7), 710–717.

_ , _ , **and** _ , "What happens before? A field experiment exploring how pay and representation differentially shape bias on the pathway into organizations.," *Journal of Applied Psychology*, 2015, *100* (6), 1678.

Mosleh, Mohsen, Cameron Martel, Dean Eckles, and David G Rand, "Shared partisanship dramatically increases social tie formation in a Twitter field experiment," *Proceedings of the National Academy of Sciences*, 2021, *118* (7).
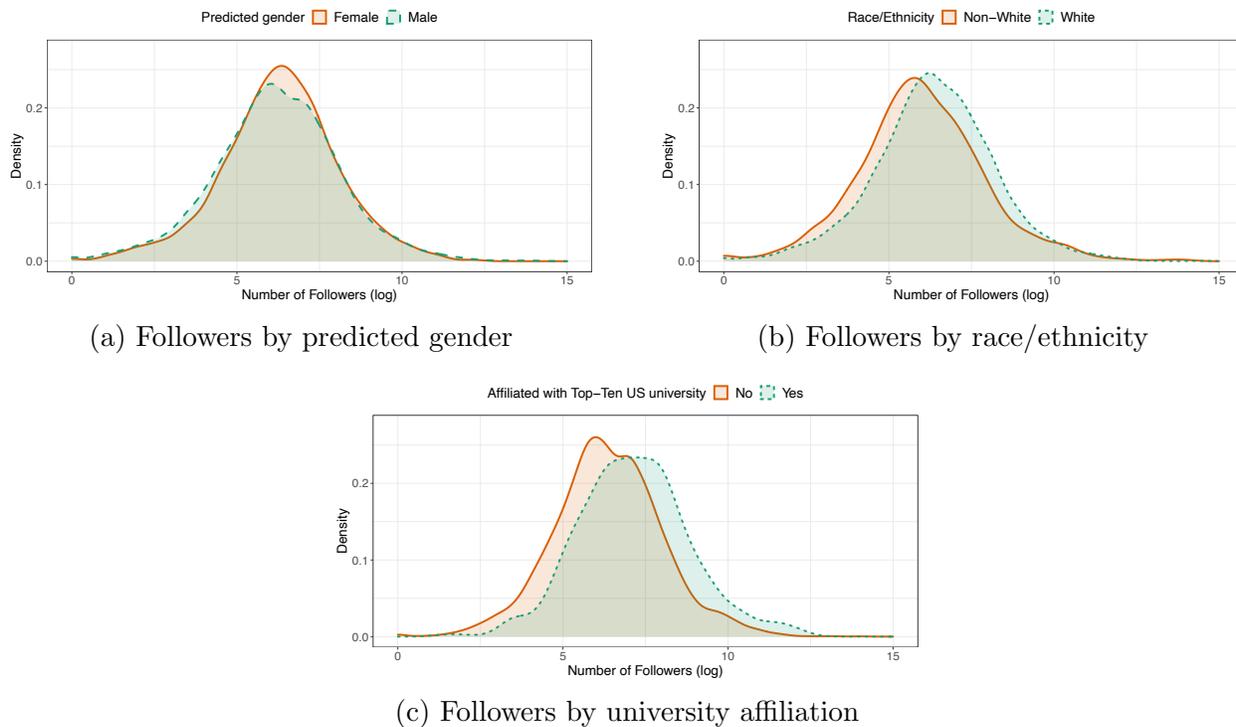
_ , **Gordon Pennycook, and David G. Rand**, "Field Experiments on Social Media," *Current Directions in Psychological Science*, 2022.

Munger, Kevin, "Tweetment effects on the tweeted: Experimentally reducing racist harassment," *Political Behavior*, 2017, *39* (3), 629–649.

Pager, Devah, "The use of field experiments for studies of employment discrimination: Contributions, critiques, and directions for the future," *The Annals of the American Academy of Political and Social Science*, 2007, *609* (1), 104–133.

_ , **Bart Bonikowski, and Bruce Western**, "Discrimination in a low-wage labor market: A field experiment," *American sociological review*, 2009, *74* (5), 777–799.

Pennycook, Gordon, Ziv Epstein, Mohsen Mosleh, Antonio A Arechar, Dean Eckles, and David G Rand, "Shifting attention to accuracy can reduce misinformation online," *Nature*, 2021, *592* (7855), 590–595.

Rich, Judith, "Do photos help or hinder field experiments of discrimination?," *International Journal of Manpower*, 2018.

Rooth, Dan-Olof, "Obesity, attractiveness, and differential treatment in hiring a field experiment," *Journal of human resources*, 2009, *44* (3), 710–735.

Rose, Michael E and Co-Pierre Georg, "What 5,000 acknowledgements tell us about informal collaboration in financial economics," *Research Policy*, 2021, *50* (6), 104236.

Rust, Niki, "A nifty guide for academics on using Twitter," *PLOS SciComm*, 2019.

Sarsons, Heather, Klarita Gërxhani, Ernesto Reuben, and Arthur Schram, "Gender differences in recognition for group work," *Journal of Political Economy*, 2021, *129* (1), 101–147.

Schultz, Robert and Anna Stansbury, "Socioeconomic diversity of economics PhDs," Technical Report 2022.

Sebo, Paul, "Performance of gender detection tools: a comparative study of name-to-gender inference services," *Journal of the Medical Library Association: JMLA*, 2021, *109* (3), 414.

Tzioumis, Konstantinos, "Demographic aspects of first names," *Scientific data*, 2018, *5* (1), 1–9.

Wolfers, Justin, "Twitter for Economists," Presentation at the American Economic Association Annual Meeting, 2015.

Wu, Alice, "Gendered language on the economics job market rumors forum," *AEA Papers and Proceedings*, 2018, *108*, 175–79.

_ , "Gender bias among professionals: an identity-based interpretation," *Review of Economics and Statistics*, 2020, *102* (5), 867–880.

Young, Alwyn, "Channeling Fisher: Randomization tests and the statistical insignificance of seemingly significant experimental results," *The Quarterly Journal of Economics*, 2019, *134* (2), 557–598.

Zinovyeva, Natalia and Manuel Bagues, "The role of connections in academic promotions," *American Economic Journal: Applied Economics*, 2015, *7* (2), 264–92.

Figures

Figure 1: Distribution of followers on *#EconTwitter*

(a) Followers by predicted gender

(b) Followers by race/ethnicity

(c) Followers by university affiliation

Notes: These figures show the distribution of the number of followers (in logs) from Twitter users in the *#EconTwitter* community, excluding users with zero followers. The sample is composed of the universe of Twitter accounts that tweeted or retweeted a status containing the term *#EconTwitter* between January and February 2022. We predicted gender from users' names using the NamSor tool, as described in the text. Figure 1a shows the distribution only for users for whom we were able to accurately predict gender ($N = 8,138$). We classified users' perceived race or ethnicity manually using profiles' metadata (profile pictures). Figure 1b shows the distribution of followers for users we were able to classify ($N =$). Finally, we obtained users' university affiliation by searching their bios. We consider a top-ten university to be the top ten universities in the 2017 USNews Ranking of universities in terms of graduate programs in economics. Figure 1c is conditional on users who are in academia, either as professors or as graduate students ($N = 5,432$).

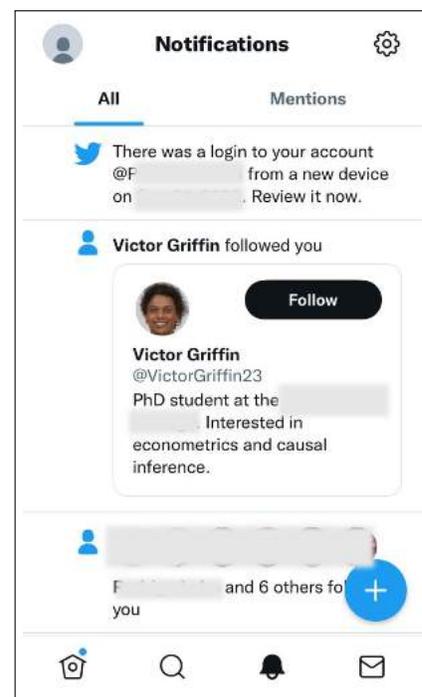Figure 2: Example of experimental accounts

