IZA Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 13475

Biased Beliefs and Entry into Scientific Careers

Ina Ganguli
Patrick Gaulé
Danijela Vuletić Čugalj

JULY 2020

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 13475

Biased Beliefs and Entry into Scientific Careers

Ina Ganguli
University of Massachusetts Amherst

Patrick Gaulé
University of Bath and IZA

Danijela Vuletić Čugalj
CERGE-EI

JULY 2020

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Biased Beliefs and Entry into Scientific Careers*

We investigate whether excessively optimistic beliefs may play a role in the persistent demand for doctoral and post-doctoral training in science. We elicit the beliefs and career preferences of doctoral students through a novel survey and randomize the provision of structured information on the true state of the academic market and information through role models on non-academic careers. One year later, both treatments lead students to update their beliefs about the academic market and impact career preferences. However, we do not find an effect on actual career outcomes 2 years post-intervention.

JEL Classification: I23, D80, D84, J24

Keywords: higher education, information, biased beliefs, career preferences, science

Corresponding author:
Patrick Gaulé
University of Bath
Claverton Down
Bath BA2 7AY
United Kingdom

E-mail: patrickgaule@gmail.com

* We thank Sandro Ambuehl, Michal Bauer, Albert Bravo Biosca, Stepan Jurajda, Nicola Lacetera, Mario Macis, Nikolas Mittag, Henry Sauermann, Paula Stephan, Hubert Wu, Basit Zafar and conference participants at CERGE-EI, the IGL Winter Research meeting, the Swiss Economists Abroad Association, CUNY Queens, and the University of Connecticut for helpful comments. We appreciate research assistance from Samantha Blaney. Gaulé and Vuletić Čugalj acknowledge financial support from the Czech Science Foundation (GACR grant no 16-05082S). Ganguli acknowledges support from the W.E. Upjohn Institute for Employment Research. This study has been approved by the University of Massachusetts IRB and pre-registered in the AEA RCT Registry (AEARCTR-0003212).

I. Introduction

Pursuing a PhD and post-doctoral training are major human capital investments involving several years of effort and substantial foregone earnings. As with earlier human capital investments, the benefits of these post-graduate investments lie in subsequent career opportunities. One such opportunity is the prospect of obtaining a tenure-track faculty position - a job that comes with considerable non-monetary attributes in terms of prestige, autonomy and flexibility, if not with greater pay..

However, becoming a tenure-track faculty member, particularly in the natural sciences in the U.S., has become incredibly difficult. In 2016, approximately 2,700 students graduated with a PhD degree in chemistry, yet there were only 152 advertised openings for chemistry faculty positions in U.S. research-intensive universities.[1] The share of PhDs that become tenure-track faculty are only around 10% or lower in chemistry and in the life and biological sciences (Gaule and Piacentini, 2018; Sauermann and Roach, 2016). Yet, despite the low likelihood of ever becoming faculty and low post-doc salaries, many graduate students pursue one or multiple postdoctoral positions, often with the hopes that it will increase their chances to obtain academic employment (Hayter and Parker, 2019).

The fact that the number of PhD graduates vastly exceeds the number of faculty openings in many STEM fields has not escaped the attention of the science policy community and has been the subject of recurring debates (e.g. Romer, 2000; Freeman et al. 2001; Cyranoski et al. 2011; Schillebeeckx, Maricque, & Lewis, 2013; Alberts et al., 2014; Sauermann and Roach, 2016).

[1] There are more than 200 research-intensive universities in the U.S. Besides being relatively easy to measure, placements in research-intensive universities are precisely those that junior scholars aspiring to an academic career with a focus on research would target. The figure of 152 openings is based on the results of a community effort to help applicants by identifying all relevant positions (see http://chemjobber.blogspot.com/).

Why do young scientists keep choosing to pursue PhD and postdoctoral training despite the dwindling academic career prospects? One possibility is that postdoctoral training improves non-academic career prospects enough to be worthwhile even in the absence of academic career options.[2] However, evidence suggests that non-academic careers vary substantially in the extent that they require doctoral training (Hayter and Parker, 2019). Alternatively, the experience of training itself may be appealing to graduate students, as scientists are drawn to the puzzle-solving nature of doing science (Merton, 1973; Dasgupta and David, 1994; Stern, 2004; Sauermann and Roach, 2012). Meanwhile, for foreigners, visa considerations may steer individuals not just towards graduate study, but also towards postdoctoral training, as universities are not to subject to the same H1-B restrictions as private sector firms, which would allow them to more easily remain in the U.S. (Stephan and Ma, 2005; Ganguli and Gaule, 2020; Amuedo-Dorantes and Delia-Furtado, 2019).

In this paper, we consider another factor that may contribute to observed human capital investment decisions: perhaps graduate students are not well informed about the state of the academic job market, and these incorrect beliefs play a role in their career decisions, particularly decisions to pursue postdoctoral training.[3] Prior studies suggest through qualitative and survey evidence that individuals already in postdoc positions were indeed overly optimistic about the likelihood of getting an academic job, and that junior scientists who had already advanced beyond the PhD reported lacking information about non-academic career options (Sauermann and Roach,

[2] For example, having completed postdoctoral training may have signaling or certification value on the labor market. Further, the knowledge gained through training may be applicable – and indeed highly valued - for working in industry (Aghion, Dewatripont, and Stein, 2008; Dasgupta and David, 1994; Sauermann and Stephan, 2010; Sauermann and Roach, 2016).

[3] Entering science involves a series of choices – from choosing a major in college to deciding to embark on a PhD and post-PhD career choices. Ideally, we would like to know how beliefs and information on the scientific labor market shape decisions to pursue a scientific career at an early stage.

2016; Hayter and Parker, 2019). Yet, whether providing information about the academic market and non-academic careers to graduate students prior to this would have a causal impact on their beliefs and subsequent career choices and preferences remains an open question.

In very different contexts, the economics literature has established that biased beliefs can drive human capital investment decisions and that providing information can causally impact subsequent educational choices (e.g. Jensen, 2010; Oreopoulos and Dunn 2013; Dinkelman and Martinez 2014; Wiswall and Zafar, 2015). In these studies, individuals typically underestimate the returns to education and thus underinvest in education or make sub-optimal education choices.

We focus on post-graduate human capital decisions and ask whether beliefs are biased and whether providing information about the academic and non-academic labor markets can have a causal impact on subsequent education investments and career aspirations, in particular preferences to do a postdoc and pursuing an academic career. We focus on a sample of U.S. chemistry doctoral students at the top 54 U.S. Chemistry departments using an original survey combined with a field experiment.[4] In the baseline survey, we first elicit beliefs about the academic market and publishing in top journals, as well as career preferences for different types of post-graduation jobs, such as postdocs, industry, government or teaching positions.

At the end of the survey, a random subsample of respondents received a message with a link to a custom-built website providing information on actual historical placement records by institution in a tabular format (historical information treatment). This treatment provides structured information about the academic labor market. Another random subsample received a message with a link to a webpage from the American Chemical Society (ACS), the main professional

[4] We focus on chemistry as this is a discipline where we are able to observe academic placements on a systematic and accurate basis thanks to the availability of a faculty directory (the ACS directory of graduate research). No comparable data for exists for biology or physics. However, tight academic labor markets and long postdoctoral training are prevalent across the life and hard sciences.

society for chemists, listing profiles with photos and career information about professional scientists in academic, industry and government positions (role model treatment). This treatment provides less-structured information about both the academic and non-academic labor markets, particularly through role models who work in non-academic sectors, with whom students would have little exposure to during their studies. Such role model interventions through various media types have been shown to impact behavior in a variety of settings, including among STEM students (e.g. Porter and Serra, forthcoming; La Ferrara, 2016).

The last randomly drawn subsample, the control group, did not receive any message. One year after the baseline survey, we conducted a follow-up survey with the respondents of the baseline survey. In order to track how beliefs changed over time and whether the information interventions caused differential adjustments in beliefs, we asked respondents the same questions about their expectations about the academic job market.

Our first result is that at baseline, doctoral students in our sample are excessively optimistic, both about the state of the academic market in their field and about publishing in top journals. When we ask respondents to state their beliefs about the share of peers from their program eventually obtaining a tenure-track position in a U.S. research-intensive university, only a third of respondents have beliefs in the correct range, with the rest being either mildly or widely overoptimistic. Being overly optimistic in turn correlates with stated preferences for doing a postdoc and academic careers more generally.

Interestingly, respondents were more optimistic about their peers' chances of obtaining a tenure-track position in a research-intensive university than about their own chances. Similar to Sauermann and Roach (2016), who show that graduate students in older cohorts are less likely to plan on doing a postdoc and are less interested in academic careers, we find students further along in their programs are less likely to hold overoptimistic beliefs about their chances on the academic job market. Foreign students were more likely to hold overoptimistic beliefs. Female students were more optimistic than male students about the prospects of their peers, but not about their own chances of becoming faculty.

Turning to the experiment, we estimate the causal impact of each information intervention on beliefs and preferences for different careers one year later. We find that both types of information (historical information treatment and role model treatment) led to a downward adjustment in beliefs about respondents' own chances of becoming faculty, particularly among those who had more optimistic initial beliefs. Yet, we observe no significant impact of either type of information on beliefs about the share of graduates from their program eventually becoming faculty.

We also examined impacts of the interventions on satisfaction with doing a PhD. We do not observe an effect of the historical information treatment on satisfaction with pursuing a PhD, but the role model treatment did lead to small decline in satisfaction. Interestingly, we do find that the historical information treatment led to an increase in the perceived attractiveness of an academic career. To the extent that the historical placement information made respondents realize that becoming a faculty member is more difficult than they expected, this may have reinforced the perceived attractiveness of academic careers. The role model treatment, meanwhile, increased the perceived attractiveness of government research and development positions and reduced the preference for doing a postdoc, suggesting that exposure to non-academic career options can impact career preferences.

We also examine longer-run outcomes by collecting data on actual placements for the subsample of students who completed their PhD after the baseline survey two years later. For this sample, we do not see any significant effects in their actual career choices, including doing a postdoc after the PhD.

In sum, we find that the beliefs of graduate students are often biased and providing historically accurate information leads to an adjustment in beliefs, especially among those who had initially higher beliefs. Moreover, providing less structured information about non-academic careers impacts preferences for these careers. Yet, these changes in beliefs lead to limited changes in career aspirations in the longer run, and we do not detect impacts on actual career outcomes. Taken together, these results provide further questions about the role of information in post-graduate human capital investments.

There are several possible reasons for the limited estimated effects on stated career aspirations and actual outcomes. First, it could be that other preferences known to drive scientists' behavior (e.g. puzzle-solving nature of doing science or prestige) are already quite strong at this point in training, so that there was minimal impact of the information on actual career preferences and choices. Second, given the sequential nature of educational choices, and that these are individuals who are already far along in their training trajectory with little option value, switching costs may be high (Stange, 2012). Third, the experience of going through postdoctoral training may be enjoyable in itself or may be desirable for visa or dual-career considerations. Finally, postdoctoral training is still valued in many industry and government positions.

While we cannot differentiate between these explanations in the current study, our findings nonetheless suggest that there is a strong rationale for departments to provide better career information, both about academic and non-academic careers, to prospective and actual students, and there seems to be demand for such information (Sauermann and Roach, 2016). Providing better information would ensure that the choices are made with full knowledge of what they imply, and the costs of collecting and sharing information on placements are low.

In addition to these implications for the post-graduate labor market, this paper contributes to the growing literature on biased beliefs and overconfidence. The prevalence and implications of biased beliefs and overconfidence has been documented across many domains (Malmendier and Taylor, 2015), such as labor supply (Mueller, Spinnewijn, and Topa, 2018), the housing market (Armona, Fuster, and Zafar, 2019), risky behavior (Dupas, 2011) and returns to schooling (Bleemer and Zafar, 2018; Wiswall and Zafar, 2015; Loyalka et al., 2013). Notably, ours is the first study which investigates the existence of biased beliefs in the educational choice to pursue post-graduate studies, postdoctoral studies in particular, and estimates how these beliefs are impacted by the provision of objective information about the labor market.

The paper proceeds as follows. Section II explains the institutional context. Section III describes the data and experimental design. Section IV presents the results, and we end with the discussion in Section V.

