Better in Person? The Effects of In-Person Screening on Hiring Outcomes*

John M. Barrios, Laura Giuliano and Andrew J. Leone†

October 2021

Abstract

How are hiring decisions affected by a reduction in the cost of in-person job screening? In theory, this innovation could improve both efficiency and equity by reducing employers' use of stereotypes and imperfect quality signals (e.g., educational pedigree). But since greater weight might be placed on attributes like speech, appearance, or social skills, biases could be introduced or magnified. We examine the introduction of a labor-market intermediary, the Accounting Rookie Camp ("ARC"), that greatly facilitated in-person screening in the academic market for PhD accountants. Using 11 years of data on the supply, demand, and market outcomes for new PhDs, we estimate models that leverage variation in the timing of ARC adoption across both recruiting and degree-granting institutions. We find that ARC adoption reduced the importance of degree-school rank and adviser connections for obtaining a high-quality job, without lowering the bar for research productivity. However, ARC's equalizing effect occurred only within the predominant demographic group: males with English-sounding names. Between groups, ARC penalized candidates with non-English names and exacerbated placement gaps by gender. It also created a premium to physical attractiveness.

Keywords: Job Matching, Screening, Signaling, Hiring, Networks, In-Person Interview, Academic Labor Market, Gender Bias, Immigrant Bias.
JEL Classification Numbers: D83, J7, J23, J44, M510.
Note: The analysis we report in Section 6 is very preliminary, and pending additional review of our data. Please do not circulate this draft.

*We thank Peter Blair, Micheal Gibbs, Abbie Smith, Jack Mountjoy, Thomas Wollmann, and workshop participants at INSEAD, National University of Singapore, Pomona College, University of Chicago Booth, UC Santa Cruz, Georgia State University, Yale SOM, Rice, and University of Minnesota. We thank Jin Deng for outstanding research assistance.

†Barrios: Washington University in St.Louis & NBER, john.barrios@wustl.edu. Giuliano: UC Santa Cruz, lgiulian@ucsc.edu. Leone: Northwestern, andrew.leone@kellogg.northwestern.edu.

1 Introduction

When hiring new workers, employers typically screen large numbers of written applications before selecting a subset for costly, in-person interviews. A broad literature suggests that information frictions lead to screening on imperfect quality signals such as educational pedigree and network-based referrals, and that such practices can perpetuate labor-market inequities.[1] In theory, a reduction in the cost of in-person interviews could boost their use as a screening tool, and thus lead to improvements in both efficiency and equity by reducing the use of blunt signals and stereotypes that favor certain groups. However, since in-person interactions may be influenced by speech, appearance, social skills, or cultural norms, elevating their role in the hiring process could introduce or magnify certain biases and put minority groups at a disadvantage.

This paper studies the introduction of a labor-market intermediary that dramatically reduced the cost of in-person screening in a high-skilled labor market: the market for PhD accountants. Before 2010, universities seeking to hire PhD accountants for research-oriented positions typically did not meet with applicants before selecting a short list of candidates to invite for campus interviews. However, in 2010, a two-day meeting known as the "Accounting Rookie Camp" (ARC) was created to facilitate in-person screening. Its stated purpose is to provide a "forum for faculty and recruiters to meet and network with PhD candidates, attend 15-minute research presentations by job-seeking candidates, and interview job-market candidates before making decisions about recruiting 'fly-outs.'"[2]

Our analysis leverages variation in the timing of ARC adoption by recruiters and degree-granting universities in order to estimate its moderating effect on several potential determinants of hiring outcomes. We start by asking whether recruiting through ARC caused employers to place less weight on a traditional quality signal and well-known correlate of job placements in academia: the research rank of one's PhD program. We also consider two proxies for referral networks: the coauthors of a candidate's PhD thesis adviser and the geographic proximity of degree schools to hiring schools. Next, we ask whether new information obtained through ARC helped employers select more productive applicants, with productivity measured by publication record. Finally, we examine ARC's impact on the role of individual characteristics that might be more salient when recruiters meet candidates

[1] The theoretical literature describes several ways in which information frictions lead to efficiency-equity tradeoffs in screening. Employers may statistically discriminate due to differences in group means (Phelps, 1972); beliefs about those means (Arrow, 1973; Coate and Loury, 1993); variances in productivity signals (Aigner and Cain, 1977); and "attention" discrimination (Bartos et al., 2016). Referral-based hiring can favor well-connected groups (Montgomery, 1991) and perpetuate inequalities (Calvo-Armengol and Jackson, 2004).

[2] http://aaahq.org/Meetings/2018/Accounting-Rookie-Camp.

in person. We use names and photographs to classify new PhDs by nationality, race, and gender, and to construct AI-based measures of physical beauty.

Consistent with prior research, we find a high degree of assortative matching in the labor market for new PhDs in accounting: degree-program reputation is a strong predictor of a candidate's initial success as measured by the rank of the hiring institution. However, our estimates also suggest this relationship is significantly weakened by participation in ARC. This finding is consistent across difference-in-difference models that use variation from both sides of the market to identify ARC's effect, and it does not appear to be driven either by pre-existing trends or by selection in the timing of ARC adoption. Instead, the results suggest that the historical return to program reputation was driven partly by its signaling value to employers, which is diminished by the new information channels created by the Rookie Camp.

We reach a similar conclusion when we investigate the role of network connections. Here, we find that even when controlling for assortative matching on institution rank, the presence of a coauthor connection between a candidate's adviser and a hiring school greatly increases the likelihood that a particular job match is realized. Again, however, the impact of adviser connections is significantly reduced after ARC adoption. Moreover, we find a similar reduction in the importance of geographic proximity in determining which job matches are realized.

If ARC reduced the need for traditional screening mechanisms and information channels by helping to improve the assessment of candidates, then its introduction should have led to more productive hires by the earlier adopters. Consistent with this prediction, we find positive overall effects of ARC adoption by recruiters on the average quantity and quality of publications authored by their new hires, with similar impacts on publications obtained prior to graduation and those within the first three years of employment. Among the highest-ranked recruiters, we find no significant impact on any measure of new hire productivity–despite the significant decline in the average rank of the degree schools from which they hire. In short, our results indicate that top-ranked recruiters placed no less weight on expected research productivity when screening through ARC. Instead, ARC seems to help them identify candidates from lower-ranked institutions who were highly qualified with respect to their research potential.

Having established that ARC helped level the playing field by reducing the role of educational pedigree and personal connections, we turn to the question of whether it also promoted equal opportunity for under-represented groups in the profession, including immigrants and women. The answer: it did not. Instead, we find that adoption of ARC led to lower-ranked jobs for Asian candidates, especially those with Chinese names, compared to their White and English-named counterparts with degrees from the same schools and similar publication records. We also find that among candidates with English names, ARC adoption exacerbated existing placement gaps by candidate gender.[3] What's more, when we estimate ARC's impact on the return to degree-school rank separately by demographic group, we find that the equalizing effect is driven entirely by male candidates, and is stronger for those with English-sounding names.

This set of results suggests that qualified immigrants and women were relatively unsuccessful at "selling" their qualifications in a setting where personal interactions played a salient role and the majority of recruiters and candidates were English-speaking men. For women, this conclusion is consistent with other recent studies of male-dominated workforces that find women and men are treated differently when presenting research seminars (Dupas et al., 2021) or pitching startups (Hu and Ma, 2021) and that gender gaps in promotion are mediated partly by differences in social interactions with male managers (Cullen and Perez-Truglia, 2019). For immigrants, additional analysis suggests that language and cultural barriers played a role.[4] Specifically, we find that the ARC-induced placement penalties for candidates with non-English names are greater among groups whose (predicted) native language is more distant from English, and among individuals who did not graduate from an English-speaking undergraduate institution. We also show that they are driven by employers in English-speaking countries, and that they do not reflect differences in the quality of candidates' *written* English as measured by textual analysis of PhD dissertations. Interestingly, the non-English penalty is especially large among candidates who have Chinese last names but adopted English first names, suggesting that the value of these names as signals of English fluency and/or assimilation with American culture was eroded once recruiters could easily meet candidates in person.[5]

Finally, we examine the role of beauty and find that ARC created a premium to physical attractiveness. This effect is not driven by any particular group, and it is independent of ARC's differential impact for women and immigrants.

A large body of research is devoted to understanding the practical importance and distributional consequences of various job screening mechanisms. In spirit, our study is similar

[3]We cannot rule out positive effects for Blacks and Hispanics; however, the estimated impacts for these groups are imprecise due to their small numbers in our sample.

[4]Lang (1986) formalizes the idea that employment discrimination can be rooted in communication difficulties across different groups. Bleakley and Chin (2004) show empirically that immigrants who are more fluent in the language of their host country have better labor market outcomes. Grogger (2019) shows that even among native English speakers, speech patterns play an important role in the sorting of workers into occupations that are intensive in interpersonal interactions.

[5]Roughly one quarter of the candidates in our sample have a Chinese last name, and one third of these have an English first name. The latter typically list their English first names in parenthesis after their Chinese first name.

to Autor and Scarborough (2008) who examine how the adoption of a new screening tool – computer-based job testing – affects both efficiency and equity in hiring. More recently, a related literature has focused on comparing machine algorithms to human judgement (e.g., Hoffman et al. 2018; Kleinberg et al. 2018; Cowgill 2020).

Much less studied, however, is the question of how human judgement in the hiring process may be affected by human interactions. The question is important for two reasons. First, concerns about machine-based hiring procedures are often based on the view that personal interaction is important for overcoming statistical discrimination. Such concerns have been amplified amid the Covid-19 pandemic and the increased use of machine-based hiring.[6] Second, while technology has produced substitutes for human judgment in the hiring process, many employers (especially in high-skilled settings) still employ in-person interviews. We advance the literature by providing a novel evaluation of the effects of in-person screening on human hiring decisions.

Because we can estimate ARC's impact on a wide range of quality signals and applicant characteristics, our findings contribute to several specific strands of the screening literature. First, we provide new evidence on the question of whether the returns to educational pedigree may reflect a signaling value (Spence, 1973) as well as a human capital effect (Becker, 1964) or pure sorting by ability. In the employer learning literature (Farber and Gibbons, 1996), a key insight is that if recruiters use education to statistically discriminate, then schooling should explain less of the variation in wages over time as employers learn about worker productivity (Altonji and Pierret, 2001). In this vein, Oyer and Schaefer (2019) interpret a large and stable return to attending an elite law school as more consistent with sorting and human capital effects than with a signaling explanation. But they also caution that the persistent elite-school premium could result from variation in initial job quality rather than the earnings potential of new graduates (Oyer, 2006; Oreopoulos et al., 2012; Kahn, 2010). Our approach circumvents this issue by showing that a change in the information available to recruiters affects the return to educational reputation in *initial* job placements. In this respect, our study is similar to MacLeod et al. (2017) who study the return to college reputation in the Colombian labor market and find that it is reduced by introduction of a

[6] For example, advocates for "ban-the-box" (or "fair chance") policies limiting the use of criminal records checks in employment screening have stressed the importance of personal contact for overcoming stereotypes (Anastasia and Natividad Rodriguez, 2016). Pager et al. (2009) provide some empirical support for the view that personal contact plays a role in mediating the effects of criminal stigma. More recently, the Covid-19 pandemic and the rise in prevalence of machine-based hiring algorithms have spotlighted concerns about the diminishing role for human interactions in hiring and led to regulation of automated employment decision tools (Weed, 2020). Ironically, a different product of the pandemic–the widespread adoption of video-conferencing technology–may ultimately lead to greater human interaction in the screening process, albeit not in person.

college exit exam giving employers new information about the skills of graduates.

Second, we add to the literature on hiring networks and employee referrals. While there is substantial credible evidence that networks and referrals influence labor market outcomes (Kramarz and Skans, 2014), less is known about the reasons why network-based hiring is common in so many settings. Support for information-based explanations has been found in the context of ethnicity-based networks in Germany (Dustmann et al., 2016), online labor market (Pallais and Sands, 2016), and in a selection of large U.S. firms (Burks et al., 2015; Brown et al., 2016). But other possible channels include reciprocity and favoritism (Beaman and Magruder, 2012). By examining the professional connections of PhD advisers to faculty at hiring institutions, we extend the evidence to the context of academic labor markets and conclude that adviser connections play a role in reducing information asymmetries.

Third, we contribute several insights to the literature on employment discrimination through our findings on ARC's differential impacts by candidate race, ethnicity, gender, and beauty. On race and ethnicity, some of the most compelling evidence of hiring discrimination has come from correspondence studies in which fictitious resumes receive fewer interview requests if they are randomly assigned black or ethnic names (e.g., Bertrand and Mullainathan 2004; Oreopoulos et al. 2012). Such findings suggest that minorities might receive more interviews if they adopted "majority" names, or if policies required employment applications to be anonymized. But the question remains whether more interviews would translate into more job offers.[7] Our findings that ARC led to worse job outcomes for immigrants – and especially for those who had adopted English-sounding names – suggests that any benefits from masking the identity of immigrants on written applications might be reversed at the interview stage.[8]

The literature has also produced clear evidence of systematic hiring discrimination by gender, most famously in the context of American symphony orchestras Goldin and Rouse (2000). Our findings contribute to a growing body of evidence of gender bias in other high-skilled, male-dominated professions including economics and surgery.[9] There is also well-

[7]Two interesting studies by Biavaschi et al. (2017); Arai and Skogman Thoursie (2009) find that immigrants who legally change their names have higher subsequent earnings compared to those who don't; however, these studies use non-experimental data and provide no direct evidence on mechanisms. On resume anonymization, a handful of field experiments in Europe produce mixed findings. Bertrand and Duflo (2016) survey this literature and note that generalizability is hard because firms generally self-select into these studies.

[8]A similar conclusion is reached by Aslund and Skans (2012) who study an experiment conducted by a local government in Sweden.