Figure 3: Example of treatment notifications on desktop and mobile Twitter apps
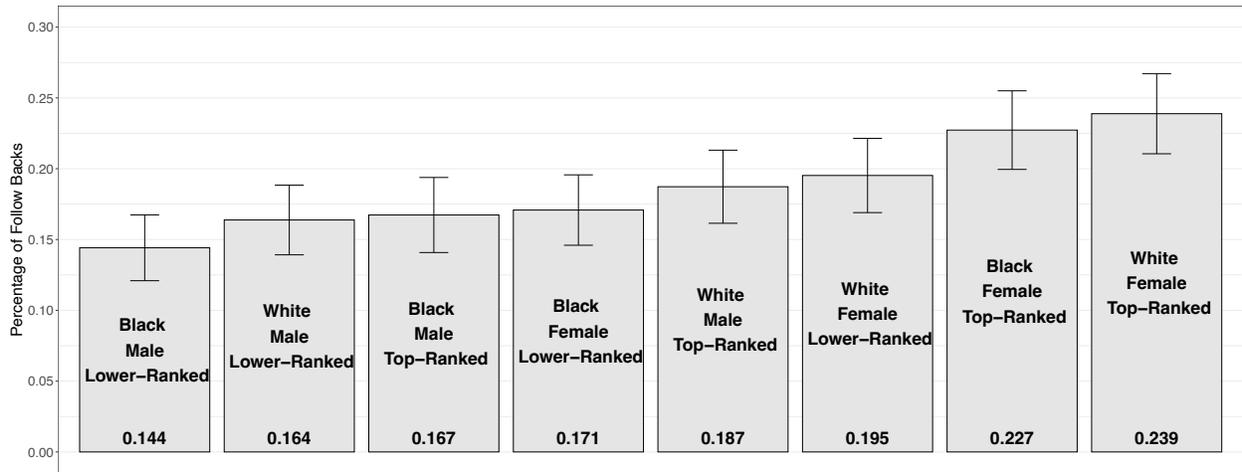

(a) Desktop notification

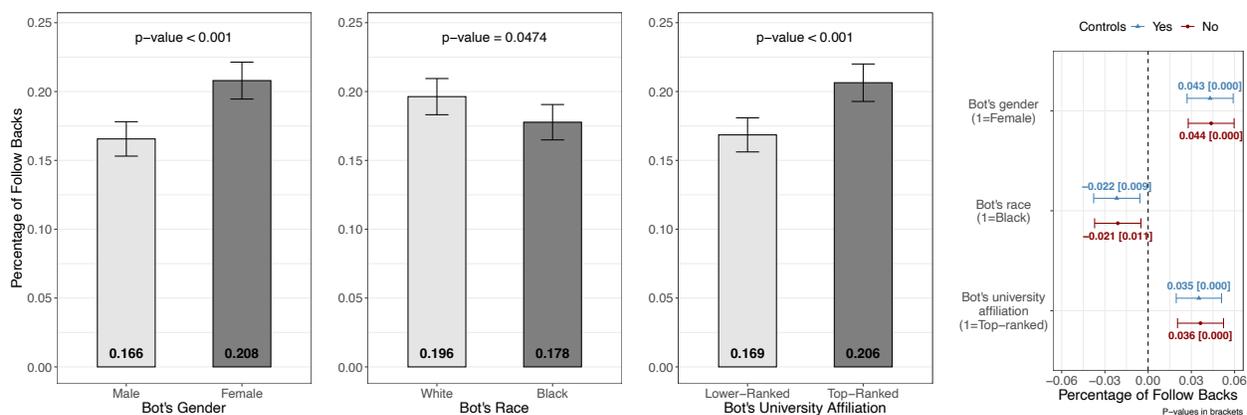(b) Desktop notification (after hovering the mouse cursor over the bot's profile)

(c) Mobile app notification

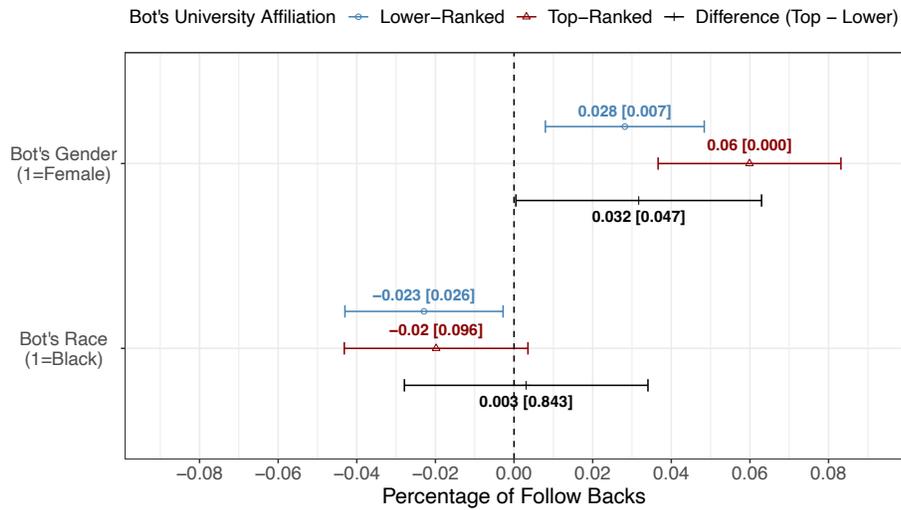Figure 4: Follow-back rates by bot group

Notes: The figure shows the average rate of follow-backs by bot type. The error bars show 95% confidence intervals. Data comes from the ten experimental waves, excluding shadow-banned accounts as discussed in the text.

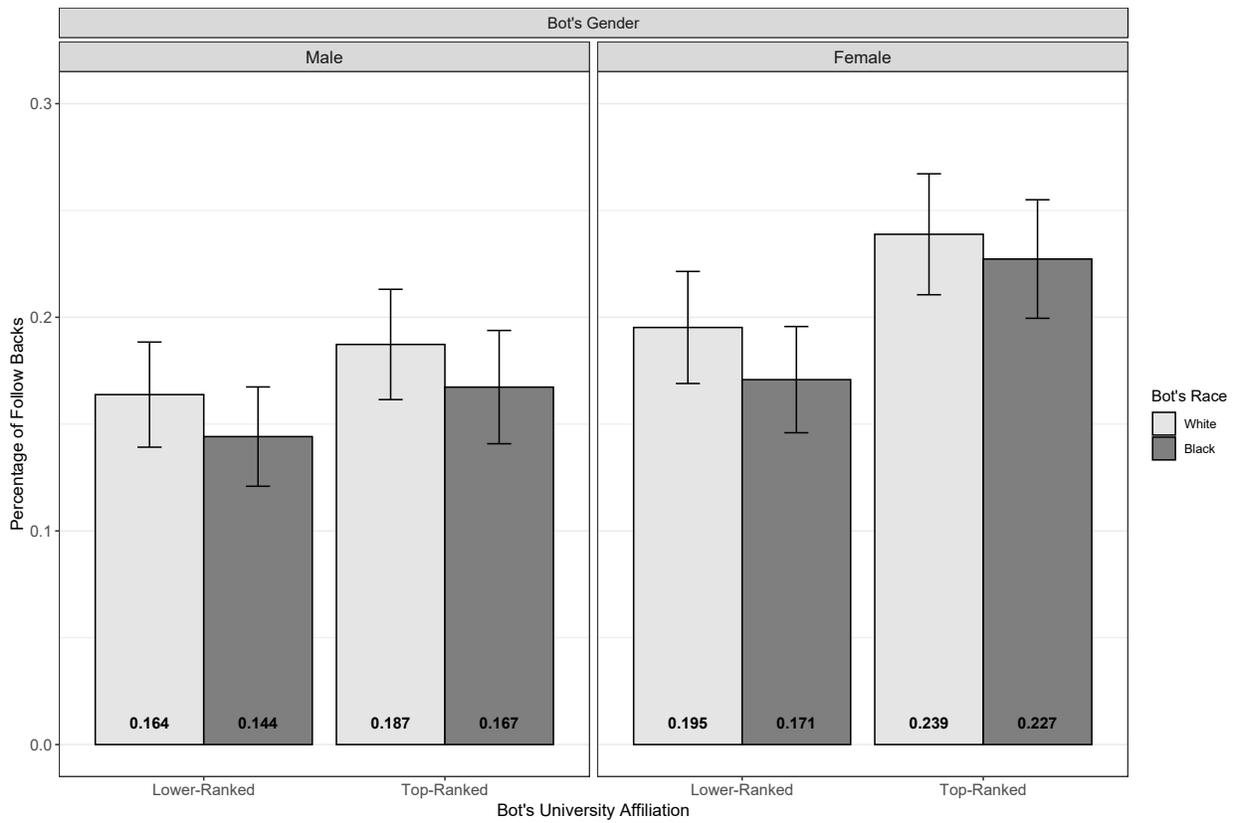Figure 5: Follow-back rates by bot group: marginal distributions

Notes: The first three figures show marginal follow-back rates for bot accounts from each group. Data comes from the ten experimental waves, excluding shadow-banned accounts as discussed in the text. The error bars show 95% confidence intervals for the mean follow-back, and the p-value displayed on top of each plot is the p-value for a standard t-test of difference in means between the two groups of that figure. The last plot shows point-estimates and 95% confidence intervals for the coefficients related to bot's gender, race and university affiliation, obtained by estimating equation (1) with and without controls. Confidence intervals are computed using standard errors clustered at the bot account level. P-values are in brackets.