II. Institutional Context

In this section, we discuss entry into scientific careers with a specific focus on chemistry and academic careers in the U.S. The entry into scientific careers is characterized by long periods of training. A PhD degree typically takes 6 years and is often followed by one or several postdocs.[5] The chemical and pharmaceutical industry, as well as the government, are major employers of chemistry PhD graduates, and graduates can industry positions before or after postdoctoral training. Despite these human capital investments into becoming a professional researcher, many

[5] In the extreme case, a small but significant proportion of postdocs end up as 'permadocs', doing several subsequent postdoctoral trainings without ever advancing to another level (Powell, 2015).

doctoral degree holders employed in industry do not actually do research in theirs jobs (Lautz et al., 2018).

A necessary condition for becoming a tenure-track professor in chemistry at a research-intensive U.S university is earning a doctoral degree. However, in chemistry and other natural sciences, a postdoctoral training has become *de facto* an additional pre-requisite, with direct transitions from obtaining a PhD degree to a tenure-track position essentially unheard of. In other words, postdoctoral training is crucial for being competitive for faculty positions. As a post-doc, junior scientists build their publication portfolio, apply for grants, and gain additional scientific and professional skills. Yet, the vast majority of postdocs do not become tenure-track faculty members. Around a third of graduate students pursue postdocs, but less than 10% of chemistry graduate students are in a tenure-track position in a research-intensive U.S university 5 years after graduation (Gaule and Piacentini, 2018). Such low odds have been documented in other disciplines and countries (Stephan, 2012b).

Postdocs receive comparatively low levels of compensation during their postdoctoral training. For example, postdocs on average receive 31% lower hourly wage than an average U.S. worker regardless of the education level (Stephan, 2013). The opportunity cost of choosing a 3-year postdoc instead of working in industry was estimated to be around $60,000 in 2012 (Stephan, 2012a). In biomedicine, compared with peers who started working outside academia immediately after finishing their graduate studies, those who finish a postdoc earn less when they actually start to work (Kahn and Ginther, 2017). According to the same authors, postdocs forgo about one-fifth of their earnings potential in the first 15 years after finishing their doctorates, which amounts to more than $200,000.

While information on career prospects for scientists is often available from professional associations and other sources, departments generally provide relatively little career information to prospective and current graduate students. Prior to the launch of this study, we visited the websites of 56 chemistry departments in our sampling frame (see appendix B) looking for their graduate degree holders' placement information. For 70% of departments, we could find no placement information at all. The remainder typically provided examples of institutions that have hired their graduates or aggregate data on placement by broad industry categories. One notable exception was the Princeton chemistry department which provided list of graduates and their placements at the conclusion of PhD. See appendix C for more details on placement information available from departmental websites.

III. Data and Experimental Design

We combine two surveys of chemistry graduate students with a field experiment, linked to the data on individual publications and career choices. The surveys provide rich descriptive data on respondents' beliefs and aspirations and how they evolve over time. To overcome potential hypothetical bias, we combine the data on hypothetical job preferences with real job preferences from hand-collected placement data of the survey respondents who finished their PhD after the baseline survey. We also leverage data from faculty directories, PhD theses and publications from an ongoing project on the production of knowledge in chemistry (see Gaule, 2014; Gaule and Piacentini, 2018; Catalini, Fons-Rosen, and Gaule forthcoming). Our research design and data collection approach is summarized in figure 1.

(Insert Figure 1 about here)

Our analysis and intervention is primarily based upon a survey we conducted in Fall 2017 (hereafter 'baseline survey'). To construct the sampling frame, we first identified the set of 54 research-intensive U.S. universities that rank highest in the Academic Ranking of World Universities (Shanghai Ranking) in its Chemistry subject ranking. These schools have large PhD programs and their students are presumably comparatively better placed for the academic job market. We gathered the names and emails of all individuals (n=9,141) that were listed as graduate students in the chemistry departments of these universities, either on graduate student directory websites or on individual laboratory websites. We then sent them email invites to complete a survey using the Qualtrics online survey platform.[6]

We received a total of 1,330 responses corresponding to a response rate of 15%.[7] The baseline survey included a set of basic demographic questions, as well as questions on undergraduate education, year of enrolment in the PhD program, progress in the PhD program and field of specialization. We asked about career preferences using both standard Likert-scale measures and counterfactual choice questions. Regarding beliefs, we asked respondents to rate their chance of publishing in *Nature*, *Science* or *Cell*, - the most prestigious science journals-, to rate their chance of becoming a tenure-track faculty in a research-intensive university, and the share of students in their program they believe eventually become tenure-track faculty in a research-intensive university (see appendix D for the exact survey questions). Finally, we asked respondents whether they would agree to be contacted in a follow up survey and if so, if they could provide us with a permanent email address that we could use to contact them again. Table A1 in appendix A shows means and standard deviations for several key variables from the baseline survey.

[6] To increase the response rate, we sent two reminder emails and offered a lottery with possibility of winning one of ten Amazon gift certificates worth $100 each. The choice of using this type of lottery was informed by Sauermann and Rauch (2013).

[7] One issue we encountered was that some of the individuals we contacted reported having already graduated, presumably reflecting the fact that some online directories and websites were not entirely up to date. We excluded such responses from our analysis sample. Adjusted for the presence of students who already graduated among the people we contacted, our response rate was around 18%.

We combined the baseline survey with an information provision experiment. After completing the baseline survey, respondents were randomly selected into either two treatment or one control group. Treatment groups received one of the two versions of a thank you message via email with information related to the labor market, while the control group received no message at all.

One of the messages contained more structured information (historical information treatment), which linked to a custom-built website providing information on historical actual academic placement rates by graduate institution in a tabular format.[8] These placement rates were well below 10% for all institutions so the information communicated was mainly an update on the difficulty of becoming a tenure-track faculty in a research department.

The second message included less-structured information about non-academic careers (role model treatment), which linked to a real webpage from the American Chemical Society (ACS), the main professional society for chemistry, called "Chemists in the Real World". This website features pictures with job titles and profiles of professional scientists in academic and (mostly) non-academic positions (see appendix F for the illustration of both websites used in this study).[9] The role model treatment was intended to provide students with information about the both academic and non-academic careers through role models. Such role model interventions through various modes, such as in person, websites, and television, have been shown to impact behavior in

[8] The historical placement records were based on previously collected data from Proquest Dissertations and Abstracts and the ACS directory of graduate research (Gaule and Piacentini, 2018). Specifically, we collected data on students graduating from U.S. chemistry graduate programs between 2008 and 2010, and matched their names to a 2015 list of chemistry faculty in research-intensive universities. We then computed the share of graduating students who had become faculty by 2015, by graduating department. For more, please see the appendix E. We published this data, together with a detailed explanation how the data was constructed on the custom-built website https://chemistryplacementdata.com/. The website was not advertised in any way. Web analytics confirm that the overwhelming majority of visits to the website originated from the survey emails.

[9] Available at: https://www.acs.org/content/acs/en/careers/college-to-career/chemists.html

a variety of settings (e.g. Porter and Serra, forthcoming; see La Ferrara, 2016) by providing exposure to individuals who students otherwise would have little interactions with in their studies.

Not all respondents clicked on the links embedded in either message. While we did not track individual usage, we estimated that roughly 35% of survey respondents who received the link visited the custom-built website, versus around 1% of respondents in the control group (we could not track clicks to the ACS website; see appendix G for details).

The randomization procedure combined block-randomization (stratified based upon a department's Shanghai Ranking) with individual-level randomization in a subset of universities (see Figure 1).[10] In order to measure the impact of the intervention on respondents' beliefs and plans, we contacted our respondents again roughly one year after the baseline survey and asked them to complete a follow up survey.[11] In the follow-up survey we repeated several questions from the baseline survey. We again incentivized responses by sending two reminder e-mails and offering a lottery to win a $100 Amazon gift certificate upon completing the survey. We obtained 500 complete responses, roughly 38% of the initial survey respondents. Table A2 in appendix A reports means and standard deviations for several variables from the follow-up survey. We complemented the follow-up survey with hand-collected information on the current position of baseline survey respondents, such as whether they were doing a postdoc or working in industry (for descriptive statistics, see table A3 in appendix A). This information was collected in the

[10] We created triads of departments of similar ranks, and within each triad assigned one department to the information treatment, one to the control, and one to individual randomization. Thus, one university of 3 in the block was randomly chosen as Treatment 1, so that all respondents to the baseline survey at this university received the first message with historical placement rate information. For the second university, respondents were in the control group. In the final university, survey respondents were individually randomized into one of the three groups (historical information, role model, or control). An advantage of this design is that for the historical information treatment, we have both individuals whose peers were also treated, and individual whose peers were not treated. This randomization design was intended to enable us to measure potential spillovers from the treatment, if the treated individuals share information with their peers. However, sample size limitations prevent us from fully leveraging this aspect of the randomization.

[11] We excluded those who indicated in the first survey not to be contacted again.

summer of 2019, roughly two years after the baseline survey. We collected this information irrespective of whether individuals answered the second survey but only for students who were expecting to graduate in 2017, 2018 or 2019 at the time when they were filling in the baseline survey.

Table A4 shows differences in the characteristics of respondents to our follow-up survey to those who completed the baseline survey only. We see some differences in observable characteristics, as students from higher-ranked programs, foreign students, and students further along in the program were less likely to respond to the follow up compared to those earlier in the program. We estimate all regressions including these controls. Importantly, we do not see differential attrition in the treatment group receiving the historical placement information treatment. We do see a small decline in the group receiving the role model treatment. However, for the actual outcomes collected, we have information for all baseline survey respondents, and therefore attrition is not a concern for those outcomes.

IV. Results

IV. a. Prevalence of biased beliefs

Do graduate students know how difficult it is to publish in the most prestigious scientific journals, and to become a tenure-track faculty member in a research-intensive university? Are individuals overconfident about their own ability, and in particular overestimate their position in the ability distribution?

One way we measure biased beliefs is by eliciting respondents' beliefs about their chances of publishing as a first author in *Nature, Science* or *Cell* before the end of their PhD. When testing the survey, we had been warned that this is a very rare event. Indeed, historically only one in 200 chemistry PhD students reaches this milestone.[12] A group of 1,301 students would thus be expected to collectively generate 6 or 7 first-authored *Nature*, *Science* or *Cell* publications. Yet, by aggregating the beliefs of the respondents, we find that they expect to collectively produce 310 first-authored *Nature*, *Science* or *Cell* publications. Figure 2 plots the distribution of the respondents' beliefs about their chances of publishing in *Nature*, *Science* or *Cell* by the end of their PhD studies.

(Insert Figure 2 about here)

We also asked respondents to rate their own chances of becoming a tenure-track faculty member in a research-intensive U.S. university. The distribution of those beliefs is displayed in Figure 2. In recent years, the share of chemistry PhD students becoming faculty members was around 5%. For instance, in 2016, a listing of chemistry faculty openings listed 152 tenure-track positions in research-intensive U.S. universities while 2,700 students graduated in this same year. Our own calculations, which are based on matching names from comprehensive lists of PhD graduates and faculty members in chemistry departments, suggest a similar rate. Again, the respondents collectively display optimistic beliefs although to a lesser degree than for *Science*/*Nature/Cell* publications. Specifically, if all the beliefs of the respondents were correct, 320 students in our sample would become tenure-track faculty members in research intensive university, while only 66 of them would actually become faculty in Chemistry departments based on historical averages.

We also asked respondents about their *peer* beliefs - their beliefs on what share of PhD students in their programs eventually become tenure-track faculty members. By asking about

[12] Authors' calculations based upon chemistry PhD graduates listed in Proquest and *Nature/Science/Cell* bibliometric data.

others in their program, we focus on information regarding the state of market. By contrast, the beliefs about the own chance to become faculty also incorporates beliefs about one's own ability as well as preferences for the academic career.

(Insert Figure 3 about here)

Interestingly, the mean beliefs about the share of students becoming faculty (24.5%) are actually slightly higher than the mean beliefs about the own chance to become faculty (24%).[13] So, what looked like an above-average effect might be incorrect beliefs about the market as a whole. The distribution of beliefs on the share of peers becoming faculty in research-intensive universities is displayed in Figure 3. While there was some variation across programs, no program had a share higher than 10% in the historic placement data. Slightly less than 30% of the respondents answered between 0% and 10%, and thus essentially had correct beliefs about the state of the market. A further 25% of respondents were mildly optimistic, answering that between 11% and 20% of peers will become faculty. The remainder – 45% of respondents - were wildly optimistic with answers far above the observed average.