[9]Recent studies find that compared their male counterparts, female economists receive less credit for coauthored work (Sarsons et al., 2021), fewer citations for comparable work (Koffi, 2021), and differential treatment in the peer review process (Card et al. 2020; Hengel 2017). They are also subject to more "patronizing and hostile" questions in research seminars (Dupas et al., 2021) and to unprofessional language on online forums (Wu 2018, 2020). Sarsons (2019) finds that among surgeons, gender influences how information

established evidence of a beauty premium in earnings (Hamermesh and Biddle, 1994), and experimental evidence suggests this premium reflects both taste-based discrimination and a positive correlation between beauty and non-cognitive skills (Mobius and Rosenblat, 2006). Our finding that ARC adoption led to a beauty premium in job placements is consistent with this literature and suggests that differential success in job interviews is a likely mediator of the beauty premium in earnings.

Finally, our study complements prior work on academic labor markets in economics, accounting, and other related fields. Several studies find that degree-school reputation and adviser connections are both strong predictors of job market success in academia, but less is known about the underlying mechanisms.[10] Other research has highlighted the difficulty of assessing future research productivity when hiring new PhDs (Conley and Onder, 2014). We provide novel evidence that the returns to program reputation and adviser connections are driven partly by their signaling value to recruiters, and institutions like ARC can reduce their importance.

2 Setting

2.1 The academic labor market for PhD accountants

Like most other academic job markets, the academic market for PhDs in accounting is characterized by a decentralized search and matching process, with timing that is driven by the academic calendar. Most degrees are awarded in May or June, and academic jobs typically begin in July or August of the following year. The recruiting process for new PhDs typically starts in September of the year before the job begins.

Most job openings are listed online through the Accounting Research Network, the American Accounting Association (AAA), and HigherEd Jobs, although a small number of schools post solely on their own websites. Universities recruiting for research-oriented jobs require candidates to submit written applications that include a CV and a research sample or "job-market paper." They also require letters of reference, the most important of which is usually

about their ability is interpreted by others.

[10] Clauset et al. (2015) establish strong patterns of assortative matching between PhD program rank and placement rank across several fields in academia. The importance of attending an elite graduate or undergraduate institution has been documented in the markets for economists Athey et al. 2007; Smeets et al. 2006; financial economists (Chan et al., 2009); lawyers (Oyer and Schaefer, 2019); and accountants (Fogarty et al., 2011). Others have found that adviser connections play a significant role. For example, see Rose and Shekhar (2018) and Krueger and Wu (2000) on the market for economists; Hadlock and Pierce (2018) on financial economists and Baruffaldi et al. (2016) on scientists and engineers. School-based networks have also been studied in the markets for lawyers (Oyer and Schaefer 2016, 2019) and for PhDs in political science (Fowler et al., 2007).

written by the candidate's primary dissertation adviser. Because many skills and personal qualities that are highly valued are hard to convey credibly in writing, campus interviews or "flyouts" – in which candidates give a formal presentation of their research and have individual meetings with current faculty – are an essential part of the process. Fly-outs occur from late January to the end of March with the market clearing by April and with most candidates finding positions.[11]

Since fly-outs are typically offered only to a small fraction of applicants, the screening process used to narrow the applicant pool is critical. Prior to 2010, universities seeking to hire PhD accountants for research-oriented positions reviewed written application materials throughout the fall and began making fly-out invitations in January. Importantly, recruiters typically did not meet with candidates in person before selecting their fly-out list. The traditional recruiting approach in accounting differed in this regard from that of larger academic fields like economics that conduct formal job-market interviews at an annual academic meeting in late fall or early winter. The lack of an interview market in accounting stemmed partly from the mismatch in timing between the annual meeting of the AAA and the market for research-oriented positions.[12]

In 2010, a new professional meeting called the "Accounting PhD Rookie Recruiting and Research Camp" (henceforth ARC or "Rookie Camp") was organized by the University of Miami and held in early December. The ARC was expressly designed to facilitate in-person screening of new PhD job candidates and it did so in three ways. First, it created a venue where recruiters could set up short interviews with candidates. Second, it provided opportunities for informal meetings among all recruiters and candidates in attendance. Finally, and perhaps most importantly, it organized a series of formal, 15-minute research presentations which were given by participating job candidates and could be attended by any recruiter. The presentation schedule was determined randomly to minimize concerns about favoritism in the allocation of preferred time slots. Beginning in 2013, the Rookie Camp was organized by the AAA and by 2015 it had become an integral feature of the job market for new PhDs in accounting.

[11] Applicants who have not secured a position by late April may seek temporary academic employment as one-year visiting assistant professors or as lecturers; or they may postpone graduation and re-enter the job market the following year.

[12] As discussed in Bergner et al. (2016) and Hunt and Jones (2016), recruiters for teaching-oriented positions in accounting often set up short interviews at the AAA meeting, which is held in August of each year. However, the early timing of the AAA meeting makes it less useful to recruiters for research-oriented positions since the quality of a candidate's dissertation is easier to assess the closer it is to the final product.

2.2 Other features of the academic labor market

Beyond the introduction of the Accounting Rookie Camp, several features of the market for new PhDs in Accounting offer practical advantages for studying the screening process. First, academic research position are human-capital intensive and initial contracts for new PhDs are typically written for at least three years, making screening important. Second, while informational frictions are high for employers, they are relatively low (and arguably much less relevant) for applicants. In short, all jobs openings are posted in a few places at the same time of year, so the cost of learning about open positions is low; and the application procedure is fairly uniform–making the marginal cost of an application low. Third, there is a fixed supply of newly qualified candidates each year and a relatively small number of research-oriented positions for new PhDs. The combination of a small number of positions and a low marginal cost of an application makes it plausible to assume that candidates who desire a research job apply to most or all positions. Finally, while we lack data on earnings, it is also a reasonable assumption that candidates for research-oriented jobs share a primary objective of placing at the highest-ranked department, subject to some idiosyncratic preferences. Not only do higher-ranked department tend to pay higher salaries and benefits; strong initial placements also confer several advantages that impact future productivity and thus future job options (Oyer, 2006).

While our analysis does not require that these assumptions hold exactly, they provide a useful framework for interpreting our findings. For example, they suggest that the assortative matching of new PhDs to similarly ranked universities is unlikely to be driven by preference-based sorting of the job seekers. It is also unlikely that ARC had any significant effect on job matching by informing job-seekers about open positions, although it may have allowed applicants and recruiters to share information about idiosyncratic preferences or features of the job.

3 Data & Sample Characteristics

3.1 Data sources and sample construction

We construct a novel database with information on the supply, demand, and market outcomes for all individuals who received a PhD in accounting from a U.S. institution and entered the market between 2005 and 2015. To construct the database, we first identify individuals who meet the sample criteria by searching the Hasselback's Accounting Faculty Directories from 2002-03 through 2016-17. These directories contain information on all full-time accounting faculty who are employed each year at over 1,000 four-year institutions worldwide. The information includes each faculty member's name, rank, teaching/research interests and highest degree earned, as well as the degree-granting school and year of the degree.[13]

We use the directory information to link new PhDs to both the schools where they received their degrees and the institutions that hired them, for all new PhDs whose first job was as a full-time faculty member at a four-year post-secondary institution. To obtain the names, degree-schools and completion years of all other individuals who completed a PhD in accounting during our sample period (i.e., those whose first job was not at a four-year academic institution), we search ProQuest's dissertation database for all dissertations in accounting that were completed between 2005 and 2015. Our full sample from these two sources combined includes 2,270 graduates from one of 102 academic institutions in the U.S. with active doctoral programs in accounting.[14]

Because we have only limited information about individuals who are not found in the Hasselback directories, we restrict our primary analysis to doctoral graduates whose first job was at a four-year academic institution. In addition to excluding those with non-academic jobs, we also exclude graduates who obtain visiting or lecturer positions in their first year after graduating as well as a small number of graduates with start dates that far pre-date their graduation year or initial placements more than two years after earning the degree.[15] These restrictions leave us with a main analysis sample of 1,740 individuals. We use data on the full sample of 2,270 PhDs in our sample period to rule out market-driven changes in sample selection as a confounder in our analysis.

For the main sample, we combine the data on individuals' educational histories and their first jobs with information from several additional sources:

Rookie Camp Participation. We obtained information on ARC participants by compiling Rookie Camp programs for the years 2010-2014.[16] The programs provide the names of all

[13]The directories were downloaded from http://www.jrhasselback.com/FacDir.html. They are available every academic year through 2016-17, and were complied from information solicited from schools each spring, effective the following academic year.

[14]We restrict our focus to individuals whose PhDs were granted by a U.S. institution because nearly all of these schools are covered by ProQuest and because we are able to check this sample against counts of accounting PhDs conferred by year and U.S. institution that are included in the Hasselback directories. For graduates from MIT, which does not publish dissertations with ProQuest, we collected rosters from MITs internal dissertation database.

[15]We have information on non-academic placements only for a limited number of cases where we obtained a CV or LinkedIn profile. The Hasselback directores contain information on non-tenure track positions in academia; however, fewer than 1% of all candidates in our sample take such positions. Graduates with pre-PhD start dates typically held non-research positions before pursuing the doctorate and then returned to those institutions to resume employment. Those with late start dates may have delayed employment due to family or other reasons. Where possible, we validated all school and date information using online CVs, professional websites, and/or LinkedIn profiles, which we were able to collect for over 90% of our sample.

[16]These years correspond to the 2011-2015 graduation and hiring years in our sample; the first Rookie Camp was held in December 2010 for candidates and recruiters who planned to participate in the 2011 job

individuals participating as job candidates each year, their school affiliations and dissertation titles, and the date and time of each individual's 15-minute presentation. The programs also contain information on the recruiting schools that registered each year.

Research Productivity. We measure research productivity using information on individuals' publication histories. This information – including publication titles, coauthors, and dates – is collected from several sources: individuals' CVs, the Social Science Research Network (SSRN), Google scholar, and the Brigham Young University (BYU) Accounting research ranking site.[17] For new PhDs, we construct a dummy variable indicating whether the candidate had at least one publication in the year of graduation or earlier, and another for having a publication in a "top-tier" journal.[18] We measure post-hire productivity using the total number of publications and top-tier publications over the first three years of employment.

Institutional Rankings. We also use BYU institutional rankings to assign reputation measures to both a candidate's degree program and their hiring university. PhD program rankings are constructed using the total number of citations received by all individuals who graduated from a program in the previous six years; programs are rank-ordered based on this total citation count. University rankings are constructed similarly using the total citations for all faculty members employed at a university in a given year, excluding those listed as emeritus or retired. In both cases, to avoid concerns about endogenous changes in rankings over time, we measure rank at a fixed point in time, using the three years prior to the Rookie Camp.[19] We also convert both rank measures to percentile ranks to facilitate interpretation of the estimated coefficients.

Adviser Networks. To measure a candidate's network connections to potential employers, we collect information on their primary PhD adviser or dissertation chair. We obtain adviser names from a combination of CVs and dissertation title pages, and we construct a measure of connectedness using information on advisers' coauthors. Specifically, we define an adviser as being connected to an institution in year t if she has ever published a paper with a researcher who is employed at that institution in year t.[20]

market.

[17]https://www.byuaccounting.net/rankings/univrank/rankings.php

[18]We define "top-tier" publications to include the five top-ranked journals that publish economics-based accounting research (*The Accounting Review*, *Journal of Accounting Research*, *Journal of Accounting and Economics*, *Contemporary Accounting Research* and *Review of Accounting Studies*), plus two top-ranked finance journals that published research by accounting PhDs in our sample: *Journal of Finance* and *Journal of Financial Economics.*

[19]The BYU ranking data is not available prior to 2009

[20]Title pages were obtained from the candidate's dissertation, published online by ProQuest. We found adviser names for 86% of our sample. Coauthor information is obtained from individual's publications, which were gathered from the same sources used to obtain candidates' publications. We consider the adviser's

Candidate Personal Characteristics. We begin by use the candidates' names and a combination of classifying algorithms and APIs to predict their gender, race, ethnicity, and national origin. For gender, we use the NamSor Gender API, which infers gender from the combination of characters in individuals' first and last names. For cases where gender is identified with less than 99% certainty, we attempt to confirm or assign gender by searching for photographs from candidates' professional websites and by using assessments made by research assistants based on those photographs. For the small number of candidates for whom gender remains uncertain, we use the NamSor classification and we check for the sensitivity of results to the exclusion of these observations.

To classify full names into one of the six race/ethnicity categories used by the U.S. Census and one of 21 nationalities, we use both NamSor and "NamePrism", a non-commercial classification tool.[21] These tools assign each name a probability, ranging from 0% to 100%, of belonging to each of the race/ethnicity or nationality groups; we then assign candidates to the category with the highest probability.

Although most categories are predicted with high probabilities (in our sample, the median probability of belonging to one's assigned nationality group is over 99%) there are, nevertheless, a sizeable number of ambiguous cases. We do not attempt to use photographs to refine our race/nationality classifications as we do for gender.[22] However, when interpreting our results, we consider the possibility that our name-based predictions proxy for what employers infer from a candidate's name (e.g., about their fluency in English) – and that this information is updated when they meet the candidates in person. In particular, we perform analyses that classify candidates with Chinese last names (the largest "nationality" category aside from those with English names) into subgroups based on: (1) the language spoken in the country where they received their undergraduate degree (collected from candidates' CVs and LinkedIn profiles) and (2) whether they have an English-sounding first name or nickname.

---

coauthors on all publications between the year 1997 and year t_{+1} (to account for a lag between coauthorship and publication date).

[21] NamSor was developed by Elian Carsenat: http://www.namsor.com; it uses classifications algorithms trained on large labeled databases in the census. NamePrism employs a training data set of 57 million contact lists from a major Internet company along with US census data on the distribution of last names by race, and trains its algorithm using the homophily principle exhibited in communication as the basis for its ethnicity classifier.

[22] After hand-checking a number of race classifications, we found classification errors when using our classification tool. We are currently in the process of refining race classifications by manually reviewing photos.

3.2 Descriptive statistics

Figure 1 shows the numbers of participants in the market each year for new PhDs (panel a), PhD-granting schools (panel b), and schools that hired at least one new PhD (panel c). In addition, panel (a) shows the counts of new PhDs whose first job was at a four-year institution (i.e., the individuals in our analysis sample), and panel (b) shows degree schools that produced at least one such job candidate. Panel (c) also shows counts for a subset of "highly ranked" universities that had a research ranking above the sample median.[23] Finally, the light blue bar in each panel shows the numbers of Rookie Camp participants each year.