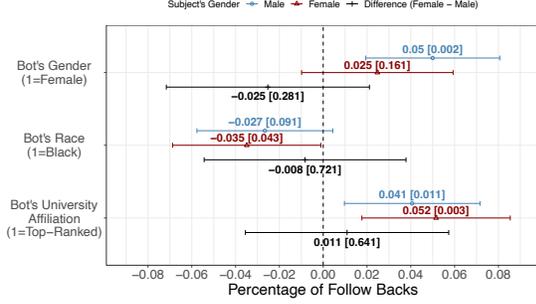Figure 6: Follow-back rate by bot university affiliation

Notes: The figure plots estimated follow-back rates by bot university affiliation. Specifically, we estimate equation (1) including interactions between bot university affiliation and gender, and between bot affiliation and race. The regression includes wave, strata and wave × strata fixed effects. This gives us an estimate of the differential follow-back rate for male and female bots conditional on university affiliation, as well as an estimate for the difference in these rates; the same analysis is then performed for bots representing Black and White students. Error bars represent 95% confidence intervals, and p-values are in brackets. Both confidence intervals and p-values are computed using standard errors clustered at the bot account level.

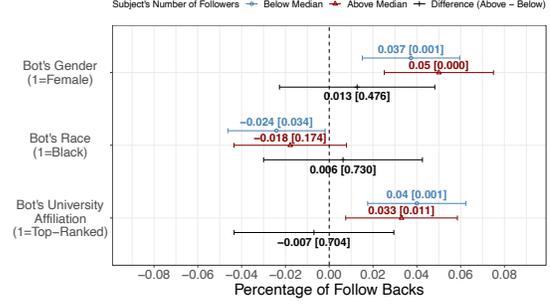## Figure 7: Follow-back rate by bot group – Full interactions

Notes: The figures show the follow-back rate for each type of bot account. The data comes from the ten experimental waves, excluding shadow-banned accounts as discussed in the text. The error bars show 95% confidence intervals for the mean follow-back.
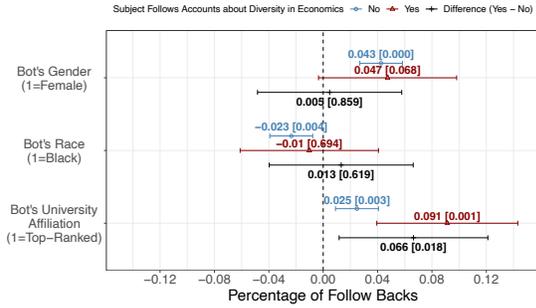
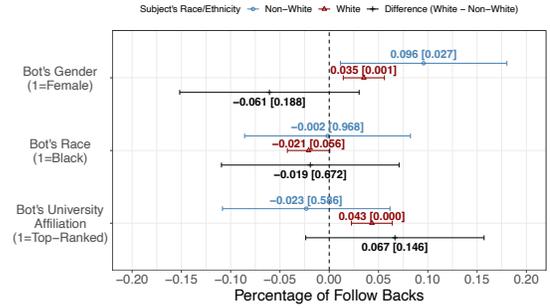Figure 8: Heterogeneity in follow-back behavior by subject characteristics

(a) Subject's gender

(b) Subject's account reach

(c) Subject's concern about the lack of
diversity in economics

(d) Subject's race/ethnicity

Notes: The figures display the estimated follow-back rate of different sets of subjects for each bot character-istic (gender, race and affiliation). Specifically, let $GROUP_i$ represent a dimension of subject heterogeneity (for instance, $GROUP_i$ equals one when a subject is predicted to be a woman and zero if they are predicted to be a man). We run a regression of the form of equation (2) and report, for each bot characteristic, the differential follow-back rate by each set of subjects and how this compares across the groups of subjects. For instance, in the case of the top three estimates reported in panel 8a, we report the estimate for β_1, which represents the differential follow-back rate by male subjects of female and male bots; the estimate for $\beta_1 + \beta_4$, which represents the differential follow-back rate by female subjects of female and male bots, and the difference in the behavior of female and male subjects, calculated as the difference in follow-back behavior between the two groups of subjects. The error bars show 95% confidence intervals for each estimate. P-values are in brackets. Confidence intervals and p-values are computed based on standard errors clustered at the bot account level.

Tables

Table 1: Procedures used to create the bot accounts

Element of Profile	Procedure
Profile Picture	Use AI generated images from Generated Media Inc.. The company's tool allows us to control several parameters when generating each picture: gender, head pose, age, emotion, skin tone, hair color, hair length, glasses and make-up. For each set of four profile pictures, we start from the same "base" face and vary gender (male or female) and skin tone (black or white).
Name	Randomly generated by matching a list of the most common first names and surnames in the US. We excluded from the list all names that are gender-neutral (specifically, we used the NamSor tool to predict the gender of the names in our list, and excluded those with less than 90% confidence in the gender prediction). We also excluded names that are specific to a certain race or ethnicity (to identify race and ethnicity specific names, we used data from Tzioumis (2018)).
Bio	The Bio from the bot accounts contains two pieces of information: first, the university where they claim to be doing their PhD; second, their research interests. To select the universities, we first considered the ten highest-ranked universities and the universities ranked between positions 79 and 100 in the same ranking that made their list of PhD students publicly available, based on the 2017 USNews rank of graduate universities in economics. We randomly selected 5 universities from each of these two sets. The interest was also randomly assigned from a list designed by the authors. Generally, the bio from a bot account was something like: "PhD student at University X. Interested in labor economics and economics of education.". We decided not to use the university's Twitter handles (for instance, @UniversityX) because this would likely affect follow recommendations made by Twitter's algorithm, which could bias the experiment (if Twitter recommended the bot accounts to users from the same university, bots from some universities could get a disproportional volume of followers for reasons unrelated to discrimination). It is harder for the algorithm to target recommendations when the Twitter handle is not used. Importantly, we do not explicitly say in the bio that the student is doing his or her PhD in economics. This is implicit from the interests listed.
Background Image	A landscape from the city where the student claims to be doing the PhD. We have a single landscape for each city.
Location	The bot accounts' profiles did not include a location.
Website	The bot accounts' profiles did not include a website.
Retweets	Before following subjects, the bot account retweets two statuses from accounts of academic journals in the field of economics. These two retweets are randomly chosen.
Followers	We asked a group of economic professors and graduate students to follow the bot accounts one day before the bot account followed the accounts randomly assigned to it.
Following	One day before following the accounts randomly assigned to it, the bot account follows all professors and graduate students we had asked to follow the account. It also follows some accounts from academic journals and other institutions related to the field of economics.

Notes: The table summarizes the procedures used to create the bot accounts.

Table 2: Descriptive statistics of the subject pool - Quantitative Variables

Variables	Mean	Std. Deviation	Median	Min	Max	Obs.
Number of followers	3,958.06	37,378.96	469	0	2,437,589	14,055
Number of accounts followed ('friends')	1,245.91	2,477.13	644	0	113,267	14,055
Number of statuses ('tweets')	22,067.7	83,014.79	2,559	0	2,696,665	14,055
Number of favorites ('likes')	21,361.06	62,001.76	3,729	0	1,250,869	14,055
Number of public lists	63.83	393.76	5	0	23,454	14,055
Share of tweets/retweets via Mobile App	0.63	0.41	0.81	0	1	13,263

Notes: The table shows summary statistics for the universe of accounts that tweeted or retweeted the hashtag *EconTwitter* between January 1st and February 28th, 2022. "Number of public lists" refers to the number of public lists that include the subject's Twitter profile. The share of tweets/retweets sent via mobile app is computed for tweets and retweets live-streamed during two periods in October 2022: between October 6th to 12th and between October 18th and 26th. For each user in our sample that tweeted at least once during the period, we computed the share of tweets sent via the Twitter mobile app as a percentage of all the tweets from that user.