In summary, these descriptive statistics suggest that overoptimistic beliefs about publishing and placement are widespread among graduate students. However, we also observe heterogeneity in beliefs, with some individuals having correct beliefs, and others being biased to various extents.

IV. b. Who holds optimistic beliefs?

We now explore descriptively whether the heterogeneity in beliefs can be related to observable characteristics. For this, we regress each of the three types of beliefs on student gender, foreign status, time since enrollment in the program, and a dummy variable for top 10 program (based on the Shanghai Ranking).

[13] As discussed earlier, both aggregate evidence and historical placement data suggest that this share is around 5%.

(Insert Table 1 about here)

Table 1 displays the results. Foreign students are considerably more optimistic about publishing and placement (table 1, columns 1 and 2). Foreign students may be higher ability on average due to a tougher selection to get into U.S. PhD programs (Gaule and Piacentini, 2013). However, they also seem to be less informed about the tightness of the U.S. academic market (table 1, column 3). Perhaps surprisingly, studying at a top 10 school is not associated with more optimistic beliefs.

While the literature has documented gender differences in overconfidence (e.g. Murciano-Goroff, 2019; Niederle and Vesterlund, 2007), we notably find few gender differences in beliefs in our sample. We find that female and male students are equally likely to hold optimistic beliefs about their chances to publish in *Nature*, *Science* or *Cell*. Female students are slightly more optimistic about the aggregate state of the academic market, i.e. their peers' chances of getting a tenure track job (see Figure 5 and A1), but we observe no gender differences in beliefs about one's own chances.

Time since enrollment in the PhD program is a strong predictor of holding optimistic beliefs: Students in their first year or second year of study are the most optimistic, though there is no statistical difference between students in their third year and subsequent years. The results are consistent with Stephan and Ma(2005), Sauermann and Roach (2012), Sauermann and Roach (2016), and Gibbs, McGready, and Griffin, (2015).

(Insert Table 2 about here)

We also investigate whether holding optimistic beliefs about the share of students becoming faculty is associated with preferences for academic careers (cf. Table 2). We measure these preferences by asking how likely respondents are to do a postdoc or to choose a prestigious postdoc vs. an industry research job or a teaching position in a hypothetical choice question.[14] We find that respondents' beliefs about the share of students becoming faculty is strongly correlated with preferences for continuing an academic path. This holds despite the fact that we are controlling for key observable correlates of holding optimistic beliefs, such as being a foreign student or being in the first or second year of study.

As discussed earlier, in this discipline, moving straight from doctoral studies to tenure-track positions is virtually impossible. However, by choosing postdoctoral training, a scientist keeps open the possibility of subsequently landing a tenure-track faculty position, a job that she often perceives to be highly desirable. The option to access this career path, while uncertain and risky, is part of the returns to doing a postdoc. Students who underestimate how difficult it is to obtain a tenure-track faculty position should thus be expected to find the postdoctoral option more attractive, which is exactly what we find.

However, as in previous studies that have documented over-optimism among scientists (e.g. Sauermann and Roach, 2016), these results are descriptive in nature. We cannot rule out that students who have optimistic or biased beliefs may also have other characteristics that drive preferences for doing a postdoc. It is thus unclear whether exogenously inducing updates in the beliefs could lead to changes in career preferences. The next section describes the results of the intervention where we provided information to a random sample of the baseline survey respondents, and then followed up with them one year later.

IV. c. Effects of the intervention

[14] See Appendix D for wording of question.

Our experimental design combined block-randomization at the university level with individual-level randomization for a subset of universities. Accordingly, survey respondents were assigned to one of the following five groups[15]:

(1) <u>Treatment 1 (historical information treatment) – Block Randomization</u>: students received the email linking to the historical information on graduates' placement, along with all other survey respondents from the same university receiving the same link.

(2) <u>Control - Block Randomization</u>: students did not receive any email along with other survey respondents from the same university not receiving any email.

(3) <u>Treatment 1 (historical information treatment) – Individual Randomization</u>: students received the email linking to the historical information on graduates' placement along with only some of respondents from the same university receiving the same link.

(4) <u>Treatment 2 (role model treatment) – Individual Randomization</u>: students received the email linking to the ACS profiles website along with only some respondents from the same university receiving the same link.

(5) <u>Control (Some peers treated) – Individual Randomization</u>: students did not receive an email but some other survey respondents from the same university received the other types of emails (Treatment 1 and 2).

We use the second group – those that did not receive any email with other survey respondents from the same university not receiving any email - as the control group and the omitted

[15] Alternatively, we could pool treatment 1- block randomization and treatment 1 – individual randomization into a single variable. Results from the alternative specification are presented in table A5. The results are qualitatively similar with this alternative specification except for the changes of the beliefs on own chance to become faculty, where the effect of the historical placement intervention is just outside the significance region (p-value-0.11 instead of 0.02 in the preferred specification) but the effect of the ACS profiles intervention is significant.

category in all specifications.[16] Our variables of interest are indicator variables for each of the other categories, or treatments, and we present specifications both, with and without controls.

We first consider the effect of the intervention on beliefs using the sample of students who answered both the initial and final survey one year later. As in the descriptive analysis, we observe two types of beliefs: the beliefs about peers (which share of students in their program become faculty) and the self-beliefs (own chance of becoming faculty). Since we asked the exact same questions on beliefs in the initial and final survey, we can track the evolution of beliefs over time and whether they were impacted by the treatment.

(Insert Table 3 and Table 4 about here)

Table 3 and 4 show the effect of the intervention on the changes in beliefs between the two surveys (final minus initial beliefs). Note that the mean change in either type of beliefs is negative, suggesting that students become more pessimistic over time. The point estimates for the effect of all treatments on beliefs about the share of peers becoming faculty are small and statistically insignificant. However, both the block-randomized historical placement information treatment and the role model treatment had a statistically significant effect on the changes in beliefs on own chances of becoming faculty, where receiving the information lowered beliefs about one's own changes of getting a tenure track faculty position (cf. Table 4). The magnitude of the effect is similar in magnitude to the mean of the dependent variable, suggesting that individuals who received the information became less optimistic about their chances to become faculty members at a faster rate than those who did not. The coefficients on both the individually-randomized historical information treatment and the 'some peers treated' group are smaller in magnitude than for the block-randomized historical information treatment. This is consistent with the effects of

[16] We also estimate the treatment effects of the historical placement information when pooling the block randomized and individually randomized groups. See Table A5.

the historical placement information being amplified when all peers received in the information, rather than only a small subset of individuals, likely by creating more opportunities for discussions that made the information more salient.

It is puzzling that we find an effect of both types of information interventions on self-beliefs, but not on beliefs about peers. Prior to the intervention, we had expected that the intervention might impact both types of beliefs and that if anything, the effect might be weaker for the beliefs of one's own chances.

We next examine whether there was differential response to the treatments in who updated their beliefs. Figure 6 shows that it appears that those with higher initial self-beliefs (those who are most optimistic regarding their own chances of becoming faculty) were more likely to update their beliefs in response to the historical information treatment. Table A6 shows that for both information treatments, the higher the baseline beliefs, the greater the decline in subsequent beliefs. In Table A7, we estimate heterogeneity in response to the treatment by our main covariates: gender, foreign status, and a dummy variable for top 20 program. Here, we see that there are not many significant differences, apart from a larger negative effect of both treatments on the beliefs about peers among foreign students.

Now that we have established that the information treatment did impact beliefs about one's own chances of becoming faculty, we proceed to investigate whether the information interventions impacted career preferences and actual career choices. For the latter, we can also include baseline survey respondents who did not complete the final survey, as we code career choices using publicly available information. Given that the historical placement information intervention led to a downward adjustment in the beliefs on their own chance of becoming faculty, we would expect postdocs to become less desirable in the treatment group (relative to the controls), and fewer people actually choosing postdocs. However, as Table 5 and 6 show, we find no effect of the historical placement information intervention on preferences for doing a postdoc or actually taking up a postdoc position after graduation.[17]

(Insert Table 5 and 6)

As for the role treatment intervention, we see a negative effect on the propensity to choose the postdoc option in the counterfactual choice option, consistent with the role model intervention making non-academic careers more salient and attractive. However, we do not find an effect on actual career outcomes.

(Insert Table 7 and 8)

Finally, we consider the effect of the interventions on additional outcomes: satisfaction with the PhD as a career choice and perceived attractiveness of a faculty position and a government research and development position. Surprisingly, we do not see an effect of either intervention on satisfaction with doing a PhD as a career choice (Table 7). However, the historical placement information did significantly increase the perceived attractiveness of an academic faculty position (Table 8A). To the extent that the historical placement information made respondents realize that becoming a faculty member is more difficult than they expected, this may have -counter-intuitively- reinforced the perceived attractiveness of academic careers. The role model treatment meanwhile increased the perceived attractiveness of a government research and development position (Table 8B). The ACS profiles page lists individuals in government research positions, so this suggests that exposure to these profiles provided information that students previously were not exposed to about government careers, which made them more attractive as potential careers.

[17] This finding echoes Sauermann and Roach (2016) who found in a descriptive analysis no systematic evidence of a relationship between perceived demand for jobs in academia and the choice of postdoctoral training.

V. Discussion

This paper studies the beliefs of science PhD students regarding the academic job market and how these beliefs impact their preferences for different types of careers and their decisions upon graduating using a novel survey of chemistry graduate students combined with the randomized information interventions.

We find considerable evidence that graduate students are excessively optimistic regarding the state of academic job market, their chances to become faculty, and their chances to publish in the very best scientific journals. Students early in the program, and foreigners, are more likely to hold excessively optimistic beliefs. Holding such beliefs is in turn associated with intentions to engage in postdoctoral training after the PhD.

Providing information on historical placement rates and non-academic career options through role models appears to influence beliefs one year later, with treated individuals adjusting their perceived chances of becoming faculty members. We find evidence that the historical information treatment led to an increase in the perceived attractiveness of faculty positions, while the role model treatment increased the perceived attractiveness of government R&D positions and reduced the preference for doing a postdoc. However, we do not observe effects on satisfaction with choosing the PhD as a career choice, nor do we see an effect of the interventions on actual career choices 2-years after the PhD (for a subsample of respondents who had graduated).

Taken together, these results provide further questions about the role of information in post-graduate human capital investments. On the one hand, the beliefs of graduate students are often biased and providing historically accurate information leads to an adjustment in beliefs, especially among those who had initially higher beliefs. On the other hand, the change in beliefs we induced experimentally lead to limited changes in career preferences and aspirations, and we do not detect impacts on actual career outcomes.

There are several possible reasons for the limited effects on stated career aspirations and actual outcomes. First, preferences for postdoctoral training may be quite strong among this group for various reasons. For example, it could be that other preferences known to drive scientists' behavior (e.g. puzzle-solving nature of doing science or prestige) are already quite strong at this point in training, so that there was minimal impact of the information on actual career preferences and choices. Moreover, given the sequential nature of educational choices, and that these are individuals who are already far along in their training trajectory, switching costs may be high. Additionally, the experience of going through postdoctoral training may be enjoyable in itself or may be desirable for visa or dual-career considerations. Finally, postdoctoral training is still valued in many industry and government positions.

Another reason may be due to the types of information we provided. Perhaps a stronger intervention impacting beliefs more strongly would lead to observable changes in actions. Only a minority of individuals who received the link to the historical information treatment actually acquired the information. Given the effects of the role model treatment, information provided directly by the American Chemical Society or the students' own department would give the information more credibility. Additionally, our sample size was relatively limited and having more statistical power would have allowed us to test for further heterogeneity in which types of students responded more or less to the information.

While we cannot differentiate between these explanations in the current study, our findings nonetheless suggest that there is a strong rationale for departments to provide better career information, both about academic and non-academic careers, to prospective and actual students.

References

Aghion, P., Dewatripont, M., & Stein, J. C. (2008). Academic freedom, private-sector focus, and the process of innovation. *The RAND Journal of Economics, 39*(3), 617-635.

Alberts, B., Kirschner, M. W., Tilghman, S., & Varmus, H. (2014). Rescuing US biomedical research from its systemic flaws. *Proceedings of the National Academy of Sciences*, *111*(16), 5773-5777.

Amuedo-Dorantes, C., & Furtado, D. (2019). Settling for academia? H-1B visas and the career choices of international students in the United States. *Journal of Human Resources, 54*(2), 401-429.