In the six years before the Rookie Camp, an average of 69 U.S. PhD programs sent PhDs to the market each year; an average of 140 new PhDs got jobs each year at an average of 111 hiring universities. The jobs numbers dip somewhat during the recession (2009-2010) and then rise to an average of 180 per year in the post-ARC years (2011-2015).

Among new PhDs who entered the market in 2011, 49 candidates (roughly 30%) had attended the inaugural Rookie Camp in the preceding December, representing 32 of the 70 PhD programs that sent candidates to the market. By 2015, participation in ARC had risen to 128 job-seekers representing nearly two-thirds of all PhDs and 70 of the 82 PhD programs with students on the market. On the demand side of the market, the number of hiring universities that participated in the ARC rose from 15 in 2011 to 61 in 2015. By 2015, 44% of hiring universities participated in ARC (panel c), and roughly half of new hires were made by a university that participated in ARC (not shown in the figure).

Because the Rookie Camp was intended primarily to facilitate the market for research-oriented jobs, participation rates – especially in the early years – skewed toward more highly-ranked schools. Trends in the composition of participants on both the supply and demand side of the market are shown in Figure 2. Starting with the supply side (panel a) we see that the average rank of degree schools participating in the first year of ARC was around the 65th percentile. However, this was followed by a steady downward trend, and by 2015, the average rank was close to the median (consistent with an 85% participation rate that year).

Turning to the demand side of the market, panel b of Figure 2 shows that the top-ranked recruiting schools were generally the first to attend ARC. Participants in the first three years were, on average, ranked at the 80th percentile of all hiring schools in the sample. Schools that began participating in 2014 and 2015 were somewhat lower-ranked, but even by the end of the sample period, the average rank of participating recruiting schools was around the 70th percentile, compared to a mean around the 25th percentile for non-ARC recruiters.

[23] As discussed below, these schools were much more likely to ever recruit through ARC, consistent with ARC's focus on the research market. When analyzing the effect of ARC participation by job candidates, we focus on the likelihood of obtaining a first job at a highly ranked school.

When analyzing the impact of ARC participation by recruiters, our preferred models will focus on the set of relatively high-ranked, research-oriented schools that had "adopted" ARC by the end of the sample period.

Descriptive statistics for our analysis samples are shown in Tables 1 and 2. Table 1 shows average characteristics of the individual job candidates and their degree programs, both for the sample overall, and separately for schools ranked above and below the median and in the years before and after the introduction of ARC. The overall sample is 42.2% female, 64.7% White, 28.8% Asian (or Pacific Islander) and 3.4% Black. The two largest groups by predicted national origin are English (44.1%) and Chinese (22.7%). Women, racial minorities, and Chinese students are all somewhat under-represented at high-ranked schools. Over time, there is a slight increase in the share of English candidates and a decline in the share of Asian candidates at both groups of schools.

With regard to research productivity, we find that 28.9% of candidates in our sample had at least one publication by the year they graduated, and 16.1% had at least on top-tier publication. Not surprisingly, these shares are higher in higher-ranked schools, although the grow more over time in lower-ranked schools. Candidates from higher-ranked schools also have more publications of all types in the first three years of employment, though again, there is some convergence over the sample period.

Turning to program characteristics, we see that the average cohort size in our sample is 3.2 students; higher-ranking programs are roughly 25% larger than lower-ranking ones, and both groups grow by around 15% over the sample period. Consistent with the research orientation of ARC, new PhDs from higher-ranking schools participate at 50% higher rate than those from lower ranking schools (59.6% vs. 39.4%). Unsurprisingly, they are also much more likely to be hired by "highly-ranked" universities over the sample period. However, while these graduates are more than twice as likely to obtain a highly-ranked job in 2005-2010, they are only 80% more likely to do so in the five years after ARC's introduction – suggesting a diminished role for program rank in placement success.

Table 2 provides descriptive statistics for four groups of recruiters in our sample: The "early adopters" who began participating between 2011 and 2013; the "late adopters" who began in 2014 or 2015; the "post-sample" adopters who attend for the first time in 2016 or 2017; and those who had never participated as of 2018. Consistent with Figure 2b, both the average rank and the fraction of schools ranked above the median is higher among the groups that participated earlier (rows 1 & 2). Row 3 shows the mean program rank of new PhDs hired by the schools in each groups. As expected, recruiters that participate in the ARC earlier, being more highly ranked themselves, tend to hire candidates from higher-ranked programs. Employers that participate earlier are also more likely to run one of the

PhD programs that produces the candidates in our sample. Interestingly, however, the "early adopter" group also has relatively large shares of non-U.S. schools (15%) and schools located in non-English-speaking countries (13%), suggesting that international schools interested in hiring PhDs from the U.S. viewed Rookie Camp participation as important.

In both tables 1 and 2, we see evidence of assortative matching on institutional reputation: consistent with prior literature on academic labor markets, there is a strong relationship between the rank of an individual's PhD program and the rank of the school where they were first hired. Appendix Figure A.1 illustrates this relationship with a binned scatter plot of average job rank against degree-school rank, along with the fitted line from a quadratic model. The relationship is close to linear with an R-squared of 0.22.

4 Empirical Strategy and Identifying Variation

The presence of some assortative matching on school rank is hardly surprising since PhDs from highly ranked programs tend to have higher research potential *on average*. But to the extent that there is unobserved individual variation within programs, the strength of the relationship might be driven partly by employers use of program rank as an imperfect proxy for applicant quality. In turn, if the Rookie Camp allowed recruiters to improve their assessments, we should expect program rank to become less predictive of job placements after recruiters and applicants begin participating in ARC. This same logic applies to the role of other imperfect quality signals such as network-based references.

To test these predictions, we use a difference-in-difference approach that exploits variation in the timing of ARC adoption by schools. We use three complementary but distinct approaches that differ in the source of variation in ARC participation (variation across recruiters vs. degree programs) and in the unit of analysis (candidate, employer, or match).

4.1 Models of hiring using recruiter variation

We begin by using variation among recruiters to estimate models in which the outcome describes the candidate that is hired. Specifically, for recruiter j in year t, we model the rank of the new hire's degree program ($\textit{Degree-School Rank}_{jt}$) as:

$$\begin{aligned} \textit{Degree-School Rank}_{jt} = &\beta_1 \cdot \textit{Recruiter Reputation}_j + \beta_2 \cdot \textit{Post ARC}_{jt} \\ &+ \beta_3 \cdot (\textit{Recruiter Reputation} \times \textit{Post ARC})_{jt} + \theta_t + \epsilon_{jt} \end{aligned} \tag{1}$$

Here, $\textit{Recruiter Reputation}_j$ is the *predicted value* of $\textit{Degree-School Rank}_{jt}$ from a model fitted to observed hires in 2005-2007, and based on the recruiter's rank and other fixed characteristics prior to 2011.[24] It can be thought of as measuring a recruiter's historical ability to attract candidates from highly-ranked programs due to its own research reputation; in the absence of any changes in reputations or hiring practices over time, we would expect $\beta_1 = 1$. $Post\ ARC_{jt}$ is a dummy variable that equals one if the hiring school sent a recruiter to the Rookie Camp in year *t or earlier*. We include year fixed effects (θ_t) to control for overall changes in market conditions, and we also estimate specifications that replace $Recruiter\ Reputation_j$ with recruiter fixed effects.

Our primary interest is in the coefficient on the interaction term, β_3, describes how participation in ARC as a recruiter moderates the relationship between the recruiter's reputation and the degree-school rank of it's new hires. In equation (1), β_3 is estimated using variation in ARC adoption among recruiters with the same predicted hiring patterns (based on years prior to the ARC's introduction). In this generalized difference-in-difference framework, a causal interpretation of β_3 requires that there is no selection on unobservables in the decision to begin recruiting at ARC. A second assumption underlying equation (1) is that there are no dynamic effects at the recruiter level; ARC participation changes recruiters' hiring in the year they participate, and not with a lag.

To assess the validity of our research design, we present a number of robustness tests. These include: (1) restricting the sample to employers that made at least one hire pre- and post ARC (to improve balance in the composition of hiring schools over time); (2) further restricting the sample to employers that eventually adopted ARC (to further improve comparability); (3) adding recruiter fixed effects to the model to control for all variation in hiring patterns due to fixed characteristics of employers; and (4) using measures of labor market tightness to control directly for potentially confounding variation due to differential effects of labor market conditions on different types of recruiters. We construct two measures of labor market tightness in each year: one based on the share of all new PhDs who obtain any job at a four-year academic institution, and one based on the share of all academic jobs that are at a "highly ranked" university. Appendix Figure A.2 illustrates the variation in these measures over time. Finally, we also present evidence from event-study models that support the identifying assumption of parallel trends in hiring patterns in the absence of ARC, and also confirm that the changes coincide with the timing of ARC adoption.

Figure 3 illustrates the research design and presents initial visual evidence in support of the key assumptions. The figure uses data on the subset of recruiters who eventually adopt the Rookie Camp during our sample period (henceforth "ever adopters"), and splits

[24] The model includes: the recruiting school's BYU ranking (see Section 3); an alternative rank measure published by UT Dallas; dummies for being in the U.S., for having a PhD program, and for being unranked (and thus assigned rank of zero) in either ranking measure; the average number of hires in 2005-2007; and the interactions of these variables with each other.

them into two groups based on the timing of adoption: the "early adopters" who begin participating between 2011 and 2013 (panel a) and the "late adopters" who begin in 2014 or 2015 (panel b). Within each of these categories, recruiters are further grouped into "higher-ranked" or "lower-ranked" based on their ranking relative to the median of the "ever-adopter" analysis sample.[25] The dashed lines plot the mean *predicted* value of $\textit{Degree-School Rank}_{jt}$ ($\textit{Recruiter Reputation}_j$ in equation 1) for each of these groups over time. Since this predicted value is based on hires made in 2005-2007 and is a fixed characteristic of each school, these group means will vary over time only if there are compositional changes in the sample that are correlated with hiring patterns. The flat trend in these lines is therefore reassuring.

The solid lines in Figure 3 plot the mean values of $\textit{Degree-School Rank}_{jt}$ for the *actual* hires made each year. In the years prior to ARC adoption (indicated by the vertical line in each panel), deviation of the solid lines from the predicted values reflect changes over time in unobserved determinants of hiring patterns. As expected, there is some fluctuation around the predicted means. Importantly, however, there is no evidence of differential trends in the years prior to ARC adoption, which is consistent with the assumption of no selection on time-varying unobservables. Further, both panels show a similar pattern following ARC adoption: among the higher-ranking employers, the average degree-school rank of new hires begins to fall relative to its predicted value. By the end of the sample period, the observed gap in average $\textit{Degree-School Rank}_{jt}$ between higher and lower-ranked employers is effectively eliminated. This indicates that there was a decline in assortative matching with timing that is strongly suggestive of a causal role for the Rookie Camp.[26]

4.2 Models of job placements using degree-school variation

Since the ability to screen through ARC requires participation by job candidates as well as by recruiters, we can also estimate ARC's impact on placements using variation in adoption by degree-granting schools. Here, we estimate linear probability models in which the outcome is defined as whether the first job obtained by candidate i with a PhD from degree school s in year t is at a highly ranked university (i.e., one that is ranked above the median of all

[25]Recall from Table 2 that the ever adopters are generally more highly ranked than non-adopters in our sample.

[26]Note that while the overall impact of ARC appears to grow over time following adoption, we cannot infer anything about the dynamic effects for individual employers from this figure; a dynamic effect for the market overall could be driven entirely by the increase in ARC participation over time.

employers in our sample). Our estimation equations take the following form:

$$\begin{aligned} \textit{High-Rank Job}_{ist} = \; & \beta_1 \cdot \textit{Degree-School Reputation}_s + \beta_2 \cdot \textit{Post ARC}_{st} \\ & + \beta_3 \cdot (\textit{Degree-School Reputation} \times \textit{Post ARC})_{st} \\ & + \beta_4 \cdot \textit{Year1}_{st} + \beta_5 \cdot (\textit{Degree-School Reputation} \times \textit{Year1})_{st} \\ & + \beta_6 \cdot X_{ist} + \beta_7 \cdot (X \times \textit{Post ARC})_{ist} + \beta_8 \cdot (X \times \textit{Year1})_{ist} \\ & + \theta_t + \epsilon_{ist} \end{aligned} \tag{2}$$

The variable *Degree-School Reputation*$_s$ is constructed using observed hires in 2005-2007 and the rank of a candidate's degree school prior to 2011 to predict the probability of *High-Rank Job*$_{ist}$.[27] Hence it measures the degree program's historical ability to place candidates at highly ranked universities due to the research reputation of its graduates. We also estimate variants of equation (2) that replace *Degree-School Reputation*$_s$ with degree- school fixed effects. In some specifications, we also include characteristics X_{ist} of the individual candidate. These include publication-based measures of productivity as well as candidate demographics.

While the unit of observation in (2) is a new PhD candidate, the treatment variable, *Post ARC*$_{st}$, is defined at the program level: a program is defined as having adopted ARC in year t if it sent at least one job candidate to the Rookie Camp in year t or earlier. Although ARC participation can vary at the individual level even within program cohorts, an individual's decision to participate is almost certainly endogenous. We therefore instrument the individual's participation decision with ARC adoption by the degree school. Further, to avoid concerns about selective timing of a degree school's adoption due to idiosyncratic variation in candidate quality between cohorts, we control for a dummy variable ($Year1_{st}$) for the school's first year of adoption.[28] Thus, the coefficients β_2 and β_3 can be interpreted as the reduced-form effect on candidate i's outcome of their school's decision, made in a prior year, to begin sending job candidates to the Rookie Camp. We are particularly interested in β_3, the coefficient on the interaction of degree-school reputation with the post-ARC dummy, as this will tell us whether program reputation carries less weight in determining a candidate's job market success after the school starts participating in ARC.

We also use models like (2) to explore whether the effect a degree school's participation in ARC differed by individual candidate characteristics X_{ist} conditional on any differential effects by the school's reputation. For example, by defining X_{ist} as an indicator for hav-

[27] Analogous to *Recruiter Reputation*$_j$ in equation (2), the prediction model for *Degree-School Reputation*$_s$ also includes other fixed school-level characteristics, such as average cohort size, that may affect a program's placement success.