Table 3: Descriptive statistics of the subject pool - Qualitative Variables

Variables	% Classified	N	%
Gender	58.17		
Female		2200	26.91
Male		5976	73.09
Continent	63.05		
Africa		344	3.88
Asia		793	8.95
Europe		3443	38.86
Latin America		612	6.91
North America (US/Canada)		3492	39.41
Oceania		177	2
Profession	60.45		
Professor		2911	34.26
Assistant Prof.		627	7.38
Associate Prof.		301	3.54
Undefined Prof.		1983	23.34
Government		426	5.01
Industry/Tech		1121	13.19
Institution		1021	12.02
Journalist		221	2.6
Non-profit/Multilateral Org.		269	3.17
PhD Student		1156	13.61
Post-Doc		272	3.2
Other Researcher		1099	12.94
Claims to be affiliated with top-ten university in bio	100		
No		13436	95.6
Yes		619	4.4
Race/Ethnicity	63.91		
White		7227	80.45
Black		713	7.94
Asian		452	5.03
Other		591	6.58
Follows Twitter account(s) addressing diversity in economics	100		
No		11926	84.85
Yes		2129	15.15
Verified	100		
No		13684	97.36
Yes		371	2.64
Has background picture	100		
No		3367	23.96
Yes		10688	76.04

Notes: The table shows the distribution of the categorical variables in the *#EconTwitter* sample (the universe of accounts that tweeted or retweeted the *EconTwitter* hashtag between January 1st and February 28th, 2022). The procedure to obtain each variable is described in Table B.2.

Table 4: Effect of bot characteristics on follow-backs

Dependent Variable:	Follow Backs (1=Yes)			
	Full Sample		Excluding Same University	
Model	(1)	(2)	(3)	(4)
Bot's Gender	0.04376***	0.04302***	0.04429***	0.0436***
(1=Female)	(0.00798)	(0.00802)	(0.00815)	(0.00819)
	[0.0000]	[0.0000]	[0.0000]	[0.0000]
Bot's Race	-0.02106**	-0.02176***	-0.02313***	-0.02329***
(1=Black)	(0.00806)	(0.00804)	(0.00824)	(0.0083)
	[0.0110]	[0.0080]	[0.0070]	[0.0080]
Bots' University Affiliation	0.03637***	0.03524***	0.03278***	0.03202***
(1=Top-Ranked)	(0.008)	(0.00789)	(0.00819)	(0.00805)
	[0.0000]	[0.0000]	[0.0000]	[0.0000]
Controls	No	Yes	No	Yes
Wave, Strata Fixed Effects	Yes	Yes	Yes	Yes
Observations	6920	6920	6735	6735
Dep. Variable Mean	0.187	0.187	0.188	0.188

Notes: The table displays regression results for the main experiment, using data from the ten experimental waves, excluding shadow-banned accounts. The first two columns display results using the full sample of subjects, while the last two exclude subjects that claim (in their Twitter bio) to be affiliated with the universities used in each wave. The dependent variable is an indicator equal to one if a Twitter user (subject) that was followed by a bot subsequently followed the bot back. The controls used in specifications (2) and (4) are (at the subject level): continent, profession, gender, affiliation to Top-10 university, year of account creation, has background picture, follows accounts addressing the lack of diversity in academia, has a verified account, number of Twitter followers and number of Twitter friends. Clustered standard errors at the bot-account level are in parentheses. P-values obtained by randomization inference, under the sharp null hypothesis of no treatment effect for each subject, are displayed in brackets. Significance codes: *** : $p < 0.01$, ** : $p < 0.05$, * : $p < 0.1$ refer to p-values computed using the clustered standard errors.

Table 5: Heterogeneity in follow-back behavior

Dependent Variable:	Follow Backs (1=Yes)	Dependent Variable:	Follow Backs (1=Yes)
Panel A: Subject's gender		**Panel C: Subject is concerned about the lack of diversity in economics**	
Bot's Gender (1=Female)	0.04767*** (0.00915) [0.0000]	Bot's Gender (1=Female)	0.04767*** (0.00915) [0.0000]
Bot's Race (1=Black)	-0.01807* (0.00933) [0.0000]	Bot's Race (1=Black)	-0.01807* (0.00933) [0.0000]
Bot's University Affiliation (1=Top-Ranked)	0.0332*** (0.00929) [0.0000]	Bot's University Affiliation (1=Top-Ranked)	0.0332*** (0.00929) [0.0000]
Bot's Gender × Female Subject	-0.0228 (0.02052) [0.1818]	Bot's Gender × Female Subject	-0.0228 (0.02052) [0.1818]
Bot's Race × Female Subject	-0.01751 (0.02011) [0.0909]	Bot's Race × Female Subject	-0.01751 (0.02011) [0.0909]
Bot's University × Female Subject	0.01855 (0.02026) [0.3636]	Bot's University × Female Subject	0.01855 (0.02026) [0.3636]
Observations	6920	Observations	6920
Dep. Variable Mean	0.189	Dep. Variable Mean	0.189
Panel B: Subject's account quality		**Panel B: Subject's race/ethnicity**	
Bot's Gender (1=Female)	0.03732*** (0.01115) [0.0000]	Bot's Gender (1=Female)	0.03732*** (0.01115) [0.0000]
Bot's Race (1=Black)	-0.0241** (0.01113) [0.0000]	Bot's Race (1=Black)	-0.0241** (0.01113) [0.0000]
Bot's University Affiliation (1=Top-Ranked)	0.03998*** (0.01125) [0.0000]	Bot's University Affiliation (1=Top-Ranked)	0.03998*** (0.01125) [0.0000]
Bot's Gender × Strong profile	0.01274 (0.01778) [0.5455]	Bot's Gender × Strong profile	0.01274 (0.01778) [0.5455]
Bot's Race × Strong profile	0.0063 (0.01818) [0.8182]	Bot's Race × Strong profile	0.0063 (0.01818) [0.8182]
Bot's University × Strong profile	-0.00696 (0.01828) [0.5455]	Bot's University × Strong profile	-0.00696 (0.01828) [0.5455]
Observations	6920	Observations	6920
Dep. Variable Mean	0.187	Dep. Variable Mean	0.187
Wave, Strata Fixed Effects	Yes	Wave, Strata Fixed Effects	Yes

Notes: The table displays estimated coefficients for interaction terms between bot and subject characteristics, using data from the ten experimental waves, excluding shadow-banned accounts. The specification used is the one in Equation (2). The dependent variable is an indicator equal to one if a Twitter user (subject) that was followed by a bot subsequently followed the bot back. Clustered standard errors at the bot-account level are in parentheses. P-values obtained by randomization inference, under the sharp null hypothesis of no treatment effect for each subject, are displayed in brackets. Significance codes: $^{***}: p < 0.01$, $^{**}: p < 0.05$, $^{*}: p < 0.1$ refer to p-values computed using the clustered standard errors.

Appendix

A Additional Information on Experimental Design

A.1 Procedure for obtaining the subject pool

(i) From January 1st, 2022, to February 28th, 2022, obtain all Twitter users that either tweeted or retweeted a status containing the term *"#econtwitter"*[23] $\rightarrow 14,449$ accounts.

(ii) Remove accounts that no longer exist, accounts that are clearly bots, and protected accounts[24] $\rightarrow 14,055$ accounts.

(iii) Compute follows/friends ratio for the remaining account. Remove accounts with a follows/friends ratio above 15 and accounts with fewer than 10 friends and institutional accounts $\rightarrow 10,226$ accounts[25]. This is our final subject pool.

A.2 Experimental Wave Timeline

Table A.1: Wave timeline

Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
...	Create Accounts Introductory Tweet	Follow Friends Rt (x2)	$d=0$ Follow subjects $d=1$	$d=2$ $d=3$	$d=4$ $d=5$	$d=6$ $d=7$
Monday	**Tuesday**	**Wednesday**	**Thursday**	**Friday**	**Saturday**	**Sunday**
$d=8$ $d=9$	$d=10$ $d=11$	$d=12$ $d=13$	$d=14$ $d=15$	$d=16$ $d=17$	$d=18$ $d=19$	$d=20$ $d=21$
Monday	**Tuesday**	**Wednesday**	**Thursday**	**Friday**	**Saturday**	**Sunday**
$d=22$ $d=23$	$d=24$ Delete Accounts	...	...	...	...	...

Notes: The table shows the timeline of an experimental wave. Accounts are active for 12 days after following the experimental subjects. Each wave starts on a Tuesday, so there will always be accounts from two waves active in the same period (eight of them in the second week and eight in the first week of the wave).