Armona, L., Fuster, A., & Zafar, B. (2019). Home Price Expectations and Behavior: Evidence from a Randomized Information Experiment. *Review of Economic Studies*, 86(4): 1371–1410.

Bleemer, Z., & Zafar, B. (2018). Intended college attendance: Evidence from an experiment on college returns and costs. *Journal of Public Economics*, *157*, 184-211.

Catalini, C., Fons-Rosen, C., Gaule, P. (forthcoming). How Do Travel Costs Shape Collaboration? Forthcoming, *Management Science*

Cyranoski, D., Gilbert, N., Ledford, H., Nayar, A., & Yahia, M. (2011). Education: the PhD factory. *Nature news*, *472*(7343), 276-279.

Dasgupta, P., & David, P. A. (1994). Toward a new economics of science. *Research Policy*, *23*(5), 487-521.

Dinkelman, T., & Martínez A, C. (2014). Investing in schooling in Chile: The role of information about financial aid for higher education. *Review of Economics and Statistics*, *96*(2), 244-257.

Dupas, P. (2011). Do Teenagers Respond to HIV Risk Information? Evidence from a Field Experiment in Kenya. *American Economic Journal: Applied Economics* 3(1):1-36

Freeman, R., Weinstein, E., Marincola, E., Rosenbaum, J., & Solomon, F. (2001). Competition and Careers in Biosciences. *Science* 294(5550):2293-4

Ganguli, I. & Gaulé, P., (2020). "Will the US Keep the Best and the Brightest (as Post-docs)? Career and Location Preferences of Foreign STEM PhDs." *In The Role of Immigrants and Foreign Students in Science, Innovation, and Entrepreneurship*. University of Chicago Press.

Gaule, P., 2014. Who comes back and when? Return migration decisions of academic scientists. *Economics Letters, 124*(3):461-464

Gaule, P., & Piacentini, M. (2013). Chinese graduate students and US Scientific Productivity. *Review of Economics and Statistics, 95*(2), 698-701.

Gaule, P. & Piacentini, M. (2018). An advisor like me? Advisor gender and post-graduate careers in science. *Research Policy, 47*(4), pp.805-813.

Hayter, C.S. & Parker, M.A., 2019. Factors that influence the transition of university postdocs to non-academic scientific careers: An exploratory study. *Research Policy*, 48(3), pp.556-570.

Jensen, R. (2010). The (Perceived) Returns to Education and the Demand for Schooling. *The Quarterly Journal of Economics, 125*(2), 515-548.

Kahn, S., & Ginther, D. K. (2017). The impact of postdoctoral training on early careers in biomedicine. *Nature Biotechnology*, *35*(1), 90.

La Ferrara, E., 2016. Mass media and social change: can we use television to fight poverty? *Journal of the European Economic Association*, 14(4), pp.791-827.

Lautz, L. K., D. H. McCay, C. T. Driscoll, R. L. Glas, K. M. Gutchess, A. J. Johnson, & G. D. Millard (2018), Preparing Graduate Students for STEM Careers Outside Academia, *Eos, 99*, https://doi.org/10.1029/2018EO101599.

Loyalka, P., Liu, C., Song, Y., Yi, H., Huang, X., Wei, J., ... & Rozelle, S. (2013). Can information and counseling help students from poor rural areas go to high school? Evidence from China. *Journal of Comparative Economics*, *41*(4), 1012-1025.

Malmendier, U. & Taylor, T., 2015. On the Verges of Overconfidence. *Journal of Economic Perspectives*, 29(4), pp.3-8.

Merton, R. K. (1973). *The Sociology of Science: Theoretical and Empirical Investigations*. University of Chicago Press.

Murciano-Goroff, R. (2019). Missing Women in Tech: The Labor Market for Highly Skilled Software Engineers, Mimeo.

Mueller, A. I., Spinnewijn, J., & Topa, G. (2018). *Job Seekers' Perceptions and Employment Prospects: Heterogeneity, Duration Dependence and Bias*. NBER Working Paper no 25294, National Bureau of Economic Research.

Niederle, M., & Vesterlund, L. (2007). “Do Women Shy Away From Competition? Do Men Compete Too Much?” *The Quarterly Journal of Economics* 122 (3): 1067–1101. https://doi.org/10.1162/qjec.122.3.1067

Oreopoulos, P., & Dunn, R. (2013). Information and College Access: Evidence from a Randomized Field Experiment. *The Scandinavian Journal of Economics*, *115*(1), 3-26.

Porter, C. & Serra, D., (2020). Gender differences in the choice of major: The importance of female role models. *American Economic Journal: Applied Economics*, 12(3), pp.226-54.

Powell, K. (2015). The future of the postdoc. *Nature*, *520*(7546), 144.

Romer, P. M. (2000). Should the Government Subsidize Supply or Demand in the Market for Scientists and Engineers? *Innovation Policy and the Economy*, 1, 221-252.

Sauermann, H., & Roach, M. (2012). Science PhD Career Preferences: Levels, Changes, and Advisor Encouragement. *PloS One*, *7*(5), e36307.

Sauermann, H., & Roach, M. (2013). Increasing Web Survey Response Rates in Innovation Research: An Experimental Study of Static and Dynamic Contact Design Features. *Research Policy 42* (1): 273–86.

Sauermann, H., & Roach, M. (2016). Why pursue the postdoc path? *Science*, *352*(6286), 663-664.

Sauermann, H., & Stephan, P. E. (2010). Twins or Strangers? Differences and Similarities between Industrial and Academic Science. NBER Working Paper No 16113. National Bureau of Economic Research..

Schillebeeckx, M., Maricque, B., & Lewis, C. (2013). The missing piece to changing the university culture. *Nature Biotechnology*, *31*(10), 938.

Stange, K.M., 2012. An Empirical Investigation of the Option Value of College Enrollment. *American Economic Journal: Applied Economics*, 4(1), pp.49-84.

Stephan, P., & Ma, J. (2005). The Increased Frequency and Duration of the Postdoctorate Career Stage. *American Economic Review*, *95*(2), 71-75.

Stephan, P. (2012a). Research efficiency: Perverse incentives. *Nature*, *484*(7392), 29.

Stephan, P. (2012b). *How Economics Shapes Science*. Cambridge, Massachusetts: Harvard University Press.

Stephan, P. (2013). How to Exploit Postdocs. *BioScience*, *63*(4), 245-246.

Stern, S. (2004). Do Scientists Pay to be Scientists? *Management Science*, *50*(6), 835-853.

Wiswall, M., & Zafar, B. (2015). How Do College Students Respond to Public Information about Earnings? *Journal of Human Capital*, *9*(2), 117-169.

Tables and Figures

Figure 1: Experimental design

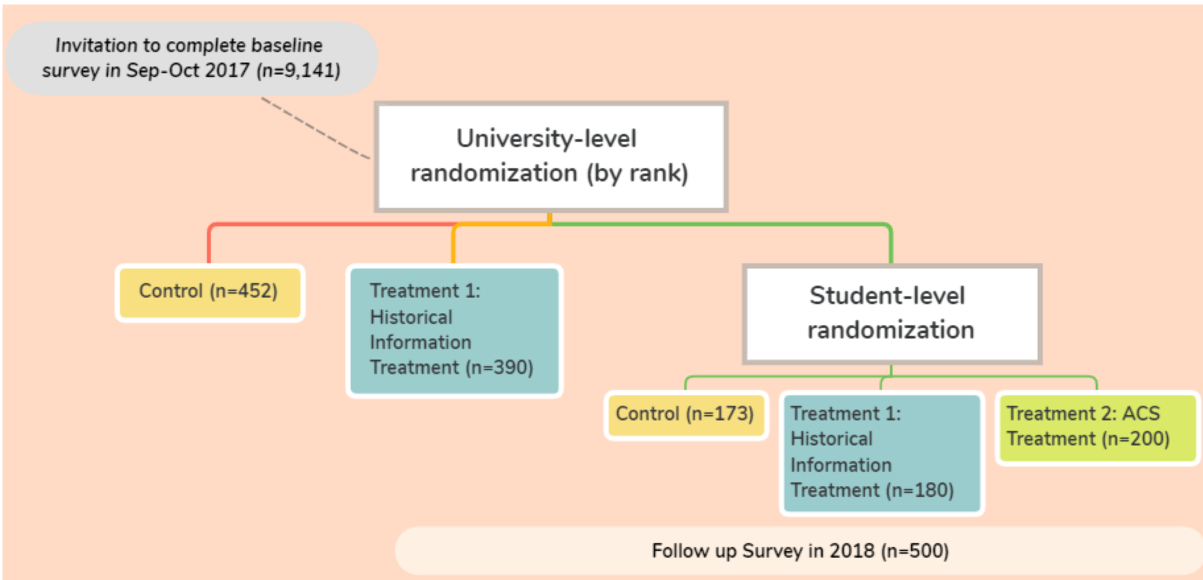


Figure 2: Respondents beliefs about their own chance of publishing in *Nature*, *Science* or *Cell*