[28] For example, a school's initial decision to participate could be motivated by an unusually promising candidate or by a last-ditch effort to help a weak candidate who has no other prospects by late November.

ing publications prior to the market, we use the coefficient β_7 (on the interaction of X_{ist} with $Post\ ARC_{st}$) to test whether this potentially observable signal of research productivity became more or less important in the context of the ARC.[29] Similarly, we can test for heterogeneous effects of ARC adoption by gender, race and nationality by defining X_{ist} as the relevant candidate characteristic.

4.3 Match-level models for the role of networks and proximity

Our third estimation approach uses variation from both sides of the market to study how adoption of ARC affected the characteristics of job matches. Different from our first two approaches, the unit of analysis is now a *potential match* between job candidate i and recruiter j, and we are interested in *match-specific* variation in the cost to recruiter j of learning about candidate i through channels other than the ARC. We focus on two proxies for information channels: adviser coauthor networks and geographic proximity.[30] We start by showing that both measures have independent effects on the likelihood that a particular match is realized, even when controlling for the effect of rank similarity between schools. We then ask whether the introduction of ARC reduced the importance of network connections and physical distance in determining the matches that are realized.

To proceed, we construct a match-level data set consisting of all possible recruiter-candidate pairs that can be formed from the set of all academic institutions that hired in a given year and all candidates that graduated and obtained a job in that year. We then estimate the following linear probability model:

$$\begin{aligned} Hire_{ijst} = {} & \beta_1 \cdot Similar\ Rank_{js} + \beta_2 \cdot Adviser\ Connection_{ij} + \beta_3 \cdot Log\ Distance_{js} \\ & + \beta_4 \cdot Post\ ARC_{j/st} + \beta_5 \cdot (Post\ ARC_{j/st} \times Similar\ Rank_{js}) \\ & + \beta_6 \cdot (Post\ ARC_{j/st} \times Adviser\ Connection_{ij}) \\ & + \beta_7 \cdot (Post\ ARC_{j/st} \times Log\ Distance_{js}) \\ & + \alpha_s + \gamma_j + \theta_t + \epsilon_{ijst} \end{aligned} \tag{3}$$

Here, the dependent variable equals one if candidate i from degree school s was hired by university j in year t and is zero otherwise. The variable $Similar\ Rank_{js}$, an indicator for whether the degree school and recruiting school are in the same rank decile, captures the effect of degree school rank on placements that is a primary focus of equations (1) and (2). The role of networks in the market is captured by $Adviser\ Connection_{ij}$, which takes on

[29] In these specifications, we also include the interaction of X_{ist} with $Year1_{ist}$.

[30] Catalini et al. (2020) show that geographic frictions play an important role in research collaborations among scientists.

a value of one if the candidate's dissertation adviser had a coauthor on the faculty of the recruiting school. And to assess the importance of geographic proximity, we include the log distance between the campuses of the degree school and recruiting school. As in equations (1) and (2), the treatment variable *Post ARC* indicates whether the recruiting school (j) or the degree school (s) had begun participating in ARC as of year t. We estimate the model using each source of variation, and in all specifications we include fixed effects for degree school (α_s), hiring school (γ_j), and year (θ_t). Our coefficients of interest are β_5, β_6 and β_7, which measure the effect of ARC participation on the importance of rank similarity, chair connections and geographic distance for determining job market matches.

5 ARC' Impact on Traditional Quality Signals

5.1 ARC's impact on the role of degree-school reputation

5.1.1 Evidence from variation in recruiter participation

Table 3 reports the estimates of equation (1), which examines how hiring outcomes are affected when recruiters participate in ARC. Here, the outcome is the percentile rank of the degree school where the hired candidate received their PhD. Column 1 shows the model estimates based on the full sample, while the specifications in columns 2-6 are restricted to recruiters that ever participated in the ARC. In columns 3-6, the sample is further restricted to recruiters that hired in both the pre and post-ARC periods.

To interpret the coefficients, recall that *Recruiter Reputation* is an index of recruiter rank variables in years prior to ARC and constructed from a model predicting the degree-school rank of new hires. Thus, in the absence of any changes in program ranks or hiring practices over time, we would expect $\beta_1 = 1$, and the coefficient of .95 in column 1 suggests that the relationship between recruiter rank and average degree-school rank of new hires was fairly stable prior to ARC adoption. In turn, the estimate of -0.40 for β_3 (the coefficient on the interaction of *Recruiter Reputation* with the *Post ARC* dummy) implies that the effect of a recruiter's "reputation" on the degree-school rank of their hires was reduced by roughly 40% when recruiters began participating in the Rookie Camp. The magnitude of this coefficient is reduced somewhat, but remains negative and significant, when estimated using only the "ever adopter" sample (column 2) and when excluding those that did not hire in both the pre- and post-ARC years (column 3). In our preferred specification that includes recruiter fixed effects (column 4), the estimate implies a 27% reduction in the effect of recruiter reputation on the average degree-school rank of new hires.

When interpreting the estimates of β_3, it is also helpful to compute the implied effect of ARC on hiring decisions for recruiters with different predicted hiring patterns. For ease of interpretation, the interaction term in Table 3 is computed using a demeaned version of *Recruiter Reputation* so that the coefficient on *Post ARC* represents the effect of ARC adoption for recruiters at the sample mean of roughly 55 (meaning on average, they are predicted to hire candidates from degree schools ranked at the 55^{th} percentile.) The small and statistically insignificant coefficient of 1.34 in column (4) thus implies that for the typical recruiter, recruiting through ARC did not significantly change the average rank of the degree schools they hired from. However, recruiters who historically hired PhDs from more highly ranked programs changed their decisions significantly. The interaction coefficient of -0.27 implies that for recruiters that historically hired from degree schools ranked at the 90^{th} percentile, the average degree-school rank of new hires fell by roughly 8 percentage points after ARC adoption.

The estimates in column (4) of Table 3 are identified off of within-recruiter changes in hiring patterns and differences in the timing of ARC adoption; they control for all fixed differences in hiring patterns across recruiters as well as annual changes in overall market conditions. While the inclusion of recruiter fixed effects reduces the precision of the estimates, it is reassuring that sign and magnitude of the estimates is fairly robust across model specifications, especially when estimated on the same sample (compare columns 3 and 4). Importantly, columns (5) and (6) show that the estimates are also highly robust to specifications that allow the effect of labor market tightness to vary with recruiter reputation.

Further support for a causal interpretation of the Table 3 results is provided in Figure 4. The four panels plot the coefficients and 95% confidence intervals from event study models that add four leads and two lags of the ARC adoption dummy and its interaction with *Recruiter Reputation* to equation (1).[31] All four specifications show a similar pattern. In the four years prior to adoption, the coefficients on the interaction terms are close to zero with no evidence of a trend; and they become negative and significant beginning in the year of ARC adoption.

Finally, in Appendix Figure A.3, we show a parallel set of event study models that use variation from a single treated cohort: recruiters that attended ARC in its inaugural year (2011). The sample is limited to the first and last adoption cohorts (recruiters who first adopted in 2015 serve as controls for those who adopted in 2011) and excludes the year 2015. Although the confidence intervals in these figures are roughly twice as large, the patterns broadly mirror those seen in Figure 4: there is no evidence of a pre-existing trend, and a clear change in hiring patterns in the first year of ARC. Since difference-in-difference designs with staggered adoption can lead to biased estimates when treatments are both dynamic and

[31] The four panels show specifications corresponding to columns 1-4 of Table 3.

heterogeneous across treatment cohorts (Goodman-Bacon 2018; Sun and Abraham 2020; Callaway and SantAnna 2020; Barrios 2021), the patterns in Appendix Figure A.3 are again reassuring.

5.1.2 Evidence from variation in degree-school participation

In Table 4, we examine the impact of ARC adoption on job market placements using variation from the supply side of the market. The table reports estimates of equation (2), in which the outcome is a dummy indicating whether a new PhD got a job at a highly ranked university. The treatment variable is an indicator for ARC adoption by the candidate's degree school, and this is interacted with the measure of degree-school reputation described in Section 3. Consistent with the results of the recruiter-based analysis, the estimates in Table 4 suggest that ARC adoption lead to a statistically significant, 21%-23% reduction in the effect of degree-school reputation on a candidate's probability of obtaining a high-ranked job. The coefficients are remarkably stable across specifications, including our preferred specification with degree-school fixed effects (column 4).[32]

In Appendix Table A.1, we demonstrate the robustness of the inferences from Table 4 to alternative measures of job market outcomes. Column 1 reproduces the estimates from column 4 of Table 4. In column 2, we replace the binary outcome with the percentile rank of the first job. And in column 3, we replace the rank measure with an indicator variable for whether the hiring school ever participated in ARC as a recruiter. Consistent with our previous results, column 2 shows that ARC adoption by degree programs led to a diminished association between the reputation of one's degree school and the rank of one's first employer. Additionally, the results in column 3 confirm that the effect of ARC participation on placements is indeed driven by changes in the likelihood of being hired by an employer that recruited through ARC.

5.2 ARC's impact on the role of connections and proximity

Table 5 reports estimates from the match-level models described by equation (3). First, column 1 confirms the importance of rank similarity for determining which matches are realized. Prior to the ARC, the probability that a candidate from degree school s is hired by school j is 0.61 percentage points higher (relative to a baseline probability of 0.79 for all

[32]All models include an indicator for the degree school's first year of adoption and its interaction with degree-school reputation. The coefficients on the interaction term (not shown) are all statistically insignificant but positive, suggesting somewhat smaller impacts in the first year. Since differential impacts in the first year may be driven by selection in the timing of ARC adoption, we focus on the estimates that control for this first-year effect, although its inclusion does not have a significant impact on our conclusions. Also for this reason, the degree-school based design does not lend itself well to an event-study approach.

other possible matches) if school s and school j are in the same rank decile. Column 2 adds our proxy for network connectedness: an indicator for whether the candidate's dissertation chair has a coauthor at school j. The presence of such a connection yields a 5.2 percentage point increase in the probability that a match is formed. Further, while the coefficient on rank similarity remains highly significant, its magnitude is reduced by about 12%, suggesting that the greater prevalence of adviser connections in similarly ranked schools explains some of the assortative matching by school rank.[33] We introduce geographic distance in column 3. Consistent with information frictions that increase with distance, the estimates suggest that a one percent increase in the physical distance between schools is associated with a reduction of 0.69 percentage points in the likelihood of a job market match.

The effect of ARC on these baseline relationships is examined in columns 4 and 5. In column 4, we use variation in the timing of degree-school participation, while column 5 uses variation in the timing of adoption by recruiters. The estimates from both models provide confirmatory evidence that ARC adoption significantly weakens the extent of assortative matching on institution rank. In column 4, we also see evidence that ARC reduces the importance of connections. The effect of an adviser connection on the probability of a match is reduced by more than 20 percent after ARC is adopted by the degree school. However, the estimates identified using recruiter variation (column 5) are much less precise and not statistically significant.

Finally, estimates from both models point to a 35-40 percent reduction in the effect of distance on match rates, and indicate that hires made after ARC adoption earned their degrees from schools that were further away, on average, from the hiring institution. These results suggest that distance-based information frictions became less critical after schools began participating in ARC.

5.3 ARC's impact on the productivity of new hires

Our results so far indicate that participating in ARC led recruiters to rely less on traditional quality signals and information channels when screening new PhD's for jobs–and that this, in turn, led to different hiring choices. We now ask whether these changes were driven by improvements in recruiters' ability to identify the most productive applicants. For example, the reduction in screening costs may have allowed recruiters to devote more attention to applicants from lower-ranked programs and those with fewer connections.[34] Personal inter-

[33]The conclusions are very similar if we use a continuous measure of rank distance rather than indicators for being in the same rank decile.

[34](Bartos et al., 2016) present a model and evidence of "attention discrimination" in which employers reduce the effort they allocate to inspecting resumes of applicants from negatively stereotyped groups.

actions with candidates at ARC may also have helped recruiters determine which candidates would be most productive at their institutions–i.e., the highest quality matches.

If ARC's impact on hiring patterns was indeed driven largely by employers ability to assess research potential, then it should have led to improvements in the productivity of hires made by ARC recruiters, relative to their competitors who had not yet adopted ARC. We test this prediction in Table 6. The table reports estimates from models like equation 1 except that the dependent variable is now a publication-based measure of individual productivity rather than the rank of a candidate's degree school. The first column reproduces the fixed-effects model for a new hire's degree-school rank (as in column 4 of Table 3); the remaining columns models for the hired candidate's productivity before they were hired (columns 2 and 3) or in the first three years of employment (columns 4 and 5). For ease of interpretation, the bottom panel of Table 6 reports the predicted effects of ARC adoption for recruiters at different rank percentiles.

For recruiters near the middle of the overall rank distribution, we find ARC adoption led to positive changes in measures of both pre-market and post-market publications–including a statistically significant, 11.5 percent increase in the number of "top-tier" publications three years after being hired. Notably, the average degree-school rank of hires made by this group of recruiters did not change significantly (column 1), which suggests that ARC allowed them to make more productive hires from among the same set of supplying PhD programs.

By contrast, recruiters ranked at the 90^{th} percentile hired candidates from significantly lower-ranked degree schools after adopting ARC. Yet ARC had no significant impact on the measured productivity of their new hires.[35] Thus, among top-ranked employers, the results suggest that ARC helped these recruiters find candidates from lower-ranked institutions who were similarly qualified with respect to their publication records (compared to hires made by these employers prior to ARC). Of course, these candidates may have been more desirable on some other dimension.