---

[23]The search considered all variations of capital and small letters for the term.

[24]Note that an account that tweeted a status containing "#econtwitter" at the beginning of January, for instance, may no longer exist at the beginning of March (the account owner may have deleted the account). We identified accounts that were clearly bots by analyzing the accounts' Twitter bios. In Twitter, "Protected" accounts are the ones that choose not to be public, restraining their information and interaction to the account's friends.

[25]15 is approximately the follows/friends ratio of the 95th percentile of the subject pool sample after step (ii). We removed accounts with too few friends because those accounts are likely to be inactive or (at least) are extremely unlikely to follow an unknown account.

A.3 AI-generated profile pictures used in the experiment

Figure A.1: Profile Pictures used in the experiment

Notes: The picture shows all AI-generated images used as profile pictures in the experiment. The four pictures in each row share the same base image, which is one of the four images (randomly chosen). To construct the other three images from the base, we kept all attributes constant apart from gender or skin tone. The last column in the table indicates the experimental wave and university affiliation for which each set of images was used. In the creation of the Twitter profiles, we cropped the images so the watermark on the left-hand side of the pictures did not appear, even if someone clicked on the profile picture.

B   Additional figures and tables

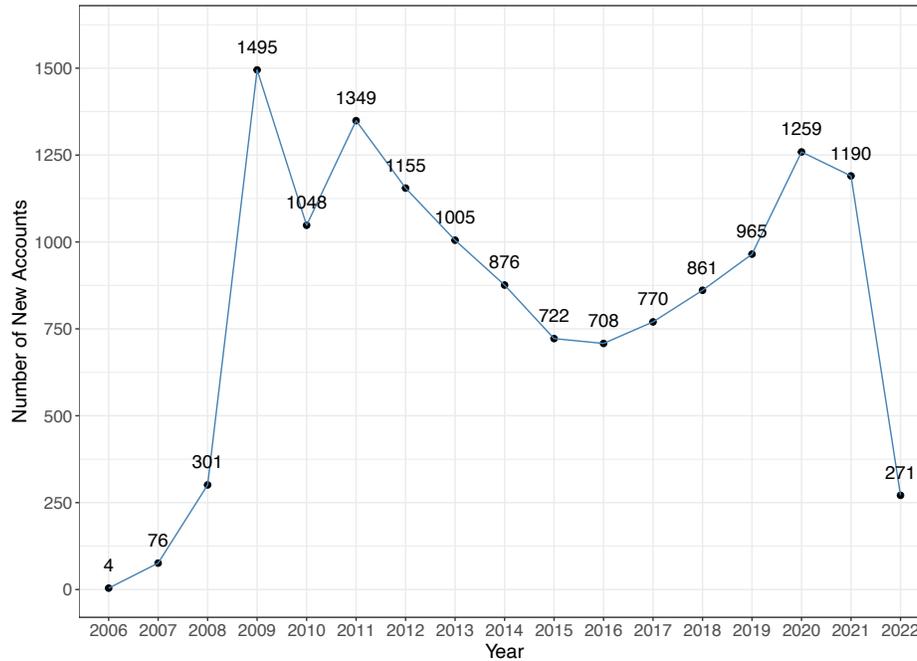## B.1   Characteristics of Twitter profiles of real PhD students

Table B.1: Summary statistics of real-life Twitter profiles of first- and second-year economics PhD students

	2021 Cohort			2020 Cohort		
	Median	Mean	Std.Dev	Median	Mean	Std.Dev
Tweets	94.5	357.33	917.26	47	174.58	252.40
Following	331.5	471.29	495.88	279	382.32	327.60
Followers	152.5	372.96	507.11	115	401.00	807.61
Website	0.0	0.29	0.46	0	0.26	0.45
Background Image	1.0	0.71	0.46	0	0.47	0.51
Location	0.0	0.42	0.50	1	0.53	0.51
Profile Pic is a Self-Portrait	1.0	0.83	0.38	1	0.89	0.32

Notes: The table shows descriptive statistics of Twitter profiles from first and second-year PhD students from three universities. For the first-year (2021) cohort, 24 out of 59 students had profiles we could find (40.67%) as of January 6th, 2022. For the second-year (2020) cohort, we could find profiles for 18 out of 56 students (32.14%).

B.2 Number of new accounts created by year in the *#EconTwitter* community

Figure B.1: Number of new accounts created by year in the experimental sample

Notes: The figure shows the year of creation of the accounts in the experimental sample. The sample is composed of all accounts that tweeted or retweeted a message containing the hashtag *#EconTwitter* between January and February 2022.

B.3   Description of subject-level variables

Table B.2: Description of variables at the subject level

Variable	Description	N (%)
Gender	Whether the account belongs to someone identified as male or female. To obtain this information, we used the users' full name to predict its gender, using the **NamSor** tool,[1] which accurately predicts gender based on full names. We only considered predictions done with above 90% confidence, and assigned as missing the gender information for the accounts with confidence below this threshold. We manually checked a randomly selected subsample of 100 accounts, and obtained 98% accuracy.	8,316 (58.4%)
Nfavorites	Number of tweets marked as "favorite" (i.e., "liked") by the user.	14,055 (100%)
Nfollows	Number of accounts the user follows.	14,294 (100%)
Nfriends	Number of friends the user has, i.e., number of accounts that follow the user.	14,055 (100%)
Verified	Indicator variable equal to one if the account is verified, a "badge" provided by Twitter to signal that the account is authentic.	14,055 (100%)
Continent	The continent in which the user lives. We obtained this information via the "location" information from Twitter. This information is provided by the user and can, in principle, be anything (it does not have to be a real location and does not have to be correct). We classified the "real" location given by region: North America, South and Central America, Europe, Asia, Africa, Oceania. If a person indicated more than one place from different continents, we classified the location as missing. At the end of the procedure, we manually checked a random subsample of 100 accounts and obtained 100% accuracy.	8,931 (62.7%)
Profession	The user's profession. We classified professions using the user's account description (or "bio"). The list of professions/areas of work is: professor (which is subdivided into "assistant", "associate" and "other"); PhD student; Post-Doc; Other academic position (for instance, Research Fellow, Research Assistant, etc.); Industry/Tech; Government; Non-profit/Multi-lateral Organization; Journalist. We first searched for keywords related to each profession, and then manually verified the matches. At the end of the procedure, we checked a random subsample of 100 accounts and obtained 99% accuracy.	8,555 (60.1%)
Race/Ethnicity	We manually classify perceived race or ethnicity from subjects' profile pictures.	8,983 (63.9%)

Table B.2: Description of variables at the subject level (Continued)

Variable	Description	N (%)
University Affiliation	Indicator variable equal to one if a user is affiliated to a highly ranked university. To obtain this information, we also consider the user's account description ("bio") and search for keywords associated with the highly ranked universities. We obtain this variable for top-ten and top-twenty US universities according to the USNews Ranking.	14,055 (100%)
Year of Account Creation	The year in which the account was created. This information is provided by Twitter's API and is therefore perfectly precise.	14,055 (100%)
Concern about the Lack of Diversity in Economics	Indicator variable equal to one if a user follows at least one Twitter account dedicated to this topic. The list of accounts we consider are: @AEACSMGEP (the American Economic Association Committee on the Status of Minority Groups in the Economics Profession); @AEACSWEP (the AEA Committe on the Status of Women in the Economics Profession); @ResearchInColor (a foundation whose objective is to "diversify economics by increasing the number and retention of scholars of color in economic disciplines through mentoring and financial support); @SadieCollective (a collective that is "addressing the pipeline pathway for Black women in economics and related fields") and @weconpol ("an inclusive community for women interested in econ, policy and development").	14,055 (100%)
Background Picture	Indicator variable equal to one if the user has a background picture (banner).	14,055 (100%)
Share of Tweets/Retweets via Mobile App	We live-streamed tweets and retweets from the sample of subjects during two weeks in October 2022 (October 06th-12th and 18th-24th). For each tweet, we collected the source (e.g., mobile app, desktop, etc.). For each user, we compute the share of tweets and retweets in this period that were sent via Twitter's mobile app.	13,263 (94.4%)

[1] We chose this tool for a few reasons: first, it has already been used in academia, including to predict names using Twitter data (e.g., Hridoy et al. (2015)); second, it has been shown to be at least as accurate as similar tools (Sebo, 2021); third, its database includes names from a variety of countries, and permits the analysis of full names.

Notes: The table lists and describes the variables obtained for the users in the subject pool. Column N (%) shows the number of accounts and the percentage of the total pool for which we were able to obtain each piece of information.