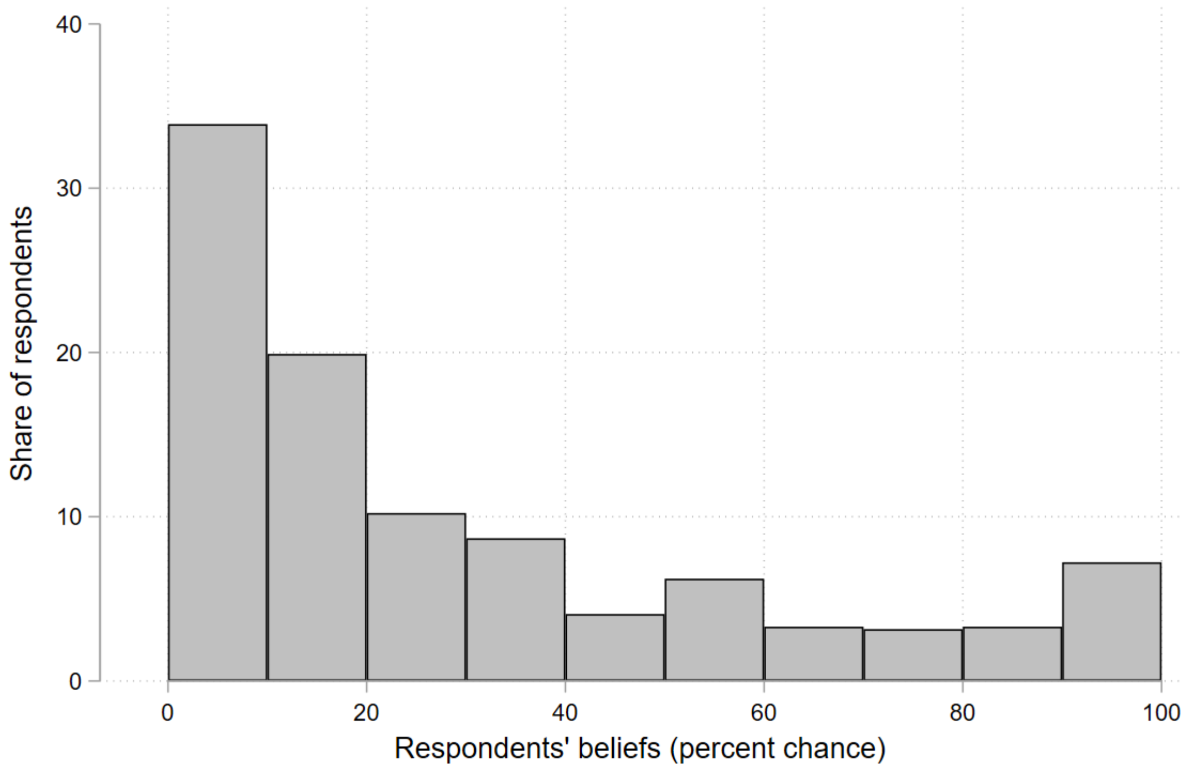


Figure 3: Respondents beliefs about their own chance of becoming faculty

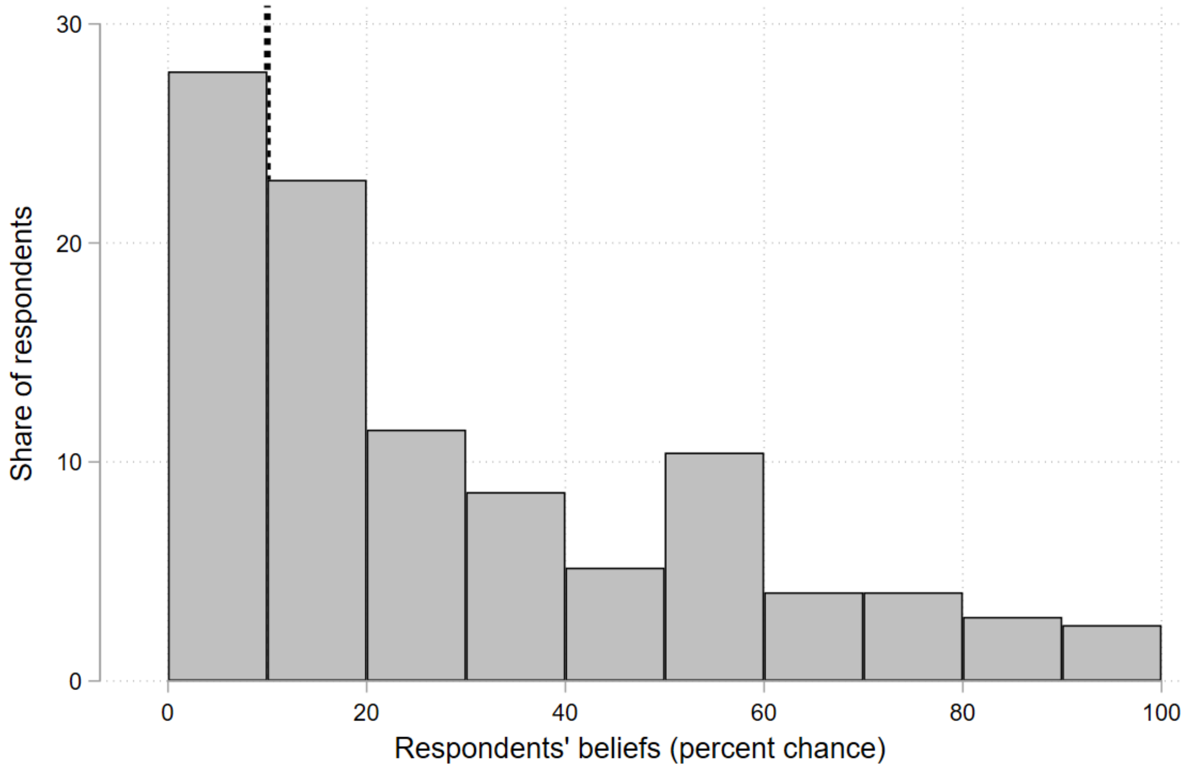


Figure 4: Respondents beliefs about the share of PhD from their program becoming faculty

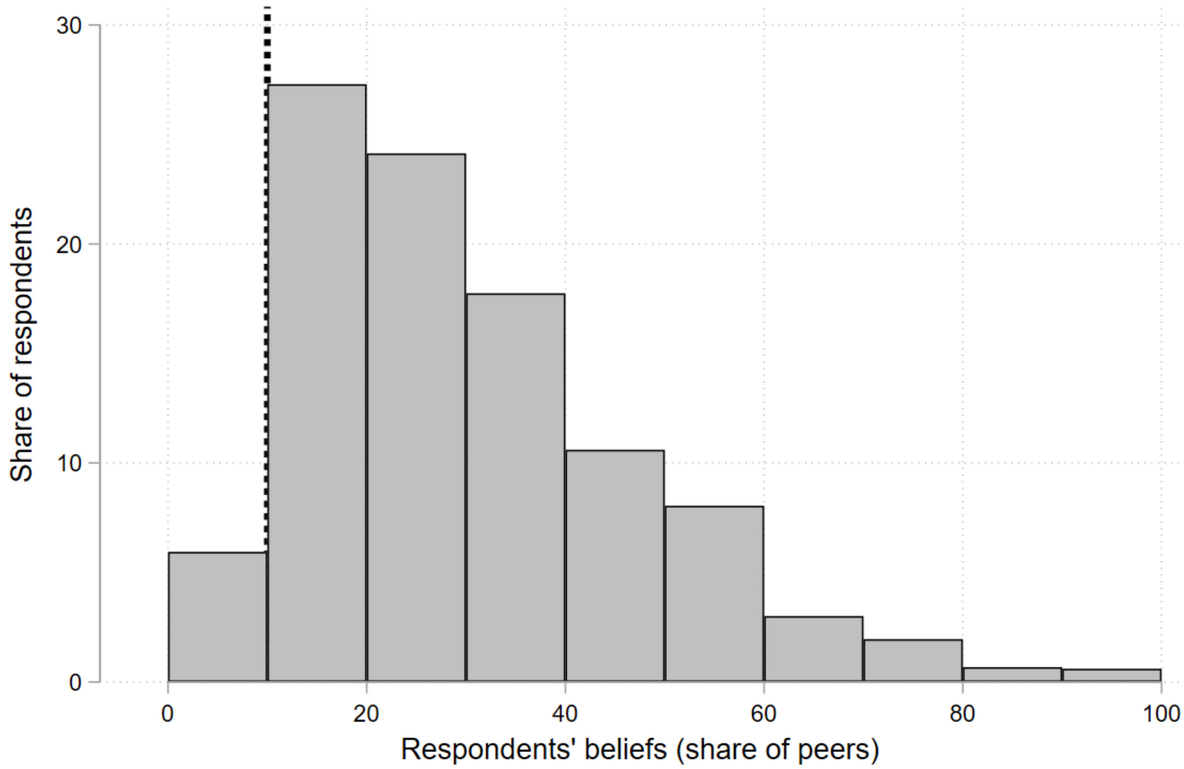


Figure 5. Beliefs of own changes and peers' chances, by Gender

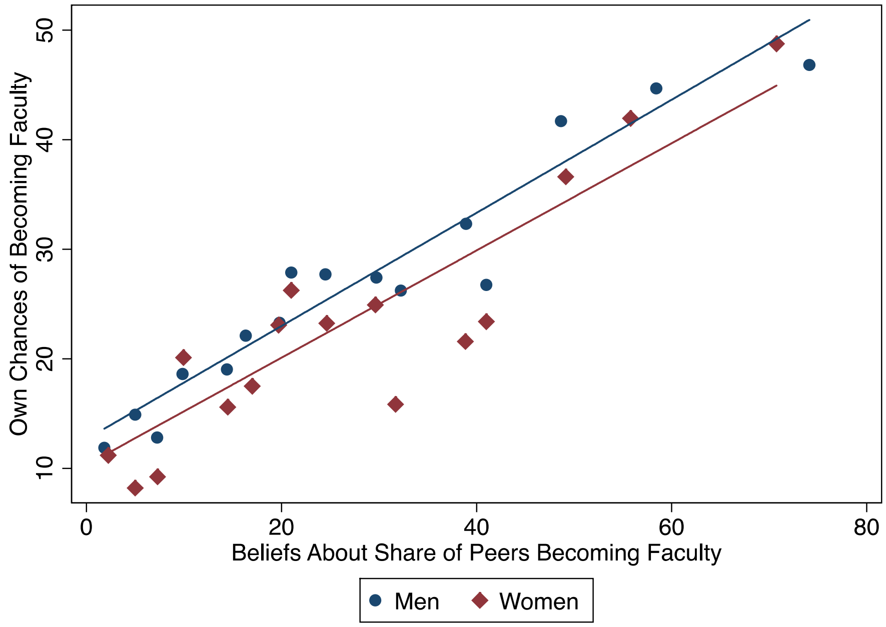


Figure 6. Initial vs. Post-Treatment Beliefs

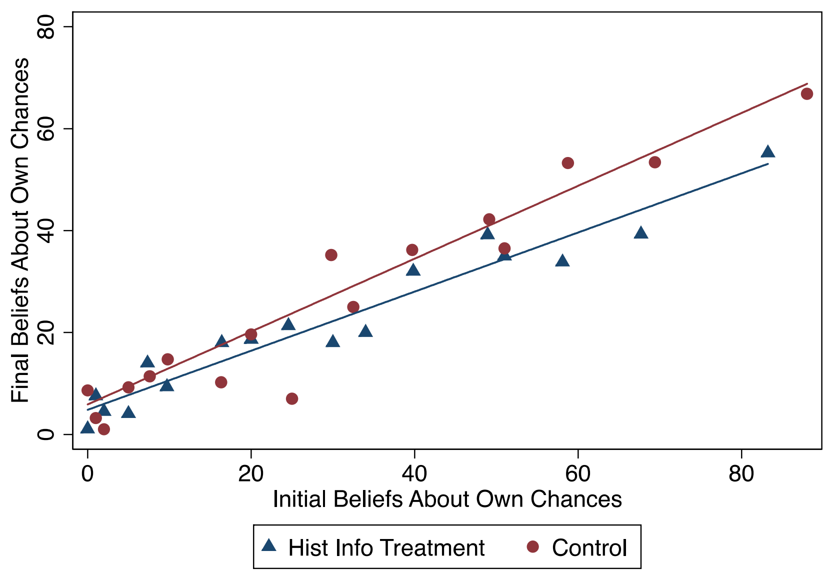


Table 1: Who holds overoptimistic beliefs?

	(1)	(2)	(3)
	D.V.= Respondents' beliefs		
	Own chance to publish in Nature/Science/Cell;	Own chance to become faculty	Percentage of students becoming faculty
Female	0.359	-1.155	2.396**
	(1.616)	(1.380)	(0.971)
Foreign student	9.400***	8.343***	3.798***
	(1.914)	(1.587)	(1.120)
Top 10 school	-1.897	-2.625	-1.349
	(1.969)	(1.679)	(1.181)
First year student	17.753***	9.789***	7.355***
	(2.233)	(1.890)	(1.331)
Second year student	9.512***	6.713***	4.558***
	(2.152)	(1.829)	(1.287)
Third year student	0.767	1.522	1.414
	(2.200)	(1.874)	(1.319)
Obs.	1301	1333	1330
Mean of D.V.	24.907	23.953	24.472
R2	0.073	0.048	0.039

The dependent variables are the respondents' beliefs regarding (1) their chance to publish in *Nature, Science* or *Cell* as a first author by the end of their PhD, (2) their chance to become tenure-track faculty in a U.S. research-intensive universe university, and (3) the percentage of students becoming become tenure-track faculty in a research-intensive universe university in a U.S. university. All the beliefs are expressed on a scale from 0 to 100. The omitted category for time in the program is fourth year and above. Robust standard errors in parentheses.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 2: Optimistic beliefs and preferences for academia

	(1) D.V.= Likelihood of doing a postdoc	(2) D.V.= Choosing postdoc among 3 options
Respondents' beliefs - share of students becoming faculty	0.205***	0.086**
	(0.050)	(0.038)
Female	-2.102	-2.559*
	(1.743)	(1.350)
Foreign student	12.085***	10.575***
	(2.012)	(1.586)
Top 10 school	-1.219	1.747
	(2.139)	(1.640)
First year student	6.000**	5.779***
	(2.401)	(1.878)
Second year student	3.566	3.599**
	(2.298)	(1.801)
Third year student	1.897	-1.419
	(2.383)	(1.832)
Obs.	1271	1312
Mean of D.V.	54.155	25.524
R2	0.055	0.056

The dependent variables are: (1) the likelihood of doing a postdoc as reported in the baseline survey (percentage out of one hundred), and (2) the likelihood (out of 100) of choosing the postdoc when offered a counterfactual choice between a postdoc, research position in industry, or a teaching position (see appendix D). The variable of interest is the respondents' beliefs on the share of students becoming faculty (also out of one hundred). The omitted category for time in the program is fourth year and above. Robust standard errors in parentheses.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 3: Effect of the interventions on beliefs regarding the share of students becoming faculty

	(1)	(2)	(3)
	D.V.= Change in beliefs of the share of students becoming faculty		
Historical placement info treatment (block)	0.008	0.612	-0.416
	(1.664)	(1.619)	(1.394)
Role model treatment	0.938	0.373	0.263
	(2.182)	(2.583)	(2.052)
Historical placement info treatment (individual)	1.184	1.000	0.154
	(2.346)	(2.469)	(2.343)
Some peers treated	1.004	0.239	-0.630
	(2.249)	(2.416)	(1.867)
Obs.	500	500	500
Controls	None	Demographics, field	Demographics, field + Initial beliefs
Mean of D.V.	-3.520	-3.520	-3.520
R2	0.001	0.081	0.374