Appendix Table A.2 presents additional evidence on the role of candidate publications from models that use supply-side variation to estimate ARC's effects . Column 1 confirms that even among candidates from similarly ranked degree schools, pre-market publications are significant predictors of the likelihood of getting a high-ranked first job. Having at least one pre-market publication is associated with a 7-8 percentage point increase in this likelihood, while the return to having a top-tier publication is roughly three times as large. The return to publications falls slightly after ARC adoption (column 2), but the change

[35] The point estimates for these high-ranked recruiters indicate negative effects of ARC adoption on pre-market publications, but positive or zero effects on post-hire publications. This pattern is consistent with improvements in match quality; however, it pattern should be interpreted cautiously given that none of the estimates is statistically significant.

is not statistically significant. In column 3, the differential return to a top-tier publication becomes slightly positive once we control for ARCs effect on the return to degree-school rank (as in equation (2)), but again, it is small and not statistically significant.

In summary, despite the diminished weight placed of degree-school rank after ARC, there was no significant change in the importance of pre-market publications for a candidate's likelihood of being hired by a highly ranked university.

6 Heterogeneity by Ethnicity, Gender, and Beauty

The results presented thus far suggest that the ARC helped level the playing field by reducing the importance of attending a prestigious degree school, network effects (advisor connections), and by drawing attention to graduates of lower-ranked programs who have strong research potential (as measured by pre-market publications). A natural next question is whether ARC also led to greater equality of opportunity in other dimensions, or whether recruiters placed more weight on characteristics that benefited certain groups. In this section, we examine placement gaps by candidate gender, race, nationality (as predicted by candidates' names), and physical appearance, and we estimate the impact of ARC participation on these gaps.

For context, Figure A.4 illustrates the evolution of several group differences in job placements over the years in our sample period. Specifically, it plots coefficients from linear probability models for probability of obtaining a first job at a highly ranked university, controlling for the rank of the candidate's degree-granting school, a dummy for group membership, and the interactions of these variables with a linear time trend. Even among students who graduate from similarly ranked degree schools; there are significant gaps in placements by gender, race, and national origin. Panel (a) shows that placements of females have deteriorated over the sample period relative to males, with women being ten percentage points less likely to place at a highly ranked school by 2015. The next two panels (b-c) show gaps by predicted race and ethnicity for minorities (defined as Black race or Hispanic origin) and for Asians relative to non-Hispanic Whites. Panel (d) shows gaps by predicted national origin for the two largest subgroups in our sample: those with Chinese names and those with English names. In all cases, the figures show gaps in recent years, and with the exception of the the minority-white gap (which has been roughly stable at 7-9 percentage points over the period), the gaps appear to widen over time. For women, whose placements were similar to men at the start of the sample, the placement gap has widened to 12 percentage points. Asian and Chinese candidates start the sample with relatively high likelihoods of placing at a highly ranked school, but by the end of the period, these gaps are reversed.

7 Conclusion

We study the impact of the introduction of a labor-market intermediary, the Accounting Rookie Camp ("ARC"), which greatly facilitated in-person screening in the academic market for accounting PhD's. We find strong evidence that the increase in in-person screening led recruiters to place less weight on degree school reputation and referrals from connected advisors, both of which are imperfect quality signals, when making hiring decisions. Reputation of a candidate's degree school strongly predicts placement, but this effect is attenuated after the introduction of ARC. ARC appears to have "leveled the playing field" for candidates, as recruiters placed less weight on candidates' degree school and advisor connections. Our findings suggest that the signal value of program reputation and network connections were diminished by new information channels provided by ARC.

To assess whether ARC improved the assessment and matching of candidates, we test whether research productivity improved for early ARC adopters. We find that the average quantity and quality of publications of new hires is higher for ARC adopters. We do not find a significant change in new-hire productivity for ARC adopters among the highest-ranked recruiters-despite the significant decline in the average rank of the degree schools from which they hire.

We next test whether nationality, race, gender, and an AI-based measure of physical beauty are related to placement, and whether the relations changed after ARC adoption. We find that adoption of ARC led to lower-ranked jobs for Asian candidates relative to White candidates, particularly those with Chinese names, controlling for degree-school and publications. We also find that among candidates with English names, ARC adoption appears to have exacerbated existing placement gaps by gender. Finally, we find that ARC created a premium to physical attractiveness. This result is not driven by any particular group, and it is independent of ARC's differential impact for women and immigrants.

References

Aigner, D. J. and G. G. Cain (1977). Statistical theories of discrimination in labor markets. *Ilr Review 30*(2), 175–187.

Altonji, J. G. and C. R. Pierret (2001). Employer learning and statistical discrimination. *The Quarterly Journal of Economics 116*(1), 313–350.

Anastasia, C. and M. Natividad Rodriguez (2016). Nelp fact sheet: Research supports fair-chance policies. Technical report, National Employment Law Project.

Arai, M. and P. Skogman Thoursie (2009, Jan). Renouncing personal names: An empirical examination of surname change and earnings. *Journal of Labor Economics 27*(1), 127–147.

Arrow, K. J. (1973). The theory of discrimination. *Discrimination in labor markets 3*(10), 3–33.

Aslund, O. and O. N. Skans (2012). Do anonymous job application procedures level the playing field? *Industrial and Labor Relations Review 65*(1), 82–107.

Athey, S., L. F. Katz, A. B. Krueger, S. Levitt, and J. Poterba (2007). What does performance in graduate school predict? graduate economics education and student outcomes. *American Economic Review 97*(2), 512–520.

Autor, D. and D. Scarborough (2008). Does job testing harm minority workers? evidence from retail establishments. *Quarterly Journal of Economics 123*(1), 219–277.

Barrios, J. M. (2021). Staggeringly problematic: A primer on staggered did for accounting researchers. *Available at SSRN 3794859*.

Bartos, V., M. Bauer, J. Chytilova, and F. Matejka (2016). Attention discrimination: Theory and field experiments with monitoring information acquisition. *American Economic Review 106*(6), 1437–1475.

Baruffaldi, S., F. Visentin, and A. Conti (2016). The productivity of science & engineering phd students hired from supervisors' networks. *Research Policy 45*(4), 785–796.

Beaman, L. and J. Magruder (2012). Who gets the job referral? evidence from a social networks experiment. *American Economic Review 102*(7), 3574–93.

Becker, G. S. (1964). Human capital. *The University of Chicago Press*.

Bergner, J. M., J. J. Filzen, and J. A. Wong (2016). Navigating the accounting academic job market and related advice. In *Advances in Accounting Education: Teaching and Curriculum Innovations*, pp. 147–176.

Bertrand, M. and E. Duflo (2016, February). Field experiments on discrimination. *National Bureau of Economics Working Paper* (22014).

Bertrand, M. and S. Mullainathan (2004). Are emily and greg more employable than lakisha and jamal? a field experiment on labor market discrimination. *American Economic Review 94*(4), 991–1013.

Biavaschi, C., C. Giulietti, and Z. Siddique (2017). The economic payoff of name americanization. *Journal of Labor Economics 35*(4), 1089–1116.

Bleakley, H. and A. Chin (2004). Language skills and earnings: Evidence from childhood immigrants. *Review of Economics and statistics 86*(2), 481–496.

Brown, M., E. Setren, and G. Topa (2016). Do informal referrals lead to better matches? evidence from a firm's employee referral system. *Journal of Labor Economics 34*(1), 161–209.

Burks, S. V., B. Cowgill, M. Hoffman, and M. Housman (2015). The value of hiring through employee referrals. *The Quarterly Journal of Economics 130*(2), 805–839.

Callaway, B. and P. H. SantAnna (2020). Difference-in-differences with multiple time periods. *Journal of Econometrics*.

Calvo-Armengol, A. and M. O. Jackson (2004). The effects of social networks on employment and inequality. *American Economic Review 94*(3), 426–454.

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

Catalini, C., C. Fons-Rosen, and P. Gaule (2020). How do travel costs shape collaboration? *Management Science forthcoming*.

Chan, K. C., C. R. Chen, and H.-G. Fung (2009). Pedigree or placement? an analysis of research productivity in finance. *Financial Review 44*(1), 87–111.

Clauset, A., S. Arbesman, and D. B. Larremore (2015, February). Systematic inequality and hierarchy in faculty hiring networks. *Science Advances 1*(1).

Coate, S. and G. C. Loury (1993). Will affirmative-action policies eliminate negative stereotypes? *The American Economic Review*, 1220–1240.

Conley, J. P. and A. S. Onder (2014). The research productivity of new phds in economics: the surprisingly high non-success of the successful. *Journal of Economic Perspectives 28*(3), 205–16.

Cowgill, B. (2020, March). Bias and productivity in humans and algorithms: Theory and evidence from resume screening.

Cullen, Z. B. and R. Perez-Truglia (2019, December). The old boys' club: Schmoozing and the gender gap. *National Bureau of Economics Working Paper* (26530).

Dupas, P., A. S. Modestino, M. Niederle, J. Wolfers, et al. (2021). Gender and the dynamics of economics seminars. Technical report, National Bureau of Economic Research.

Dustmann, C., A. Glitz, U. Schönberg, and H. Brücker (2016). Referral-based job search networks. *The Review of Economic Studies 83*(2), 514–546.

Farber, H. S. and R. Gibbons (1996). Learning and wage dynamics. *The Quarterly Journal of Economics 111*(4), 1007–1047.

Fogarty, T. J., D. V. Saftner, and J. R. Hasselback (2011). Knowing one's place: The distribution of new accounting academics into a segmented labor market. *Journal of Accounting Education 29*(2-3), 89–99.

Fowler, J. H., B. Grofman, and N. Masuoka (2007). Social networks in political science: Hiring and placement of ph.ds, 1960–2002. *PS: Political Science & Politics 40*(4), 729–739.

Goldin, C. and C. Rouse (2000). Orchestrating impartiality: The impact of "blind" auditions on female musicians. *American Economic Review 90*(4), 715–741.

Goodman-Bacon, A. (2018). Gender and the dynamics of economics seminars. Technical report, National Bureau of Economic Research.

Grogger, J. (2019). Speech and wages. *Journal of Human Resources 54*(4).

Hadlock, C. J. and J. R. Pierce (2018). Hiring your friends: Evidence from the market for financial economists. *Industrial and Labor Relations Review forthcoming.*

Hamermesh, D. S. and J. E. Biddle (1994). Beauty and the labor market. *The American Economic Review 84*(5), 1174–1194.

Hengel, E. (2017). Publishing while female. are women held to higher standards? evidence from peer review.

Hoffman, M., L. B. Kahn, and D. Li (2018). Discretion in hiring. *The Quarterly Journal of Economics 133*(2), 765–800.

Hu, A. and S. Ma (2021, January). Persuading investors: A video-based study.

Hunt, S. and K. T. Jones (2016). Interviewing potential accounting faculty at the aaa national meeting and rookie camp: Surveys of schools'experiences and impressions. *Academy of Educational Leadership Journal 20*(1), 65–84.

Kahn, L. B. (2010). The long-term labor market consequences of graduating from college in a bad economy. *Labour Economics 17*(2), 303–316.

Kleinberg, J., H. Lakkaraju, J. Leskovec, and S. Ludwig, Jens anf Mullainathan (2018). Does job testing harm minority workers? evidence from retail establishments. *Quarterly Journal of Economics 133*(1), 237–293.

Koffi, M. (2021). Innovative ideas and gender inequality.

Kramarz, F. and O. N. Skans (2014). When strong ties are strong: Networks and youth labour market entry. *Review of Economic Studies 81*(3), 1164–1200.

Krueger, A. B. and S. Wu (2000). Forecasting job placements of economics graduate students. *The Journal of Economic Education 31*(1), 81–94.

Lang, K. (1986). A language theory of discrimination. *The quarterly journal of economics 101*(2), 363–382.

MacLeod, W. B., E. Riehl, J. E. Saavedra, and M. Urquiola (2017). The big sort: College reputation and labor market outcomes. *American Economic Journal: Applied Economics 9*(3), 223–61.

Mobius, M. M. and T. S. Rosenblat (2006). Why beauty matters. *American Economic Review 96*(1), 222–235.

Montgomery, J. D. (1991). Social networks and labor-market outcomes: Toward an economic analysis. *The American Economic Review 81*(5), 1408–1418.

Oreopoulos, P., T. Von Wachter, and A. Heisz (2012). The short-and long-term career effects of graduating in a recession. *American Economic Journal: Applied Economics 4*(1), 1–29.

Oyer, P. (2006). Initial labor market conditions and long-term outcomes for economists. *Journal of Economic Perspectives 20*(3), 143–160.

Oyer, P. and S. Schaefer (2016). Firm/employee matching: An industry study of us lawyers. *ILR Review 69*(2), 378–404.

Oyer, P. and S. Schaefer (2019). The returns to elite degrees: The case of american lawyers. *ILR Review 72*(2), 446–479.

Pager, D., B. Western, and N. Sugie (2009). Sequencing disadvantage: Barriers to employment facing young black and white men with criminal records. *The Annals of the American Academy of Political and Social Science 623*(1), 195–213.

Pallais, A. and E. G. Sands (2016). Why the referential treatment? evidence from field experiments on referrals. *Journal of Political Economy 124*(6), 1793–1828.

Phelps, E. S. (1972). The statistical theory of racism and sexism. *The american economic review 62*(4), 659–661.

Rose, M. E. and S. Shekhar (2018). Informal contacts in hiring: The economics job market. *Max Planck Institute for Innovation & Competition Research Paper No. 18-12*.

Sarsons, H. (2019). Interpreting signals in the labor market: evidence from medical referrals.

Sarsons, H., K. Gërxhani, E. Reuben, and A. Schram (2021). Gender differences in recognition for group work. *Journal of Political Economy 129*(1), 000–000.

Smeets, V., F. Warzynski, and T. Coupé (2006). Does the academic labor market initially allocate new graduates efficiently? *Journal of Economic Perspectives 20*(3), 161–172.

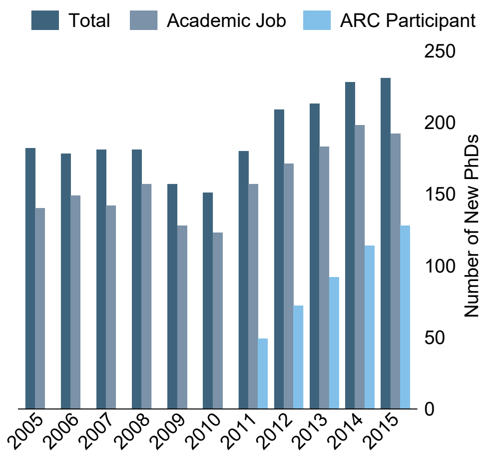
Spence, M. (1973). Job market signaling. *The Quarterly Journal of Economics 87*(3), 355–374.