B.4 Balance and evidence of no differential attrition

Table B.3: Balance table

	Treatment Arm (Bot's characteristics)								
	Male				Female				
	White		Black		White		Black		
	Top-ranked	Lower-ranked	Top-ranked	Lower-ranked	Top-ranked	Lower-ranked	Top-ranked	Lower-ranked	
Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	F Stat [p-value]
Number of followers	1,176.3	1,268.6	1,190	1,356.8	1,297.3	964.7	1,323.7	1,467.7	0.0079 [1.00]
	(3,841.1)	(4,655.4)	(3,688.7)	(5,375)	(5,739)	(3,282.6)	(5,105.4)	(7,180)	
Number of friends	1,221.3	1,254.5	1,406.2	1,341.4	1,344.6	1,227.9	1,403.2	1,421.4	0.0091 [1.00]
	(1,590.3)	(1,918.4)	(2,193.9)	(3,087)	(4,099.7)	(1,370.4)	(2,780.2)	(2,276.2)	
Number of statuses ('tweets')	20,004.7	17,642.6	22,446.5	19,747.7	16,310.8	18,150.2	20,993.9	21,859.3	0.007 [1.00]
	(92,373.2)	(65,217.5)	(69,613.1)	(68,805)	(47,260.7)	(56,881)	(76,448.8)	(105,441.2)	
Number of favorited statuses ('likes')	21,156.6	20,104.4	23,046.6	25,402	23,033.7	20,377.7	23,116.1	21,344	0.0073 [1.00]
	(64,822.6)	(54,783.8)	(65,082.2)	(72,248.4)	(60,450.7)	(53,227.9)	(65,656)	(56,885.5)	
Number of lists	24.766	27.324	25.844	25.796	27.185	19.585	32.052	42.824	0.0226 [1.00]
	(97.3)	(113.1)	(101.8)	(91.1)	(112)	(65.9)	(155.7)	(254.5)	
Account is verified	0.008	0.014	0.008	0.01	0.008	0.006	0.015	0.009	0.009 [1.00]
	(0.089)	(0.117)	(0.088)	(0.101)	(0.089)	(0.075)	(0.121)	(0.095)	
Year of account creation	2,014.6	2,014.5	2,014.5	2,014.3	2,014.6	2,014.6	2,014.6	2,014.4	0.0063 [1.00]
	(4.176)	(4.164)	(4.178)	(4.118)	(4.112)	(4.059)	(4.179)	(4.158)	
Has background picture	0.734	0.724	0.771	0.773	0.741	0.755	0.78	0.759	0.019 [1.00]
	(0.442)	(0.447)	(0.42)	(0.419)	(0.439)	(0.43)	(0.415)	(0.428)	
Follows diversity accounts	0.154	0.156	0.176	0.159	0.176	0.16	0.172	0.146	0.0079 [1.00]
	(0.362)	(0.363)	(0.381)	(0.366)	(0.381)	(0.367)	(0.377)	(0.353)	
Female	0.174	0.179	0.178	0.178	0.175	0.173	0.175	0.167	8e-04 [1.00]
	(0.379)	(0.383)	(0.383)	(0.383)	(0.38)	(0.378)	(0.38)	(0.374)	
Affiliated to Top-10 University	0.031	0.052	0.033	0.042	0.042	0.037	0.034	0.044	0.0111 [1.00]
	(0.172)	(0.221)	(0.178)	(0.201)	(0.201)	(0.19)	(0.182)	(0.205)	
Has rainbow flag on profile	0.019	0.014	0.016	0.014	0.011	0.014	0.017	0.008	0.0077 [1.00]
	(0.138)	(0.117)	(0.124)	(0.116)	(0.106)	(0.116)	(0.13)	(0.089)	
States preferred pronouns	0.057	0.057	0.069	0.056	0.064	0.074	0.08	0.066	0.0109 [1.00]
	(0.232)	(0.232)	(0.254)	(0.229)	(0.245)	(0.262)	(0.271)	(0.248)	
Profession									
Graduate Student	0.216	0.208	0.229	0.216	0.216	0.23	0.211	0.223	0.0032 [1.00]
	(0.412)	(0.406)	(0.42)	(0.412)	(0.412)	(0.421)	(0.409)	(0.416)	
Professor	0.211	0.214	0.193	0.2	0.21	0.195	0.205	0.2	0.0031 [1.00]
	(0.408)	(0.411)	(0.395)	(0.4)	(0.408)	(0.397)	(0.404)	(0.4)	
Works in Industry/Tech	0.067	0.104	0.084	0.09	0.091	0.074	0.081	0.081	0.015 [1.00]
	(0.25)	(0.306)	(0.277)	(0.286)	(0.288)	(0.262)	(0.273)	(0.274)	
Other	0.076	0.049	0.058	0.07	0.055	0.076	0.073	0.067	0.0152 [1.00]
	(0.265)	(0.217)	(0.233)	(0.256)	(0.228)	(0.265)	(0.26)	(0.25)	
Region									
Europe	0.233	0.243	0.234	0.251	0.257	0.236	0.26	0.25	0.0053 [1.00]
	(0.423)	(0.429)	(0.424)	(0.434)	(0.437)	(0.425)	(0.439)	(0.433)	
Canada/US	0.243	0.228	0.255	0.254	0.24	0.224	0.255	0.232	0.0076 [1.00]
	(0.429)	(0.42)	(0.436)	(0.436)	(0.427)	(0.417)	(0.436)	(0.422)	
Other	0.131	0.147	0.159	0.151	0.133	0.159	0.142	0.149	0.0081 [1.00]
	(0.337)	(0.354)	(0.366)	(0.358)	(0.339)	(0.366)	(0.349)	(0.357)	
Number of treated observations	881	873	765	881	875	881	880	884	
%	0.127	0.126	0.111	0.127	0.126	0.127	0.127	0.128	
Attrition (not treated)	16	25	16	17	22	17	18	19	0.0048 [1.00]
% of assigned to treatment	0.018	0.028	0.02	0.019	0.025	0.019	0.02	0.021	

Notes: This table presents descriptive statistics of treated subjects in the experiment for each of the eight treatment arms. For each treatment arm and each pre-treatment variable, the table presents the mean value among treated subjects, as well as its standard deviation (in parentheses). The last column in the table displays an F-statistic of a joint test of difference in means across the treatment arms, along with the test's p-value. The F-statistic is computed from a regression of the pre-treatment variable on the treatment indicators. For all pre-treatment variables, we cannot reject the null hypothesis of equality of means across all eight treatments. The description of the variables is in the Appendix. The row "Number of treated obs." shows the number of treated observations (i.e., accounts followed by a bot) for each treatment arm, while "%" shows the percentage treated among all treated participants. The row "Attrition" shows the number of participants assigned to each treatment that could not be treated (either because they deactivated their account, were suspended by Twitter, or chose to make their profile private). The F-statistic displayed for this row is obtained from a regression of the attrition indicator on the treatment indicators. The last row shows the percentage of participants assigned to treatment that could not be treated.
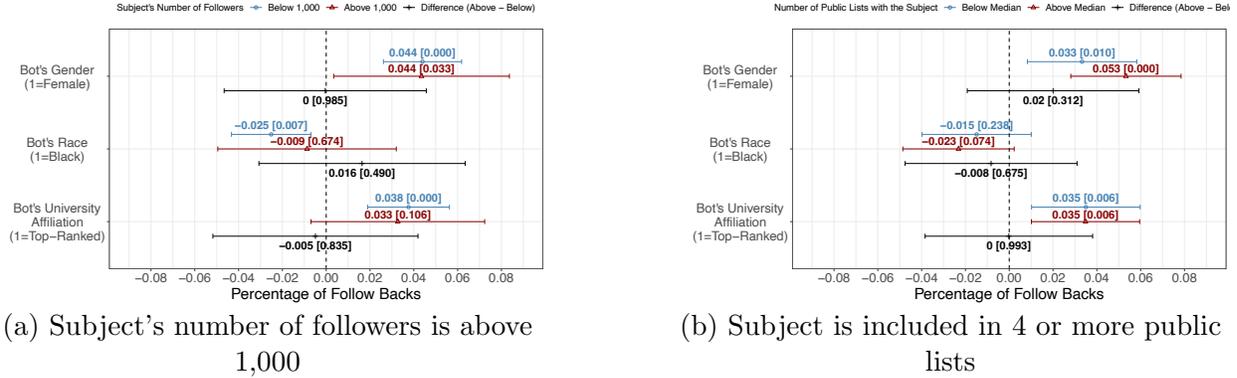
Table B.4: Differential attrition