These regressions are run on the sample of survey respondents who answered both, the initial and follow-up survey. The dependent variable is the change in beliefs on the percentage of students who will become faculty (belief in the final survey minus belief in the initial survey). The coefficients reported correspond to 4 different indicators for each treatment status (see main text for description). The omitted group is the group of survey respondents who did not receive a thank you message in universities where other respondents also did not receive a thank you message. The specification (1) does not include any controls. Specification (2) includes controls for gender, foreign status, field of study, time in the program and university rank. In specification (3) we additionally control for the initial level of beliefs. Clustered standard errors in parentheses. The cluster is a group of three universities of similar rank which was used to stratify the block-randomization.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 4: Effect of the interventions on beliefs regarding own chance to become faculty

	(1)	(2)	(3)
	D.V.= Changes in beliefs of own chance to become faculty		
Historical placement info treatment (block)	-5.995***	-5.002**	-3.071**
	(1.625)	(1.807)	(1.428)
Role model treatment	-5.083*	-6.888**	-5.982**
	(2.624)	(2.655)	(2.213)
Historical placement info treatment (individual)	-2.882	-3.194	-2.015
	(2.402)	(2.559)	(2.959)
Some peers treated	-2.144	-3.540	-2.689
	(2.743)	(2.957)	(2.787)
Obs.	500	500	500
Controls	None	Demographics, field	Demographics, field + Initial beliefs
Mean of D.V.	-3.736	-3.736	-3.736
R2	0.015	0.092	0.273

These regressions are run on the sample of survey respondents who answered both the initial and follow-up survey. The dependent variable is the change in beliefs on the respondents' own chance to become faculty (belief in the final survey minus belief in the initial survey). The coefficients reported correspond to 4 different indicators for each treatment status (see main text for description). The omitted group is the survey respondents who did not receive a thank you message in universities where other respondents also did not receive a thank you message. Specification (1) does not include any controls. Specification (2) includes controls for gender, foreign status, field of study, time in the program and university rank. In specification (3) we additionally control for the initial level of beliefs. Clustered standard errors in parentheses. The cluster is a group of three universities of similar rank which was used to stratify the block-randomization. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 5: Effect of the interventions on post-PhD career choice

	(1)	(2)	(3)
	D.V.= Started a postdoc after PhD		
Historical placement info treatment (block)	0.008	0.018	0.029
	(0.030)	(0.032)	(0.028)
Role model treatment	-0.066	-0.032	-0.026
	(0.045)	(0.042)	(0.041)
Historical placement info treatment (individual)	-0.054	-0.003	-0.007
	(0.047)	(0.043)	(0.050)
Some peers treated	-0.043	-0.007	0.008
	(0.048)	(0.053)	(0.057)
Obs.	574	574	574
Controls	None	Demographics, field	Demographics, field + Initial beliefs
Mean of D.V.	0.181	0.181	0.181
R2	0.006	0.118	0.231

These regressions are run on the sample of survey respondents who as of September 2017 were expecting to graduate in 2017, 2018 and 2019, irrespective of whether they answered the final survey afterwards. The dependent variable is whether the person actually started a postdoc as determined by manual searches. The coefficients reported correspond to 4 different indicators for each treatment status (see main text for description). The omitted group is the survey respondents who did not receive a thank you message in universities where other respondents also did not receive a thank you message. The specification (1) does not include any controls. The specification (2) includes controls for gender, foreign status, field of study, time in the program and university rank. In specification (3) we additionally control for the initial level of beliefs. Clustered standard errors in parentheses. The cluster is a group of three universities of similar rank which was used to stratify the block-randomization. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 6: Effect of the interventions on preference for doing a postdoc

	(1)	(2)	(3)
	D.V.=Change in the propensity to choose postdoc among 3 options		
Historical placement info treatment (block)	-1.872	-0.958	-0.680
	(1.809)	(2.235)	(2.338)
Role model treatment	-3.670	-4.084**	-4.311**
	(2.323)	(1.904)	(1.901)
Historical placement info treatment (individual)	-0.680	-1.191	-1.683
	(3.485)	(3.798)	(3.830)
Some peers treated	-2.778	-2.686	-2.562
	(3.291)	(3.211)	(2.936)
N	500	500	500
Controls	None	Demographics, field	Demographics, field + Initial beliefs
Mean of D.V.	-2.401	-2.401	-2.401
R2	0.005	0.088	0.103

These regressions are run on the sample of survey respondents who answered both the initial and follow-up survey. The dependent variable is the change in the propensity of choosing the postdoc among three options in the counterfactual choice question (see appendix D). The coefficients reported correspond to 4 different indicators for each treatment status (see main text for description). The omitted group is the survey respondents who did not receive a thank you message in universities where other respondents also did not receive a thank you message. The specification (1) does not include any controls. The specification (2) includes controls for gender, foreign status, field of study, time in the program and university rank. In specification (3) we additionally control for the initial level of beliefs. Clustered standard errors in parentheses. The cluster is a group of three universities of similar rank which was used to stratify the block-randomization. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 7: Effect of the interventions on satisfaction with the PhD as a career choice

	(1)	(2)	(3)
	D.V.= Changes in satisfaction with the PhD as a career choice		
Historical placement info treatment (block)	0.281	0.024	0.031
	(0.311)	(0.358)	(0.348)
Role model treatment	-0.648	-0.774	-0.814*
	(0.374)	(0.455)	(0.442)
Historical placement info treatment (individual)	0.006	-0.075	-0.068
	(0.535)	(0.583)	(0.535)
Some peers treated	0.714**	0.410	0.351
	(0.333)	(0.326)	(0.288)
N	496	496	496
Controls	None	Demographics, field	Demographics, field + Initial beliefs
Mean of D.V.	2.613	2.613	2.613
R2	0.016	0.084	0.106

These regressions are run on the sample of survey respondents who answered both the initial and follow-up survey. The dependent variable is the change in respondents' satisfaction with choosing PhD as career track (satisfaction reported in the final survey – satisfaction in reported in the baseline survey). The coefficients reported correspond to 4 different indicators for each treatment status (see main text for description). The omitted group is the survey respondents who did not receive a thank you message in universities where other respondents also did not receive a thank you message. The specification (1) does not include any controls. The specification (2) includes controls for gender, foreign status, field of study, time in the program and university rank. In specification (3) we additionally control for the initial level of beliefs. Clustered standard errors in parentheses. The cluster is a group of three universities of similar rank which was used to stratify the block-randomization. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 8A: Effect of the interventions on perceived attractiveness of faculty position

	(1)	(2)	(3)
	D.V.= Changes in the attractiveness of faculty positions		
Historical placement info treatment (block)	0.237***	0.298***	0.298***
	(0.077)	(0.094)	(0.088)
Role model treatment	0.102	0.129	0.132
	(0.166)	(0.181)	(0.188)
Historical placement info treatment (individual)	0.196*	0.216*	0.214*
	(0.111)	(0.110)	(0.111)
Some peers treated	0.081	0.151	0.154
	(0.190)	(0.204)	(0.210)
N	500	500	500
Mean of D.V.	-0.288	-0.288	-0.288
R2	0.009	0.089	0.096

Table 8B: Effect of the interventions on perceived attractiveness of gov't R&D position

	(1)	(2)	(3)
	D.V.= Changes in the attractiveness of gov't R&D positions		
Historical placement info treatment (block)	0.025	0.003	0.029
	(0.092)	(0.118)	(0.119)
Role model treatment	0.184	0.259*	0.272**
	(0.119)	(0.128)	(0.128)
Historical placement info treatment (individual)	0.136	0.177	0.173
	(0.164)	(0.174)	(0.171)
Some peers treated	0.168	0.170	0.191
	(0.136)	(0.153)	(0.159)
N	500	500	500
Mean of D.V.	-0.084	-0.084	-0.084
R2	0.006	0.056	0.076

These regressions are run on the sample of survey respondents who answered both the initial and follow-up survey. The dependent variable is the change in perceived attractiveness of faculty positions for Panel A and of government R&D positions for Panel B (reported attractiveness in the final survey minus reported attractiveness in the initial survey). Attractiveness is measured on a 1-5 scale. The coefficients reported correspond to 4 different indicators for each treatment status (see main text for description). The omitted group is the survey respondents who did not receive a thank you message in universities where other respondents also did not receive a thank you message. The specification (1) does not include any controls. The specification (2) includes controls for gender, foreign status, field of study, time in the program and university rank. In specification (3) we additionally control for the initial level of beliefs. Clustered standard errors in parentheses. The cluster is a group of three universities of similar rank which was used to stratify the block-randomization. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Appendix Tables and Figures

Appendix A: Descriptive statistics and covariate balance

Table A1: Descriptive Statistics on baseline survey respondents (n=1,330)

	Mean	S.D.
Chances of publishing in Nature/Science/Cell	24.91	29.90
Chance of becoming TT faculty in a U.S. research intensive university	24.47	17.76
Share of students becoming faculty in U.S. research-intensive university	23.95	25.38
Likelihood of doing a postdoc	54.13	31.32
Likelihood of choosing postdoc among three options	25.52	24.75
Female	0.42	0.49
Foreign	0.28	0.45
Top 10 school	0.20	0.40
Year in doctoral program:		
First year	0.19	0.39
Second year	0.21	0.40
Third year	0.19	0.40
Field of study:		
Analytical Chemistry	0.11	0.32
Biological/Biochemistry	0.18	0.38
Inorganic Chemistry	0.16	0.37
Medical/Clinical/Pharmaceutical Chemistry	0.01	0.12
Organic Chemistry	0.18	0.38
Physical Chemistry	0.16	0.36
Polymer Chemistry	0.04	0.20
Theoretical/Computational Chemistry	0.07	0.25
Other	0.09	0.28
Obs.	1330	

Table A2: Descriptive Statistics on final survey respondents (n=500)

	Mean	S.D.
Change in beliefs on the share of students becoming faculty	-3.52	15.70
Changes in beliefs on own chance to become faculty	-3.74	20.28
Historical placement info treatment (block)	0.31	0.46
Role model treatment	0.12	0.33
Historical placement info treatment (individual)	0.12	0.33
Some peers treated	0.12	0.33
Female	0.47	0.50
Foreign	0.17	0.38
Top 10 school	0.25	0.43
Year in doctoral program:		
First year	0.21	0.40
Second year	0.28	0.45
Third year	0.22	0.41
Field of study:		
Analytical Chemistry	0.11	0.32
Biological/Biochemistry	0.17	0.38
Inorganic Chemistry	0.17	0.37
Medical/Clinical/Pharmaceutical Chemistry	0.01	0.12
Organic Chemistry	0.17	0.38
Physical Chemistry	0.17	0.38
Polymer Chemistry	0.04	0.20
Theoretical/Computational Chemistry	0.07	0.26
Other	0.07	0.26
Obs.	500	

Table A3: Descriptive Statistics on sample with actual placement data (n=574)

	Mean	S.D.
Started a postdoc	0.18	0.39
Change in beliefs on the share of students becoming faculty	0.29	0.45
Changes in beliefs on own chance to become faculty	0.16	0.36
Historical placement info treatment (block)	0.12	0.32
Role model treatment	0.13	0.34
Female	0.44	0.50
Foreign	0.26	0.44
Top 10 school	0.20	0.40
Year in doctoral program:		
First year	0.01	0.10
Second year	0.04	0.20
Third year	0.38	0.49
Field of study:		
Analytical Chemistry	0.11	0.32
Biological/Biochemistry	0.17	0.38
Inorganic Chemistry	0.18	0.38
Medical/Clinical/Pharmaceutical Chemistry	0.02	0.12
Organic Chemistry	0.17	0.38
Physical Chemistry	0.14	0.35
Polymer Chemistry	0.06	0.23
Theoretical/Computational Chemistry	0.07	0.25
Other	0.08	0.27
Obs.	574	

Table A4: Is there differential selection into the follow-up survey?