Sun, L. and S. Abraham (2020). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics x*(x), x.

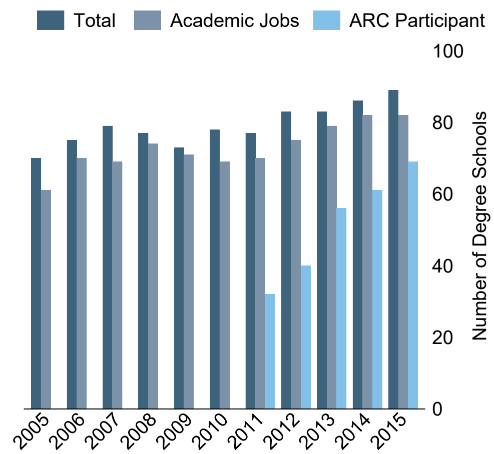
Weed, J. (2020, Nov). Job interviews without interviewers, products of the pandemic.

Wu, A. H. (2018). Gendered language on the economics job market rumors forum. In *AEA Papers and Proceedings*, Volume 108, pp. 175–79.

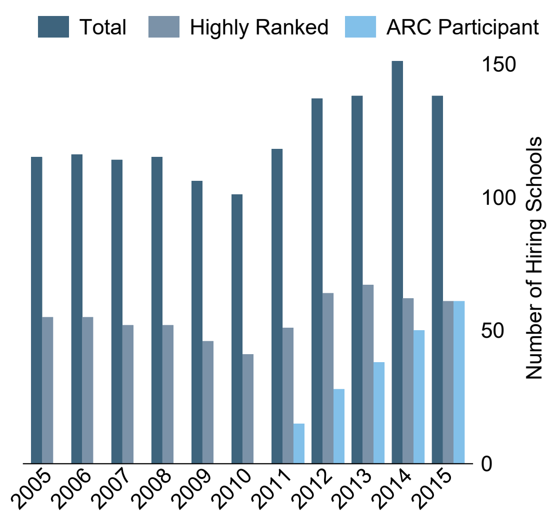
Wu, A. H. (2020). Gender bias among professionals: An identity-based interpretation. *Review of Economics and Statistics 102*(5), 867–880.



(a) Supply: New PhDs

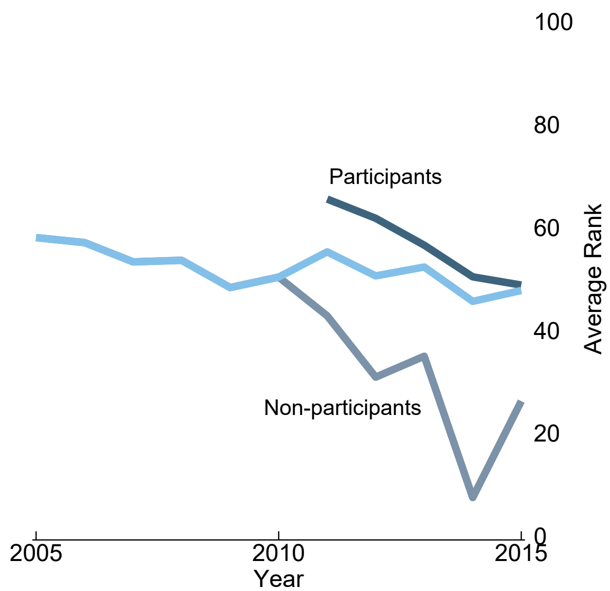


(b) Supply: PhD-Granting Schools

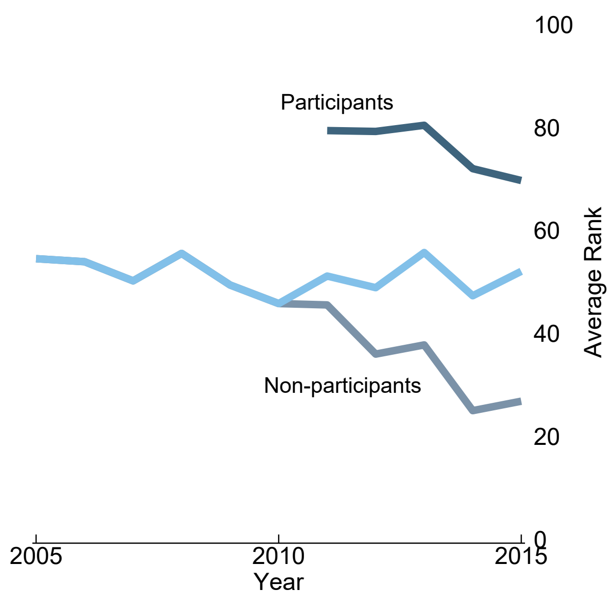


(c) Demand: Hiring Schools

Figure 1: This figure plots annual numbers of individual and institutional participants in the labor market for PhD accountants. Panel (a) shows the total number of new PhDs who graduated from one of 102 accredited U.S. PhD programs and the number hired by an academic institution. Panel (b) shows the number of U.S. universities that produced at least one new PhD each year and the number that placed at least one new PhD in an academic job. Panel (c) plots the number of academic institutions (worldwide) that hired at least one of the new PhDs in the sample. In panel(c), highly ranked schools are those ranked above the sample median prior to ARC (see text for details). Each panel also shows the number of Rookie Camp (ARC) participants each year.

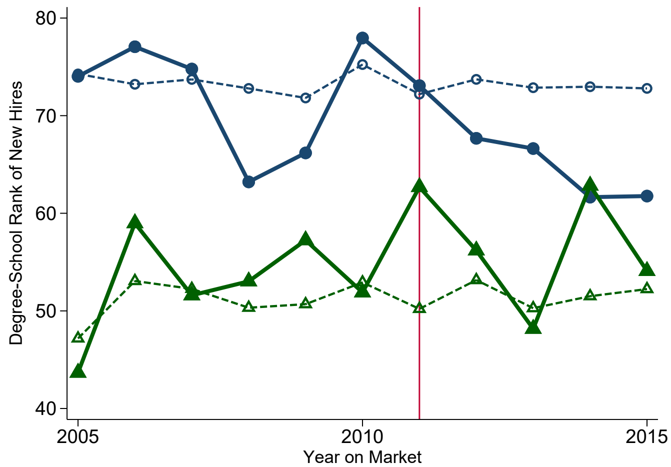


(a) Degree Schools

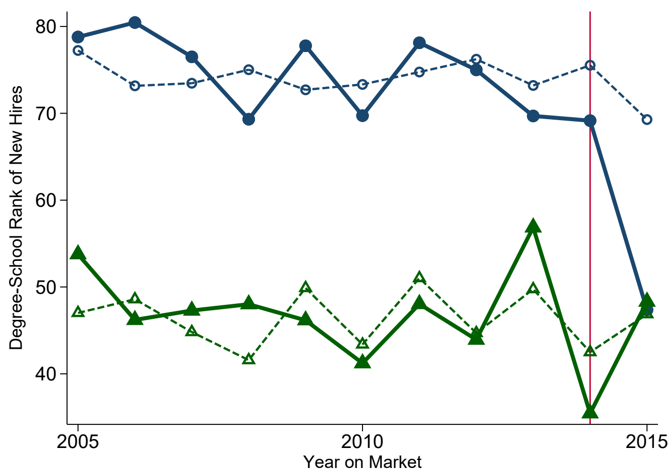


(b) Hiring Schools

Figure 2: This figure plots the average percentile rank of degree-granting schools (a) and hiring schools (b) by year and by participation in the Rookie Camp (ARC).



(a) Early adopters (2011-2013)



(b) Late adopters (2014-2015)

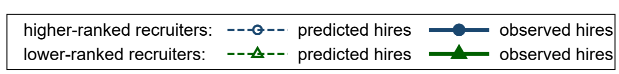
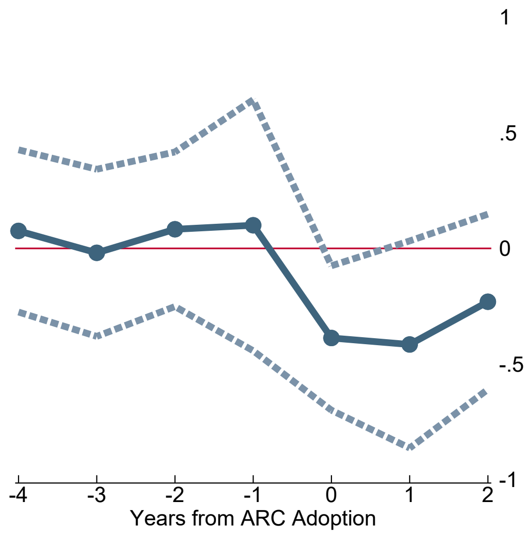
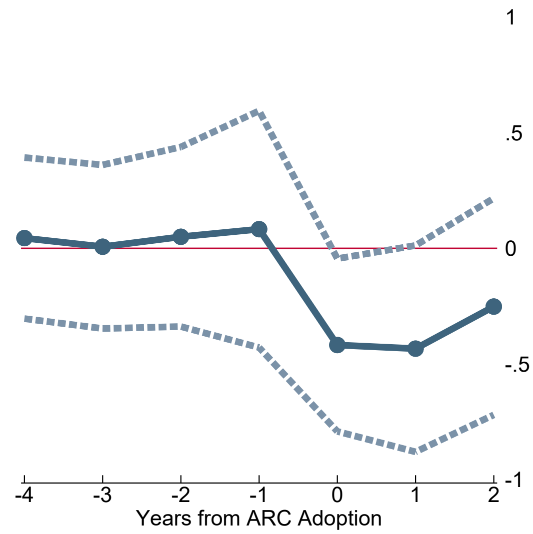


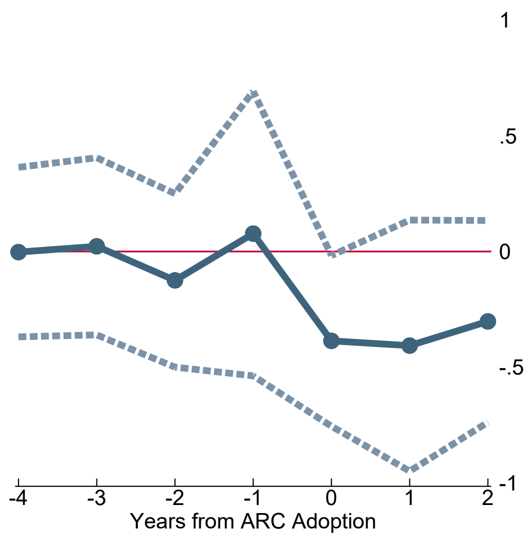
Figure 3: This figure plots the mean rank of the degree schools from which recruiters hired each year, along with the mean predicted rank ("recruiter reputation"), for all hiring schools that eventually adopted the Rookie Camp ("ever adopters"). The sample is split into two groups based on the first year the hiring school participated in ARC as a recruiter: "early adopters" (panel a) began participating between 2011 and 2013 and "late adopters" (panel b) began in 2014 or 2015 . Within each of these categories, recruiters are further grouped into "higher-ranked" or "lower-ranked" based on their reputation (i.e., their predicted degree-school rank of new hires) relative to the median of the "ever adopter" sample.



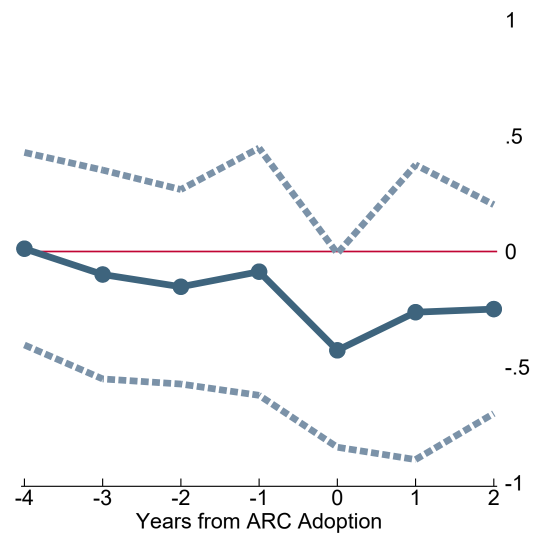
(a) Model 1



(b) Model 2



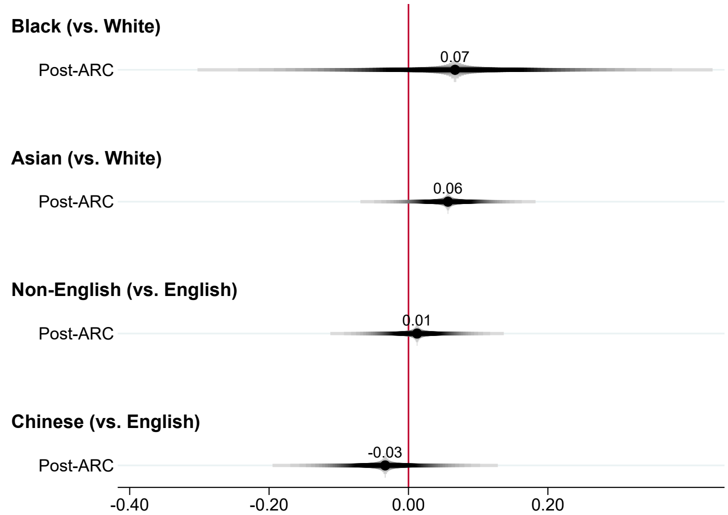
(c) Model 3



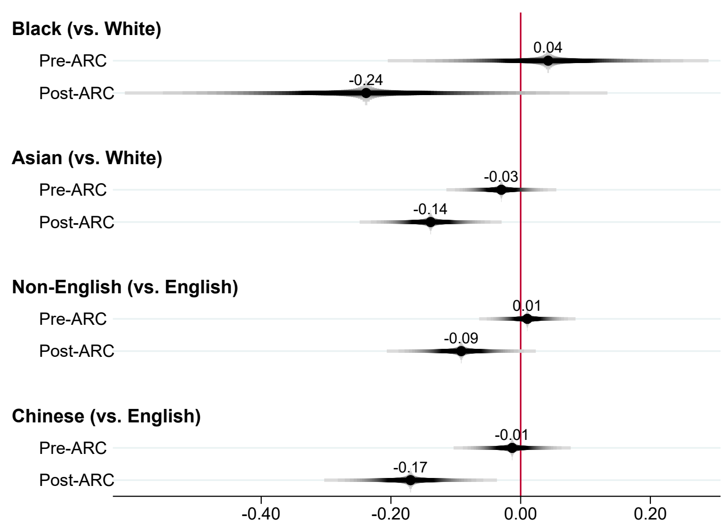
(d) Model 4

Figure 4: This figure illustrates the evolution over time, before and after recruiters adopted ARC, of the relationship between *Recruiter Reputation* and the degree-school rank of new hires. Each panel plots coefficient estimates and 95% confidence intervals for leads and lags of the ARC adoption dummy and its interaction with *Recruiter Reputation* from models that add four leads and two lags to equation (1). Panels (a)-(d) correspond to the specifications reported in columns (1)-(4) of Table 3.