	Treatment Arm (Bot's characteristics)								
	Male				Female				
	White		Black		White		Black		
	Top-ranked	Lower-ranked	Top-ranked	Lower-ranked	Top-ranked	Lower-ranked	Top-ranked	Lower-ranked	
Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	F Stat [p-value]
Number of followers	845.4	760.8	1,686.8	1,204.2	947.6	1,444.6	621.5	1,094.5	0.113 [0.997]
	(852.1)	(1,384.3)	(3,703.5)	(3,590.2)	(1,769.5)	(4,352)	(766.1)	(2,048.1)	
Number of friends	1,857.2	1,479.1	2,020.1	1,454.7	1,700.1	1,879.8	1,568	1,232.2	0.0586 [1.00]
	(1,741.9)	(1,802.1)	(3,518.2)	(3,314.5)	(1,820.6)	(4,549.1)	(1,589.7)	(1,388.5)	
Number of statuses ('tweets')	72,216.1	74,999.8	33,770.6	19,775.5	64,219	31,836.8	30,754.6	39,077.7	0.1703 [0.99]
	(170,092.4)	(190,548.1)	(56,954)	(41,656.9)	(160,830.6)	(74,665.9)	(62,575.3)	(127,206.9)	
Number of favorited statuses ('likes')	76,527.5	57,006.5	26,767.8	21,817.2	43,991.5	27,468.9	67,903.7	27,099.6	0.2199 [0.979]
	(192,158.3)	(122,428.2)	(34,879.2)	(31,661)	(81,500.1)	(84,383.6)	(152,098.8)	(46,041.1)	
Number of lists	42.625	15.68	15.75	5.647	8.136	7.765	8.722	32.579	0.3441 [0.929]
	(110.6)	(49.8)	(34.6)	(9)	(23.2)	(12.3)	(16.4)	(86.8)	
Account is verified	0	0	0	0	0	0	0	0	0 [-]
	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	
Year of account creation	2,015.5	2,015.6	2,016.1	2,018.1	2,016.5	2,014.8	2,014.6	2,016.3	0.363 [0.919]
	(5.007)	(4.803)	(2.778)	(3.436)	(4.351)	(4.72)	(4.461)	(4.042)	
Has background picture	0.625	0.72	0.875	0.706	0.773	0.882	0.944	0.789	0.3792 [0.91]
	(0.5)	(0.458)	(0.342)	(0.47)	(0.429)	(0.332)	(0.236)	(0.419)	
Follows diversity accounts	0.062	0.12	0.062	0.059	0.091	0.118	0	0.105	0.1283 [0.996]
	(0.25)	(0.332)	(0.25)	(0.243)	(0.294)	(0.332)	(0)	(0.315)	
Female	0.125	0.16	0	0.176	0.182	0.059	0.111	0	0.3491 [0.926]
	(0.342)	(0.374)	(0)	(0.393)	(0.395)	(0.243)	(0.323)	(0)	
Affiliated to Top-10 University	0	0.04	0	0.059	0.091	0	0	0	0.3074 [0.947]
	(0)	(0.2)	(0)	(0.243)	(0.294)	(0)	(0)	(0)	
Has rainbow flag on profile	0.125	0.04	0	0	0.091	0	0	0.053	0.3414 [0.931]
	(0.342)	(0.2)	(0)	(0)	(0.294)	(0)	(0)	(0.229)	
States preferred pronouns	0.062	0.04	0	0	0.273	0	0	0	1.3772 [0.238]
	(0.25)	(0.2)	(0)	(0)	(0.456)	(0)	(0)	(0)	
Profession									
Graduate Student	0.062	0.16	0.188	0.176	0.227	0.059	0.056	0.158	0.2176 [0.979]
	(0.25)	(0.374)	(0.403)	(0.393)	(0.429)	(0.243)	(0.236)	(0.375)	
Professor	0.062	0.08	0	0	0	0.176	0	0.053	0.5069 [0.825]
	(0.25)	(0.277)	(0)	(0)	(0)	(0.393)	(0)	(0.229)	
Works in Industry/Tech	0.188	0.04	0.062	0	0.136	0.118	0.111	0.316	0.571 [0.776]
	(0.403)	(0.2)	(0.25)	(0)	(0.351)	(0.332)	(0.323)	(0.478)	
Other	0	0.04	0	0	0.045	0.059	0	0	0.1939 [0.985]
	(0)	(0.2)	(0)	(0)	(0.213)	(0.243)	(0)	(0)	
Region									
Europe	0.188	0.32	0.188	0.059	0.091	0.118	0.056	0.158	0.4176 [0.886]
	(0.403)	(0.476)	(0.403)	(0.243)	(0.294)	(0.332)	(0.236)	(0.375)	
Canada/US	0.312	0.2	0.188	0.235	0.273	0.176	0.167	0.158	0.1004 [0.998]
	(0.479)	(0.408)	(0.403)	(0.437)	(0.456)	(0.393)	(0.383)	(0.375)	
Other	0.188	0.16	0.25	0.294	0.136	0.294	0.111	0.263	0.188 [0.987]
	(0.403)	(0.374)	(0.447)	(0.47)	(0.351)	(0.47)	(0.323)	(0.452)	
Number of untreated observations	16	25	16	17	22	17	18	19	0.0048 [1.00]

Notes: This table presents descriptive statistics of subjects that attrited, i.e., that were assigned to treatment but that were not treated in the experiment, for each of the 8 treatment arms. This could happen for one of three reasons: the users deactivated their account, were suspended by Twitter, or chose to make their profile private. For each treatment arm and each pre-treatment variable, the table presents the variable mean among attrited subjects, as well as the standard deviation (in parentheses). The last column in the table displays an F-statistic of a joint test of difference in means across the treatment arms, along with the test's p-value. The F-statistic is computed from a regression of the pre-treatment variable on the treatment indicators. For all pre-treatment variables, we cannot reject the null hypothesis of equality of means across all eight treatments, i.e., there is no differential attrition. The description of the variables is in the Appendix. The F-statistic displayed in the last row is obtained from a regression of the attrition indicator on the treatment indicators.

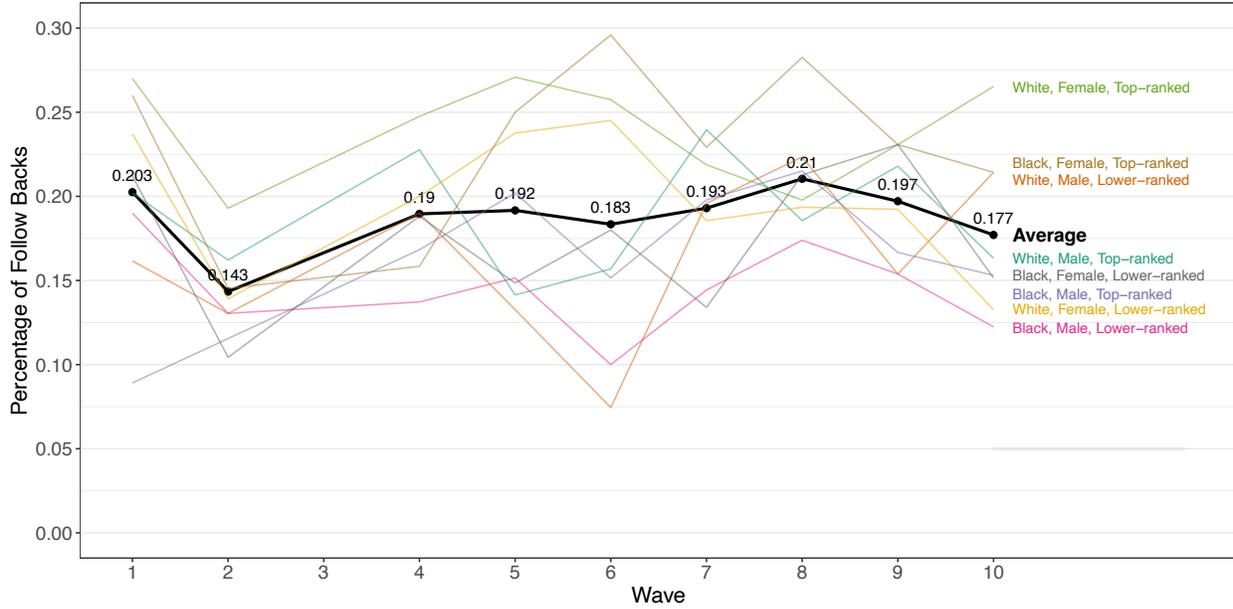# B.5 Robustness to alternative definitions of account reach

Figure B.2: Heterogeneity in follow-back behavior by subject's account reach – alternative definitions of reach

(a) Subject's number of followers is above 1,000

(b) Subject is included in 4 or more public lists

Notes: The figures show estimates of the follow-back behavior of strong and weak accounts analogous to those of Figure 8b, using different definitions of strength. The first panel considers accounts with more than 1,000 followers as strong, while the second panel considers accounts listed in at least four public lists as strong. A Twitter list is a selection of Twitter accounts that talk about the same topic. Lists can be created by Twitter users themselves, and then be left public – so that other users can subscribe to "follow" this list. Thus, if an account is included in a public list, this means that other users find this account worth following. Four lists was chosen as a cutoff because this is the median value of the variable in this sample.