	(1)	
	Responded follow-up survey	
Historical placement info treatment (block)	-0.044	(0.034)
Role model treatment	-0.094^{**}	(0.042)
Historical placement info treatment (individual)	-0.044	(0.045)
Some peers treated	-0.058	(0.044)
Foreign student	-0.147^{***}	(0.031)
Female	0.022	(0.027)
Top 10 school	0.091^{***}	(0.033)
First year student	0.127^{***}	(0.036)
Second year student	0.194^{***}	(0.035)
Third year student	0.128^{***}	(0.036)
Field study:		
Analytical Chemistry	0.020	(0.050)
Biological/Biochemistry	0.006	(0.044)
Inorganic Chemistry	0.007	(0.044)
Medical/Clinical/Pharmaceutical Chemistry	0.022	(0.113)
Physical Chemistry	0.044	(0.045)
Polymer Chemistry	-0.012	(0.070)
Theoretical/Computational Chemistry	0.043	(0.058)
Other	-0.042	(0.054)
Constant	0.322^{***}	(0.043)
Obs.	1330	
Mean of D.V.	0.375	

Organic chemistry excluded.
Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A5: Effects of the interventions pooling the historical placement info treatment into one variable

	(1)	(2)	(3)	(4)	(5)
	Change in beliefs of the share of students becoming faculty	Changes in beliefs of own chance to become faculty	Started a postdoc after PhD	Changes in satisfaction with the PhD as a career choice	Changes in the attractiveness of faculty positions
Historical placement info treatment (block + individual)	-0.245	-2.761	0.018	-0.081	0.261***
	(1.453)	(1.663)	(0.028)	(0.347)	(0.056)
Role model treatment	0.254	-6.012**	-0.025	-0.710	0.133
	(2.045)	(2.194)	(0.041)	(0.410)	(0.184)
Some peers treated	-0.640	-2.712	0.009	0.333	0.145
	(1.862)	(2.750)	0.056	(0.248)	(0.189)
Obs.	500	500	574	496	500
Controls	Demographics, field + Initial beliefs	Demographics, field + Initial beliefs	Demographics, field + Initial beliefs	Demographics, field + Initial beliefs	Demographics, field + Initial beliefs
Mean of D.V.	-3.520	-3.736	0.181	2.613	-0.288
R2	0.374	0.273	0.230	0.171	0.129

These regressions correspond to the column of tables 3-7 except that Historical placement info treatment (block) and Historical placement info treatment (individual) are pooled instead of being entered separately. The omitted group is the group of survey respondents who did not receive a thank you message in universities where other respondents also did not receive a thank you message. All specification control for gender, foreign status, time in the program, university rank and the initial level of beliefs. Clustered standard errors in parentheses. The cluster is a group of three universities of similar rank which was used to stratify the block-randomization.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A6. Heterogeneity: effects of the interventions on peer and own beliefs by baseline beliefs

	(1) Change in beliefs of the share of students becoming faculty	(2) Changes in beliefs of own chance to become faculty
Historical placement info treatment (block)	3.498	-0.049
	(3.075)	(2.092)
Role model treatment	-2.419	-0.781
	(2.805)	(2.650)
Historical placement info treatment (individual)	5.823	-3.320
	(4.352)	(3.164)
Some peers treated	6.758**	-3.037
	(2.823)	(2.838)
Historical placement info treatment (block) x Baseline beliefs	-0.161	-0.153*
	(0.157)	(0.081)
Role model treatment x Baseline beliefs	0.117	-0.233**
	(0.118)	(0.089)
Historical placement info treatment (individual) x Baseline beliefs	-0.205	0.028
	(0.170)	(0.072)
Some peers treated x Baseline beliefs	-0.272*	0.023
	(0.146)	(0.114)
N	500	500
Mean of D.V.	-3.520	-3.736
R2	0.351	0.263

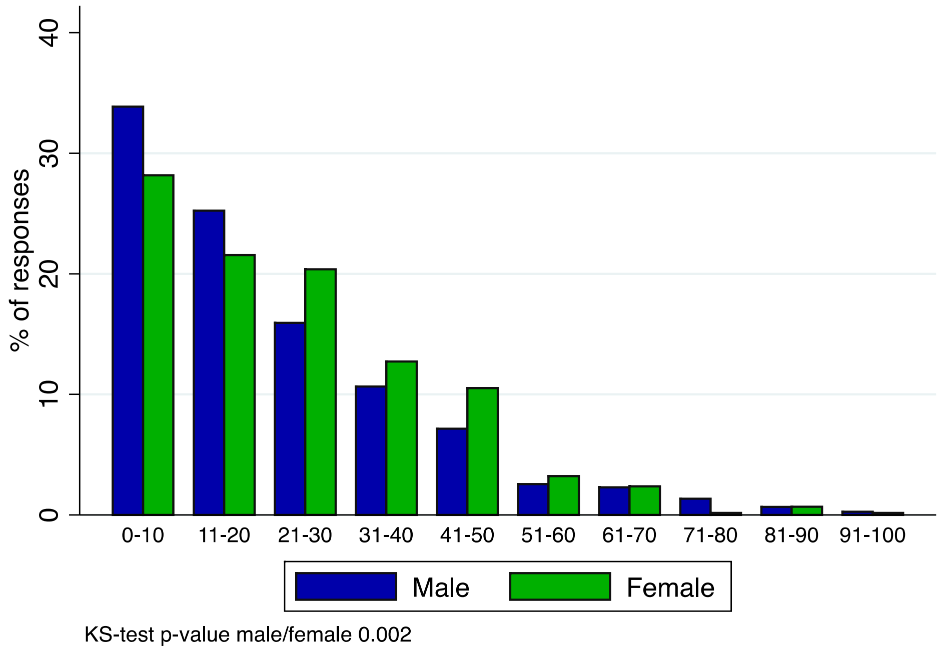
All specification control for gender, foreign status, time in the program, university rank and the initial level of beliefs. Clustered standard errors in parentheses. The cluster is a group of three universities of similar rank which was used to stratify the block-randomization.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A7: Heterogeneity: effects of the interventions on peer and own beliefs

	(1)	(2)	(3)	(4)	(5)	(6)
	Change in beliefs of the share of students becoming faculty			Changes in beliefs of own chance to become faculty		
Covariate →	**Female**	**Foreign**	**Top Univ.**	**Female**	**Foreign**	**Top Univ.**
Historical placement info treatment (block)	0.951	2.127	1.796	-7.061***	-4.784**	-1.857
	(1.967)	(1.409)	(3.872)	(1.438)	(1.698)	(2.928)
Role model treatment	1.891	1.297	1.886	-2.481	-6.478**	-5.600
	(2.520)	(2.805)	(5.928)	(2.981)	(2.672)	(4.711)
Historical placement info treatment (individual)	-2.088	1.498	-2.837	-2.272	-2.273	5.261
	(3.429)	(2.756)	(4.154)	(5.055)	(2.481)	(3.582)
Some peers treated	0.119	3.870*	1.166	-1.200	-2.286	-0.792
	(2.861)	(2.118)	(3.159)	(2.810)	(4.038)	(2.549)
Historical placement info treatment (block) x Covariate	-0.416	-7.759*	-1.444	4.495	-1.516	-3.991
	(3.858)	(3.948)	(4.176)	(3.522)	(7.078)	(3.995)
Role model treatment x Covariate	-4.322	-7.278**	-2.823	-9.201	-0.387	-2.144
	(3.752)	(3.006)	(6.441)	(5.585)	(5.730)	(5.593)
Historical placement info treatment (individual) x Covariate	6.212	-0.155	4.921	-1.655	-7.616	-11.503**
	(5.117)	(16.610)	(5.047)	(7.842)	(9.300)	(4.304)
Some peers treated x Covariate	2.337	-12.565	0.163	-5.043	-5.137	-4.173
	(4.722)	(7.954)	(4.256)	(6.729)	(9.168)	(5.328)
N	500	500	500	500	500	500
Mean of D.V.	-3.520	-3.520	-3.520	-3.736	-3.736	-3.736
R2	0.067	0.071	0.064	0.098	0.088	0.096

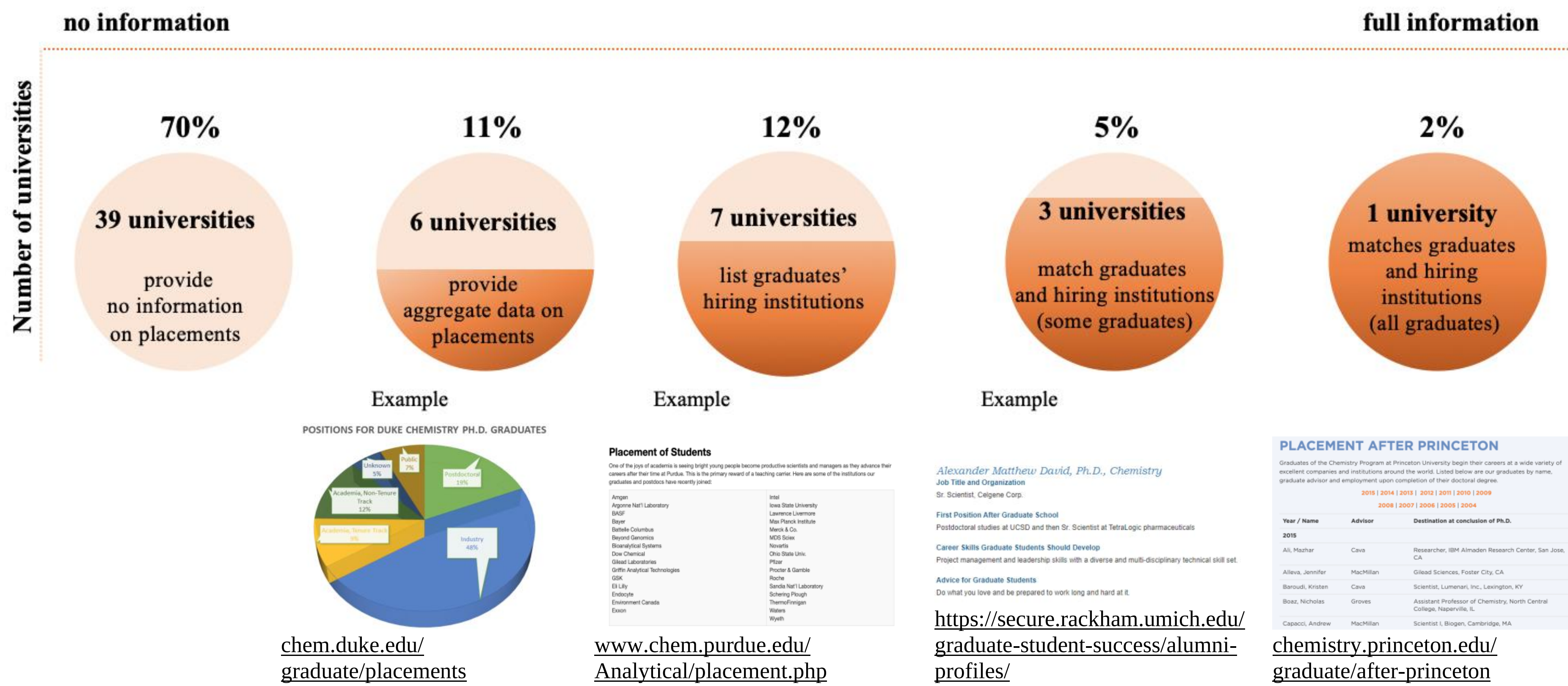
Figure A1. Gender differences in beliefs about the share of PhD from their program becoming faculty



Appendix B: Universities included in the sampling frame

Arizona State University
California Institute of Technology
Carnegie Mellon University
Colorado State University
Columbia University
Cornell University
Duke University
Emory University
Georgia Institute of Technology
Harvard University
Indiana University
Iowa State University
Johns Hopkins University
Massachusetts Institute of Technology
North Carolina State University
Northwestern University
Princeton University
Purdue University
Rice University
Stanford University
State University of New York at Buffalo
Texas A&M University
The Ohio State University
The Pennsylvania State University
The University of Texas at Austin
University of California, Berkeley
University of California, Davis
University of California, Irvine
University of California, Los Angeles
University of California, Riverside
University of California, San Diego
University of California, Santa Barbara
University of Chicago
University of Colorado
University of Delaware
University of Florida
University of Houston
University of Illinois at Urbana-Champaign
University of Maryland, College Park
University of Massachusetts Amherst
University of Michigan
University of Minnesota
University of North Carolina at Chapel Hill
University of Pennsylvania
University of Pittsburgh
University of South Florida
University of Southern California
University of Utah
University of Virginia
University of Washington
University of Wisconsin-Madison
Washington State University
Washington University in St. Louis
Yale University

Appendix C: Information on graduates' placements from university webpages[18]



---

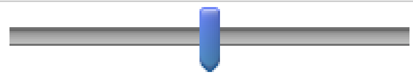
[18] We visited websites of 56 U.S. chemistry research-intensive universities in search for the information they publish on their graduates' placements. We looked through their graduate studies' main pages, graduate student handbooks, career pages, alumni profiles, and news section.