.

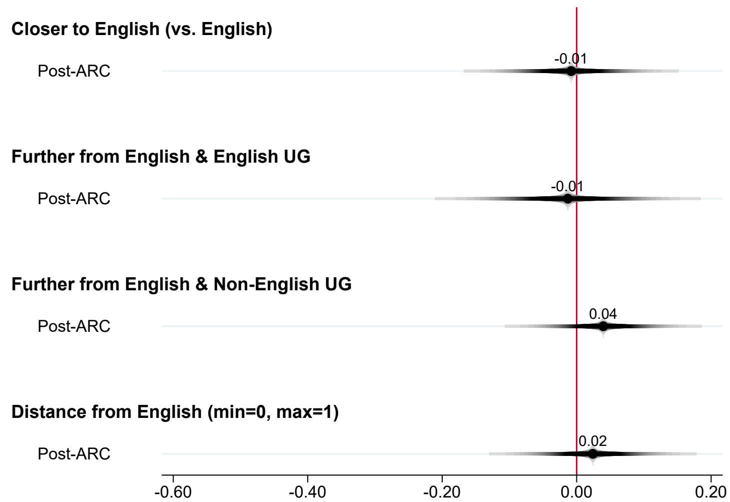


(a) Participated in ARC

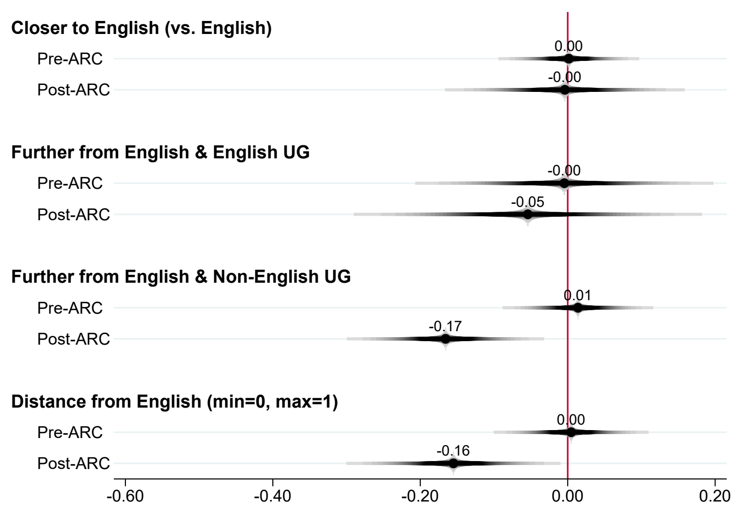


(b) First Job at Highly Ranked School

Figure 5: This figure shows estimated gaps in the likelihood of participating in ARC (Panel a) and the likelihood of obtaining a first job at a highly ranked university before and after a candidate's degree school adopts ARC (Panel b). The bands show graduated confidence intervals for up to 99% confidence (the lightest shaded segments). Estimates are constructed from regression models like equation (2) in which X_{ist} is an indicator for the specified candidate characteristic. The corresponding coefficients and standard errors are reported in columns 1, 4, 7, and 10 of Appendix Table A.3.

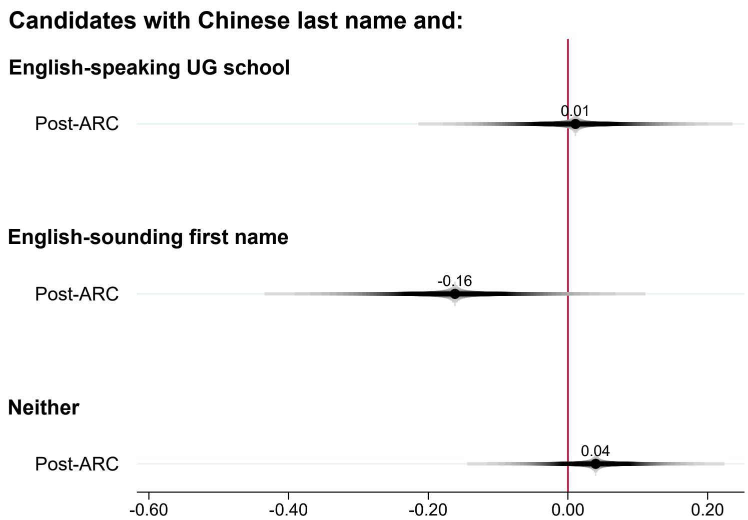


(a) Participated in ARC

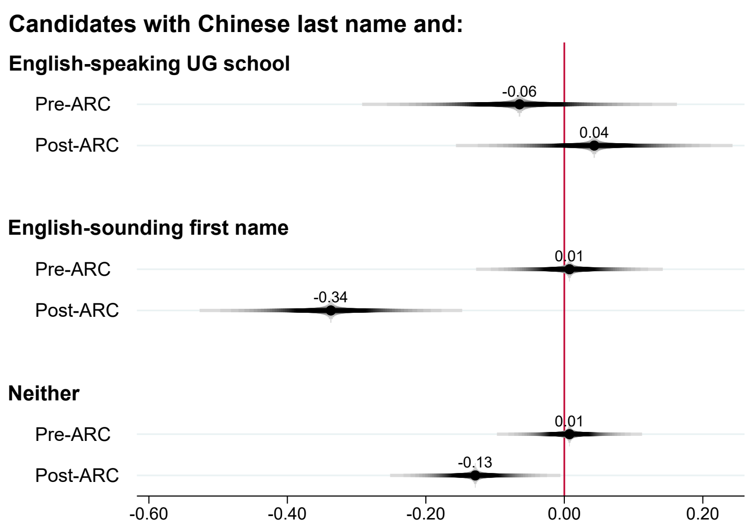


(b) First Job at Highly Ranked School

Figure 6: This figure shows estimated gaps, relative to candidates with English names, in the likelihood of participating in ARC (Panel a) and the likelihood of obtaining a first job at a highly ranked university before and after a candidate's degree school adopts ARC (Panel b). The bands show graduated confidence intervals for up to 99% confidence (the lightest shaded segments). Estimates are constructed using coefficients from the regression models shown in columns 1, 4, 7, and 10 of Appendix Table A.5. *Distance from English* is a continuous variable, ranging from zero to one; it is constructed using the candidate's predicted native language (see text for details and Appendix Table A.4 for the values corresponding to each nationality). The category *Closer to English* indicates .80 $> DistancefromEnglish >$ 0, while *Further from English* indicates *Distance from English*$\geq$ .80. *English UG* and *Non-English UG* indicate whether the candidate graduated from an undergraduate institution where the primary language of instruction is English.

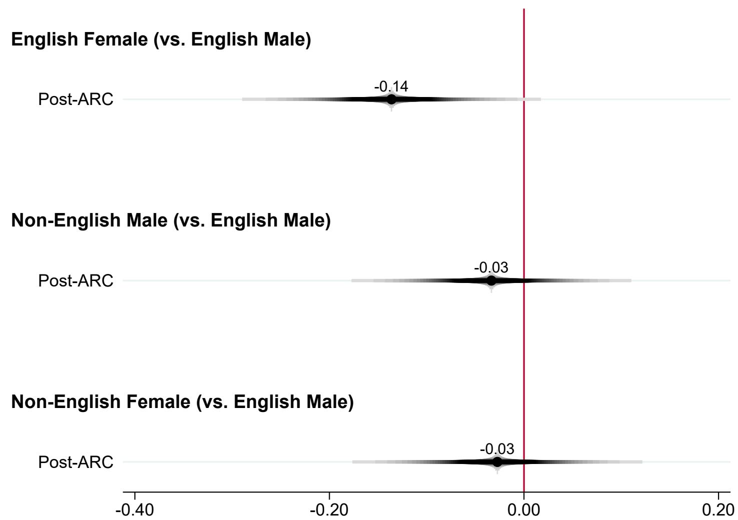


(a) Participated in ARC

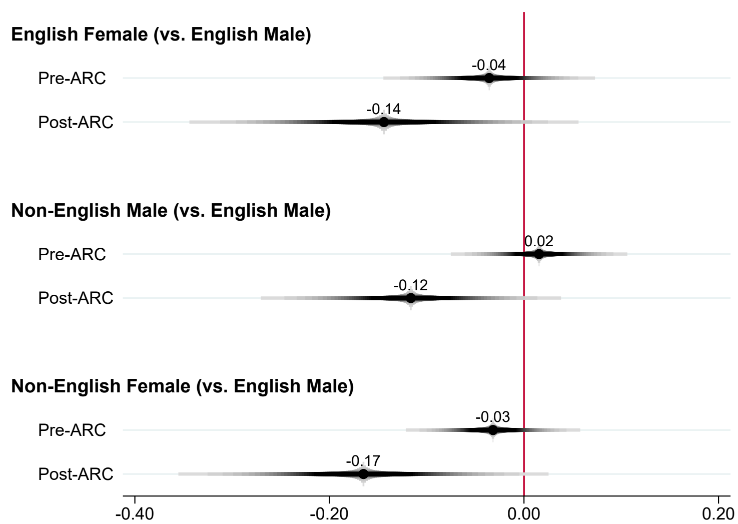


(b) First Job at Highly Ranked School

Figure 7: This figure shows estimated gaps, relative to candidates with English names, in the likelihood of participating in ARC (Panel a) and the likelihood of obtaining a first job at a highly ranked university before and after a candidate's degree school adopts ARC (Panel b). The bands show graduated confidence intervals for up to 99% confidence (the lightest shaded segments). Estimates are constructed from regression models like equation (2) in which X_{ist} is an indicator for the specified candidate characteristic. The corresponding regression coefficients and standard errors are reported in columns 1, 4, 7, and 10 of Appendix Table A.8.



(a) Participated in ARC



(b) First Job at Highly Ranked School

Figure 8: This figure shows estimated gaps, relative to male candidates with English names, in the likelihood of participating in ARC (Panel a) and the likelihood of obtaining a first job at a highly ranked university before and after a candidate's degree school adopts ARC (Panel b). The bands show graduated confidence intervals for up to 99% confidence (the lightest shaded segments). Estimates are constructed from regression models like equation (2) in which X_{ist} is a dummy variable for the specified demographic group. The corresponding regression coefficients and standard errors are reported in columns 1, 4, 7, and 10 of Appendix Table A.10.

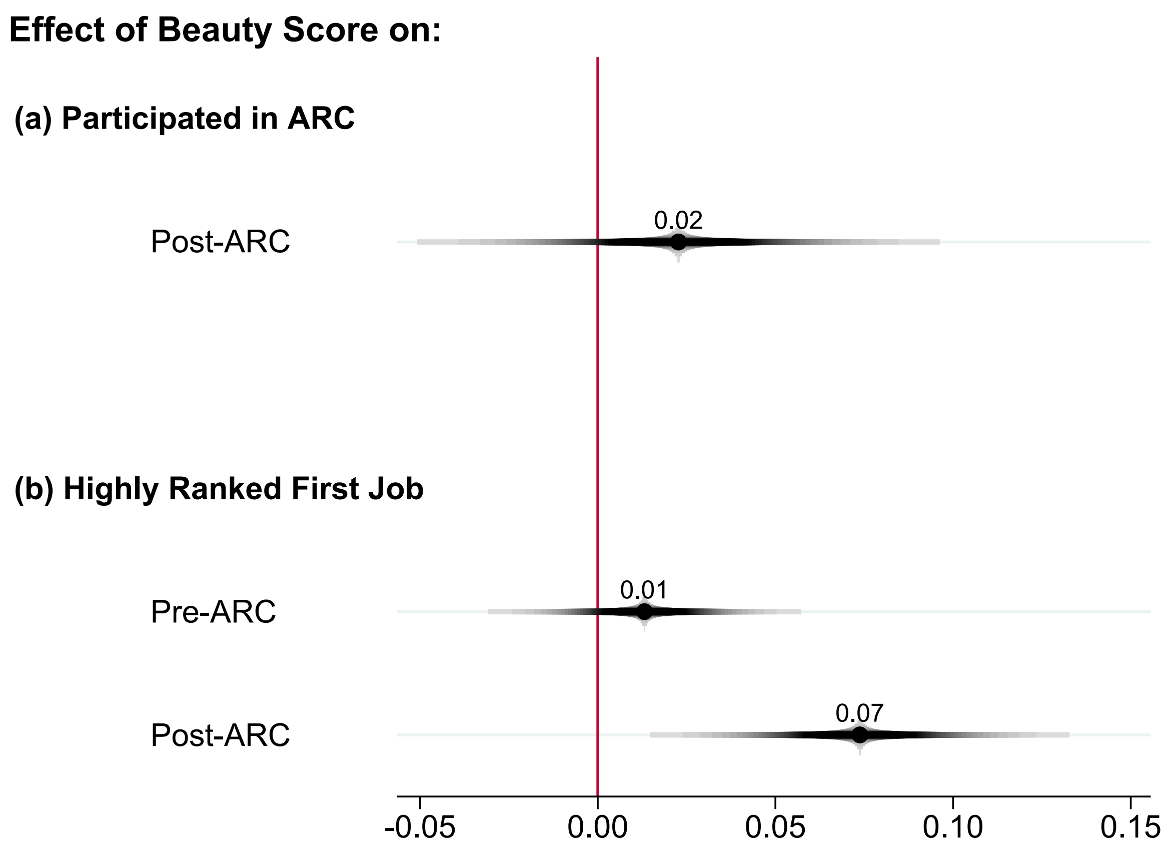


Figure 9: This figure shows the estimated effect of a one standard deviation increase in a candidate's beauty score on the likelihood of participating in ARC and the likelihood of obtaining a first job at a highly ranked university before and after a candidate's degree school adopts ARC. The bands show graduated confidence intervals for up to 99% confidence (the lightest shaded segments).