B.6 Evolution of experimental uptake

Figure B.3: Evolution of the follow-back rate across experimental waves

Notes: The figure plots the evolution of the follow-back rate in the experiment across experimental waves. The thick black line is the average follow-back rate for all accounts in the wave, while the colored lines represent the follow-back rate in each wave for each bot type. The first wave started on May 26th, 2022, and the last wave started on July 28th, 2022.
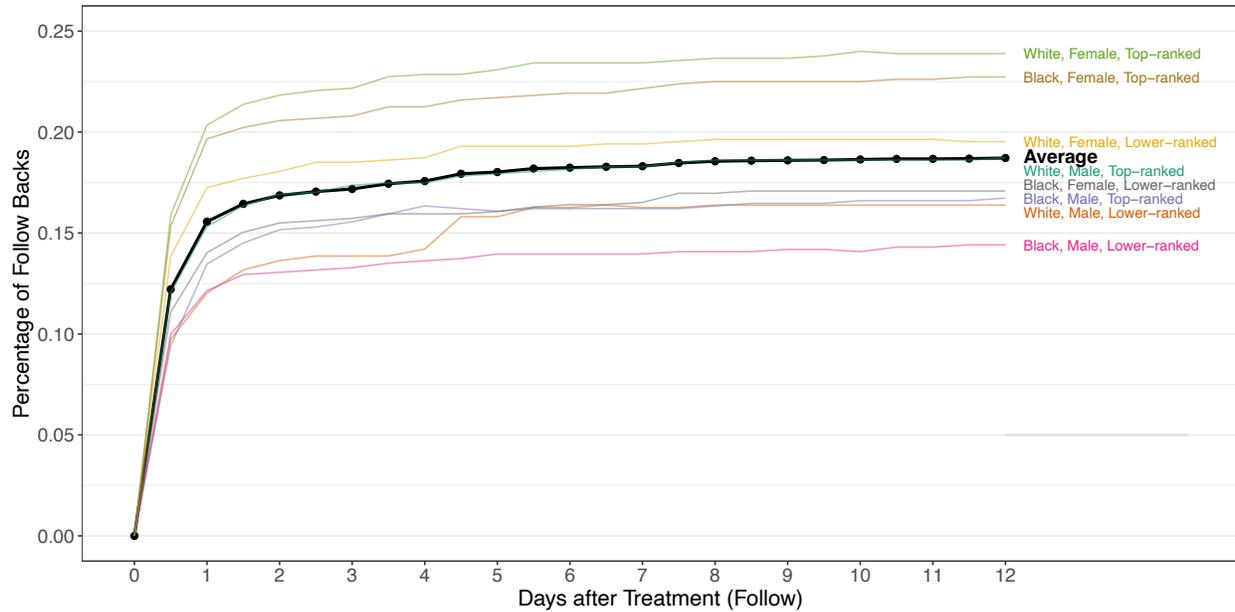
B.7 Organic follows by bot group

Table B.5: Average number of organic follows by bot group

Group	Average Number of Organic Follows	P-value
Bot's Gender		
Female	1.472	0.959
Male	1.457	
Bot's Race		
Black	1.343	0.413
White	1.583	
Bot's Affiliation		
Lower-ranked	1.333	0.364
Top-ranked	1.600	

Notes: The table displays the average number of organic follows (i.e., follows by users who were not followed by the bots) for each group of experimental accounts, across all ten waves (excluding shadow-banned accounts). The p-value in the last column refers to the p-value of a simple test of difference in means between the two types of each group.

B.8 Evolution of follow-backs within waves of the experiment

Figure B.4: Evolution of follow-backs within an experimental wave

Notes: The figure plots the evolution on the follow-back rate of the experimental accounts within each wave. The *x* axis represents the number of days after the bot follows the subjects (which happens on day 0). The thick black line represents the average follow-back rate across all types of bots, while the colored lines are the follow-back rate for each bot type. Data is pooled across the ten experimental waves, excluding shadow-banned accounts as discussed in the text. In all waves, we collected follow-back data up to twelve days after the treatment.

C Implementation Details

This Appendix describes how we implemented the experiment in detail, focusing on the activation of accounts in each round and how we followed subjects assigned to each treatment. Most procedures related to account activation and following subjects were done manually to avoid detection by Twitter.

In each wave, we activated eight accounts (one of each group). First (on day "zero"), each account followed a set of "elite" accounts from *#EconTwitter* (accounts from journals and associations), and the accounts of the approximately 30 colleagues who knew about the experiment and who we asked to follow the accounts back. This was done so that the experimental accounts looked more realistic (since they would already have a set of followers and friends). Note that this set of followers and friends was the same for all accounts in a given wave. At this initial moment, all accounts also retweeted two statuses from accounts of economic journals.[26] The statuses were chosen randomly from all tweets published by these accounts on the week the bot accounts were activated that had garnered more than three retweets. We also randomized the order of account activation in order to avoid introducing any bias related to timing.

After these procedures, the accounts were ready to follow the subjects assigned to them. This happened on the following day (Day "one"). At this time, each account followed the subjects randomly assigned to it. We randomized the order of follows both within and across accounts (i.e., the order of subjects to be followed by each account was random, as was the order of accounts following the subjects). This procedure always happened on Thursdays, and the distance between the first and last subject followed in each wave was at most 1h30 (this ensured that the timing at which the treatment notification was received is comparable across waves and subjects within waves). Apart from following all of its subjects, every bot account followed an account of someone who was aware of the experiment. This person would then inform us if he or she received a notification of this follow. If not, we considered the respective bot to have been "shadow-banned" and excluded it from the analysis (as pre-registered). This happened with a single account of the 80 we created during the experiment. We also had eight accounts (one from each group) be suspended by Twitter. These accounts were also excluded from the experiment because the suspension happened before we could complete the treatment (this hypothesis was also pre-registered). Thus, the final experimental data includes data on 71 (non-shadow banned and non-suspended) accounts across ten waves.

During the first five experimental waves, we created new Twitter accounts for the bots at the beginning of every single wave and deleted the accounts from the previous round. After

---

[26]The accounts we chose to retweet from are: @AEAJournals (journals from the American Economic Association), @ecmaEditors (Econometrica), @JPolEcon (Journal of Political Economy), @QJEHarvard (Quarterly Journal of Economics), @RevEconStudies (Review of Economic Studies), @restatjournal (Review of Economics and Statistics), @JEEA_News (Journal of the European Economic Association), @EJ_RES (Economic Journal), @JPubEcon (Journal of Public Economics), @nberpubs (National Bureau of Economics Working Papers), @qe_editors (Quantitative Economics), @EconTheory (Theoretical Economics), @J_HumanResource (Journal of Human Resources), @RevOfFinStudies (Review of Financial Studies, @Jof-Finance (Journal of Finance), and @J_Fin_Economics (Journal of Financial Economics).

these first five waves, we realized that this procedure was unnecessary since we could simply delete all account information and start over with a new identity. Specifically, at the end of the activation period, we first removed all followers and friends from the account (except the people who knew about the experiment and who we asked to follow all accounts) and deleted all of its posts. We then changed the account's name, description, Twitter handle, profile picture, and background picture, so that the account started over with a new identity. This new procedure saved time and reduced the likelihood of suspensions by Twitter, so we adopted it starting at wave 6. To guarantee that no spillovers from previous identities were correlated with the bot type, at each new wave, we randomized which type of bot account would be created using each account (thus, if a specific account represented a White male from a top-ranked university in wave 6, the same account could represent a Black female from a lower-ranked university in wave 8). Every account had an email and a phone number associated with it (for the purpose of Twitter verification).

During each wave, we collected information on follows using Twitter's API. We collected this information twice a day for twelve days counting from the beginning of the treatment.