Appendix D: Selected Survey Questions

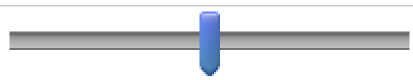
<u>Measuring beliefs about the academic job market</u>

Q. What do you think is the percent chance (or chances out of 100) that you will eventually have a tenure-track position in a U.S. research-intensive university?

	Not likely	Somewhat likely	Very likely

	0 10 20 30 40 50 60 70 80 90 100
How likely you will have a tenure-track position in the US? ()	

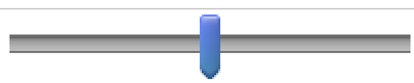
Q. Approximately what share of PhD graduates from your PhD program do you think eventually obtain a tenure-track position in a US research-intensive university? (0 means “None” and 100 means “All”).

	0 10 20 30 40 50 60 70 80 90 100
Share of students with a tenure-track position in the US ()	

<u>Measuring beliefs about postdoctoral training</u>

Q. What do you think is the percent chance (or chances out of 100) that you will do a postdoc after your PhD?

	Not likely	Somewhat likely	Very likely

	0 10 20 30 40 50 60 70 80 90 100
How likely are you to do a postdoc? ()	

Measuring career preferences – counterfactual choice question

Q. Now we want to ask you to do some simple evaluations of potential job offers. Imagine that you have just completed your dissertation and are looking for **a full-time position**.
First, suppose you have the following job offers and you need to choose between them. Please rate how likely you are to accept one of them rather than the other. For each job offer, choose the percent chance (out of 100) of choosing each one. **The total chances given to each offer should add up to 100.**

_______ **Job Offer #1:** Research Scientist/Engineer at Private Sector Firm (e.g. DuPont, Novartis) **Annual Salary:** $90,000 (1)

_______ **Job Offer #2:** Postdoctoral Research Fellow at Top U.S. university (e.g. Berkeley, MIT) **Annual Salary:** $50,000 (2)

_______ **Job Offer #3:** Assistant Professor at top liberal arts college (e.g. Swarthmore College) **Annual Salary:** $70,000 (3)

Q. Putting job availability aside, how attractive do you personally find each of the following careers?

	Not at all attractive (1)	Mostly not attractive (2)	Neutral (3)	Mostly attractive (4)	Very attractive (5)
Academic faculty with an emphasis on research (1)	○	○	○	○	○
Academic faculty with an emphasis on teaching (2)	○	○	○	○	○
Government research and development position (3)	○	○	○	○	○
Government (other) (6)	○	○	○	○	○
Industry position with an emphasis on research and development (4)	○	○	○	○	○
Industry (other) (5)	○	○	○	○	○

Appendix E: measuring historical placement rates

Overview

The objective of this data collection effort was to understand what share of PhD graduates from U.S. chemistry departments become faculty members themselves (in research-intensive universities), and differences across schools. To reach this objective, we collected data on students graduating from U.S. chemistry graduate programs between 2008 and 2010, and matched their names to a 2015 list of chemistry faculty in research-intensive universities. We then computed the share of graduating students who had become faculty by 2015, by graduating department.

Data sources

The database "Proquest Dissertations and Abstracts" was used to obtain the list of chemistry dissertations completed between 2008 and 2010. Proquest Dissertations and Abstracts includes the names of students, the year and university of graduation as well as a subject classification for the thesis, among other information. While the database itself is generally thought to be quite comprehensive, it does not clearly indicate from which department the student graduated. This implies that one must deduce whether it was a chemistry dissertation from the subject classification.

For lists of chemistry faculty, we relied on the "ACS Directory of Graduate Research" available online at dgr.rints.com. This resource, meant to help prospective graduate students choose a graduate program, has an extensive listing of faculty members in U.S. PhD-granting chemistry, chemical engineering and biochemistry programs. The ACS Directory of Graduate Research was used to create a list of faculty members in U.S. research intensive universities, where research intensive is defined as "R1" or "R2" in the Carnegie classification.

An important limitation is that it does not list faculty members outside the U.S. as well as in non-chemistry departments where PhD chemistry graduates may find employment as university faculty with a focus on research.

Matching

The list of graduate students was matched to the list of faculty using last names, initials, first names, year of graduation and university of graduation. The matching algorithm is robust enough to handle cases of variations in spelling of first names, inconsistent reporting of middle names or individuals changing last names.

Limitations of the placement data

The placement data presented here have a few important limitations.

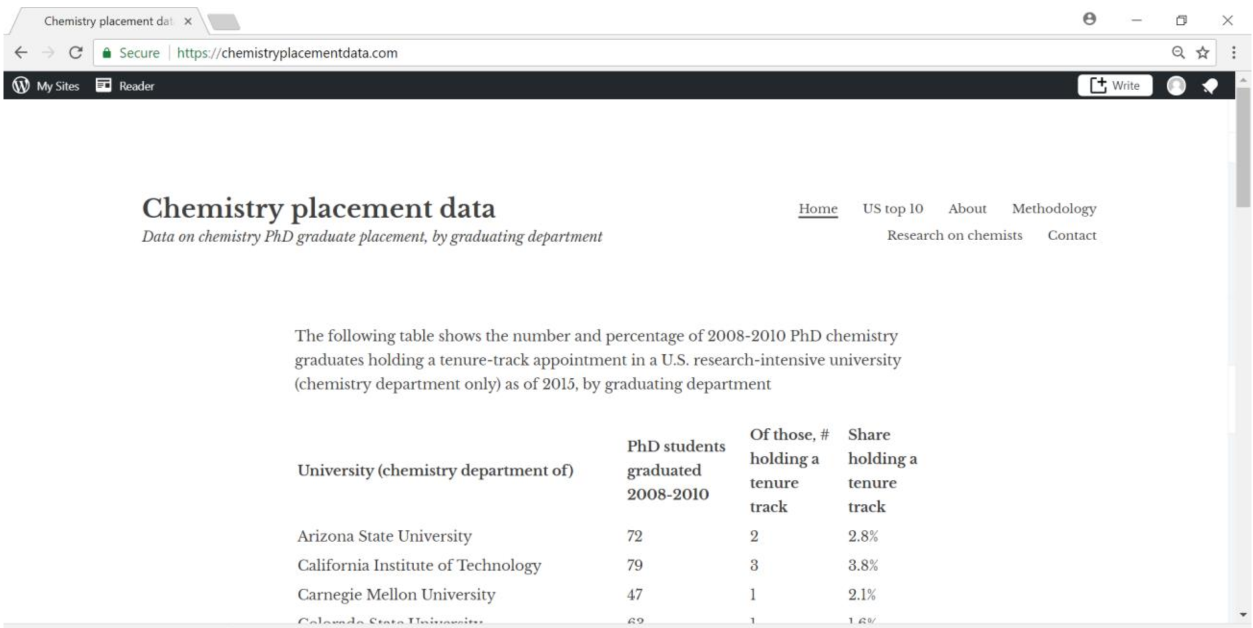
First, some truncation bias arises from the fact that faculty placements are observed as of 2015, while the list of students include students who graduated relatively recently (say 2010) and may have obtained a faculty position in 2016 or 2017, or may obtain a faculty position in the future.

Second, the placement data fails to capture placement in non-chemistry departments that may employ chemistry PhD students, as well as placements outside the U.S.

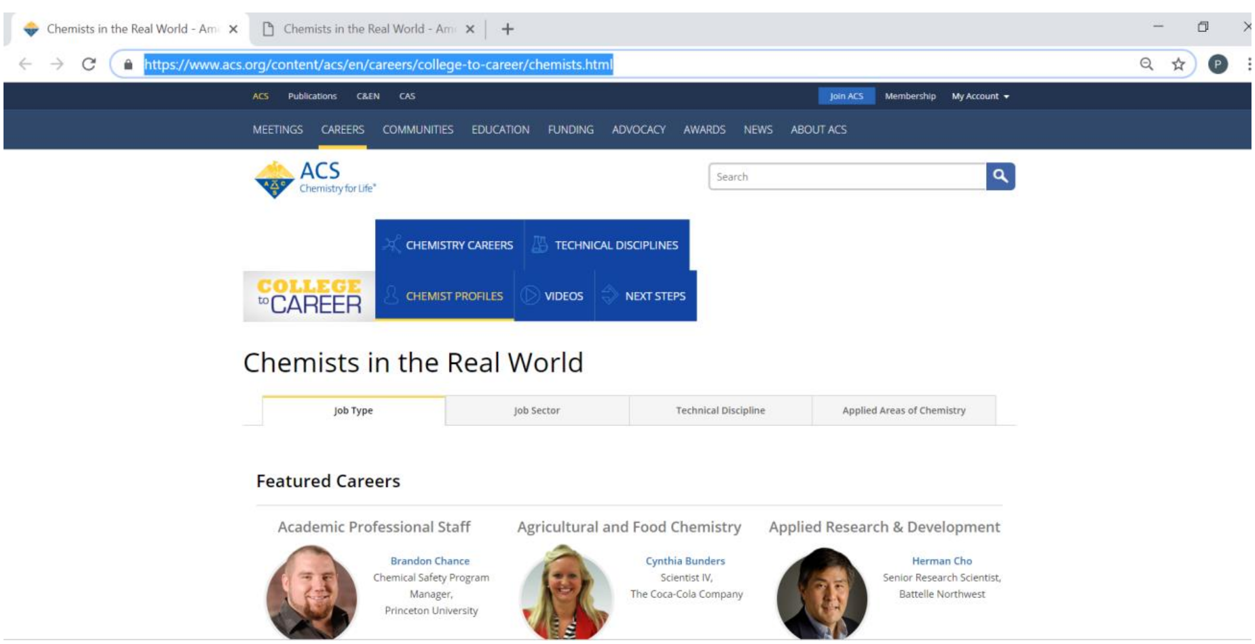
Third, students outside chemistry departments may be mistakenly assigned to the chemistry department if the subject classification of their thesis is close to chemistry; and this could impact the placement measures.

Appendix F: Websites linked in the thank you emails

<u>Custom-built website with historical placement information</u>

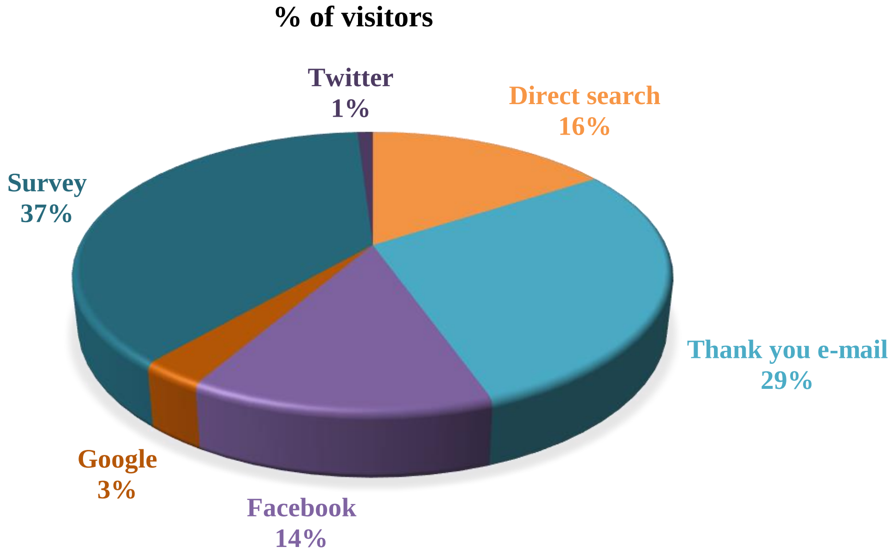


<u>American Chemical Society "Chemists in the Real World" website listing profiles of professional scientists in both academic and industry occupations</u>



Appendix G: Web analytics on visits to the website with historical placement information

Figure 1G: Share of visitors accessing the website: www.chemistryplacementdata.com by source



assessed

Figure 2G: Share of respondents who visited website according to treatment status