Table 1: Characteristics of New PhDs in Sample (2005-2015)

	Lower-Ranked Degree Schools		Higher-Ranked Degree Schools		
Candidate Characteristics	2005-2010	2011-2015	2005-2010	2011-2015	Total
Female (%)	43.0	45.1	40.0	41.3	42.2
Perceived Race:					
White (%)	62.1	62.6	66.0	67.0	64.7
Black (%)	5.4	5.1	2.1	2.0	3.4
Asian (%)	29.3	28.8	29.5	27.7	28.8
Predicted Nationality:					
English (%)	41.9	46.3	41.2	46.7	44.1
Chinese (%)	25.4	24.6	22.1	19.8	22.7
Other Asian (%)	8.0	7.9	8.6	9.5	8.6
Other European (%)	17.9	16.7	22.5	16.0	18.4
Other (e.g. Middle Eastern) (%)	6.0	3.2	5.1	6.9	5.3
Undergraduate Institution:					
Non-missing (%)	77.8	86.0	86.5	90.5	85.7
In U.S. (%)	60.4	68.5	54.5	67.4	62.7
In English-speaking country (%)	65.9	70.8	60.9	70.5	67.0
Publications:					
Has 1+ pre-market publication (%)	19.7	23.9	33.4	34.7	28.8
Has 1+ pre-market top-tier pub. (%)	6.3	10.6	21.1	22.8	16.1
Has 1+ post-hire publication (%)	29.1	26.4	47.3	43.4	37.6
Has 1+ post-hire top-tier pubs (%)	10.5	13.3	34.0	29.1	23.0
Hiring School (First job) Quality:					
Percentile Rank	36.9	37.7	62.6	61.9	51.4
Highly Ranked Job (%)	29	35	64	63	50
Participated in Rookie Camp (%)	–	39.4	–	59.6	–
Cohort size	2.6	2.9	3.2	3.7	3.2
Number of PhDs	351	406	488	495	1740
Number of degree schools	54		45		99

Notes: This table reports average characteristics for the new PhDs in the analysis sample. The first four columns show means for sub-groups based on whether the percentile rank of one's degree school is below the sample median (Lower-Ranked) or above it (Higher-Ranked), and on whether the candidate entered the job market in the years before (2005-2010) or after (2011-2015) the Rookie Camp was introduced. The final column shows overall sample means.

Table 2: Characteristics of Hiring Schools by Participation in ARC as a Recruiter

	First Year Participated in ARC as Recruiter				
Recruiter Characteristics	Early adopters: 2011-2013	Late adopters: 2014-2015	Post-sample: 2016-2017	Never as of 2018	Total
BYU Research Ranking (percentile)	79.2	61.8	46.5	23.4	51.4
Highly Ranked (%)	86.5	61.9	39.4	14.6	50.0
Degree-School Rank of New Hires	64.6	59.7	44.9	38.1	51.8
Has PhD Program and is Located in US (%)	69.6	57.4	41.4	11.1	42.0
Located in US (%)	85.0	95.8	86.9	91.8	89.9
Located in English-Speaking Country (%)	92.7	97.0	88.9	94.9	94.2
Number of hires	703	601	336	99	1739
Number of recruiters	302	88	65	31	486

Notes: This table reports average characteristics for the 487 hiring schools (recruiters) in the analysis sample, weighted by the number of hires. The first four columns show statistics for sub-samples that are defined based on the year the university first participated in the rookie camp as a recruiter. The final column shows overall sample means. A hiring school is defined as "highly ranked" if its BYU ranking is above the median of all placements in the sample, or roughly the 60^{th} percentile of hiring schools. Hiring schools with degree programs in the U.S. are also degree schools that produced job candidates in the analysis sample.

Table 3: Effects of ARC Adoption by Recruiter on Degree-School Rank of New Hires

	(1)	(2)	(3)	(4)	(5)	(6)
Recruiter Reputation	0.94**	0.91**	0.88**			
	(0.05)	(0.07)	(0.08)			
Post ARC Adoption (By Recruiter)	6.84**	5.73+	4.41	1.34	1.35	1.33
	(2.49)	(3.24)	(3.30)	(3.84)	(3.92)	(3.79)
Recruiter Reputation×Post ARC	-0.40**	-0.38**	-0.34**	-0.27+	-0.27+	-0.27+
	(0.11)	(0.14)	(0.13)	(0.16)	(0.16)	(0.15)
Recruiter Reputation×LM tightness					0.02	-0.10
					(1.30)	(1.02)
Observations	1,740	1,036	956	956	956	956
Adjusted R^2	0.262	0.207	0.198	0.258	0.257	0.257
Sample	All	Ever ARC	Ever ARC & Pre-Post	Ever ARC & Pre-Post	Ever ARC & Pre-Post	Ever ARC & Pre-Post
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Recruiter FE	No	No	No	Yes	Yes	Yes
LM Tightness Measure	None	None	None	None	Jobs/ PhDs	High-Rank/ Total Jobs

Notes: The table reports coefficient estimates from models as in equation (1) for the effect of recruiter participation in ARC on the rank of the degree-granting schools from which they hire. *Recruiter reputation* measures the predicted degree-school rank of new hires from a model fit to pre-ARC data (see Section 4.1 of text for details). This variable is demeaned before interacting with the ARC adoption dummy so the coefficients on *Post ARC Adoption* represent effects for recruiters ranked at the full sample mean of roughly 55 (meaning they are predicted to hire candidates from degree schools ranked at the 55^{th} percentile). Column (1) is estimated on the full sample, column (2) is limited to recruiters that eventually adopted ARC, and columns (3)-(6) are restricted to recruiters that adopted ARC and also hired new PhDs in both the pre- and post-ARC period. The models in columns (4)-(6) include recruiter fixed effects. Columns (5) and (6) each control for a measure of labor market tightness interacted with recruiter reputation. Labor market tightness is measured as either the ratio of academic jobs to new PhDs on the market (column 5) or the ratio of highly ranked jobs to total academic jobs (column 6). Parentheses report bootstrapped standard errors clustered on recruiter. + p<.10 * p<.05 ** p<.01.

Table 4: Effects of ARC Adoption by Degree School on a New PhD's Probability of Placement at a Highly Ranked University

	(1)	(2)	(3)	(4)
Degree-School Reputation	0.90**	0.91**	0.91**	
	(0.06)	(0.06)	(0.06)	
Post ARC Adoption (By Degree School)	0.16**	0.14**	0.15**	0.04
	(0.05)	(0.05)	(0.06)	(0.05)
Degree-School Reputation×Post ARC	-0.21*	-0.22*	-0.22*	-0.23**
	(0.08)	(0.09)	(0.10)	(0.08)
Observations	1,740	1,662	1,658	1,658
Adjusted R^2	0.241	0.229	0.230	0.282
Sample	All	Ever ARC	Ever ARC & Pre-Post	Ever ARC & Pre-Post
First Year ARC Controls	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Degree School FE	No	No	No	Yes

Notes: This table reports coefficient estimates from linear probability models as in equation (2) for the probability that a new PhD's first job is at a highly ranked university. Highly ranked universities are defined as those ranked above the median of all placements in the sample, or roughly the 60th percentile of hiring schools. *Degree-School Reputation* measures the predicted probability of a highly ranked job from a model fit to pre-ARC data (see Section 4.2 text for details). Column (1) is estimated on the full sample, column (2) is limited to degree schools that eventually adopted ARC, and columns (3)-(4) are restricted to schools that adopted ARC and also supplied new PhDs to the market in both the pre- and post-ARC period. All specifications include a dummy for the first year of ARC adoption by the degree school plus its interaction with *Degree-School Reputation*. Parentheses report bootstrapped standard errors clustered on degree school.
+ p<.10 * p<.05 ** p<.01.

Table 5: The Role of Connections in Hiring

	(1)	(2)	(3)	(4)	(5)
Same Rank Decile	0.608**	0.505**	0.517**	0.708**	0.667**
	(0.12)	(0.11)	(0.11)	(0.14)	(0.10)
Adviser Connection		5.207**	5.033**	5.746**	4.798**
		(0.46)	(0.46)	(0.58)	(0.41)
Log Distance			-0.685**	-0.814**	-0.765**
			(0.06)	(0.08)	(0.06)
Post ARC				-1.870**	-2.226**
				(0.44)	(0.48)
Same Rank Decile×Post ARC				-0.486**	-0.686**
				(0.17)	(0.18)
Adviser Connection×Post ARC				-1.315*	0.777
				(0.64)	(0.81)
Log Distance×Post ARC				0.285**	0.330**
				(0.06)	(0.07)
Observations	216,804	216,804	216,804	216,804	216,804
Adjusted R^2	-0.002	0.006	0.009	0.010	0.010
First Year ARC Controls	No	No	No	Yes	No
Year FE	Yes	Yes	Yes	Yes	Yes
Degree School FE	Yes	Yes	Yes	Yes	Yes
Recruiter FE	Yes	Yes	Yes	Yes	Yes
Source of variation in timing of ARC adoption	-	-	-	Degree School	Hiring School
Mean of Outcome×100	0.791	0.791	0.791	0.791	0.791

Notes: This table reports coefficient estimates from linear probability models for realized job matches as in equation (3). The sample of potential matches is constructed by matching each job candidate in the main analysis sample with every recruiting school that hired at least one new PhD in the year the candidate was on the market. The dependent variable is equal to one if the match was realized and zero otherwise. All coefficients are multiplied by 100. *Same Rank Decile* is an indicator for whether the candidate's degree school and recruiting school are in the same rank decile. *Adviser Connection* is an indicator for whether the candidate's dissertation chair had a coauthor on the faculty of the recruiting school. *Log Distance* is the log distance between the campuses of the degree school and recruiting school. *Post ARC* is an indicator for whether the degree school (in column 4) or the recruiter (in column 5) had participated in ARC in the current year or earlier. All specifications include year, degree school, and hiring school fixed effects. Column (4) also controls for the differential effect of the first year of ARC participation by the degree school. Parentheses report standard errors clustered on the degree school (column 4) or the recruiter (column 5). + p<.10 * p<.05 ** p<.01.

Table 6: The Effects of ARC on New Hire Productivity

		$\geq$ 1 Pre-Market Pub		$\geq$ 1 Post-Hire Pub	
	DS Rank	Any	Top-tier	Any	Top-tier
	(1)	(2)	(3)	(4)	(5)
Post ARC Adoption (by Recruiter)	1.343	0.109+	0.068	0.121*	0.092
	(3.899)	(0.066)	(0.056)	(0.059)	(0.056)
Post ARC×Recruiter Reputation	-0.267+	-0.004	-0.004+	-0.003	-0.004
	(0.149)	(0.003)	(0.002)	(0.003)	(0.003)
Observations	956	956	956	956	956
Adjusted R^2	0.258	0.060	0.094	0.118	0.192
Sample	Ever ARC & Pre-Post	Ever ARC & Pre-Post	Ever ARC & Pre-Post	Ever ARC & Pre-Post	Ever ARC & Pre-Post
Year FE	Yes	Yes	Yes	Yes	Yes
Recruiter FE	Yes	Yes	Yes	Yes	Yes
Mean of dependent variable	62.140	0.347	0.240	0.461	0.349
ARC Impacts by Recruiter Reputation:					
50th pctile	2.811	0.132+	0.092	0.136*	0.116+
	(4.394)	(0.072)	(0.061)	(0.067)	(0.063)
75th pctile	-3.861	0.029	-0.017	0.067	0.005
	(3.389)	(0.069)	(0.061)	(0.067)	(0.060)
90th pctile	-7.864+	-0.032	-0.082	0.026	-0.062
	(4.485)	(0.093)	(0.080)	(0.101)	(0.088)

Notes: The table reports coefficient estimates from models as in equation (1) for the effect of recruiter participation in ARC on the quality of new hires. All models include recruiter and year fixed effects as in Table 3, column (4), and are estimated on the sample of recruiters that adopted ARC and hired new PhDs in both the pre- and post-ARC periods. Column (1) reproduces Table 3 column (4); the dependent variable is the percentile rank of the school where the new hire received their PhD. In columns (2) and (3), the dependent variables are measures of candidate productivity prior to being hired: an indicator for having at least one publication (column 2) and an indicator for having at least one publication in a "top-tier" journal (column 3). The final two columns present models for post-hire productivity, using indicators for having at least one publication (column 4) or at least one top-tier publication (column 5) within the first three years of employment. The *Recruiter Reputation* is demeaned before interacting with the ARC adoption dummy so the coefficients on *Post ARC Adoption* represent effects for recruiters ranked at the full sample mean of roughly 55 (meaning they are predicted to hire candidates from degree schools ranked at the 55^{th} percentile). The bottom panel calculates ARC impacts for recruiters ranked at the 50^{th}, 75^{th}, and 90^{th} percentiles. Parentheses report bootstrapped standard errors clustered on recruiter.
+ p<.10 * p<.05 ** p<.01.

Table 7: Heterogeneity in ARC's Impact on Role of Degree-School Rank

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Post ARC Adoption	0.04	0.07	-0.04	0.02	0.06	0.12	-0.09
(By Degree School)	(0.05)	(0.08)	(0.09)	(0.13)	(0.18)	(0.11)	(0.10)
Degree-School Reputation	-0.23**	-0.33**	-0.00	-0.49**	-0.06	-0.22	0.16
× Post ARC	(0.08)	(0.11)	(0.18)	(0.16)	(0.35)	(0.18)	(0.23)
Observations	1,658	960	696	444	262	505	415
Adjusted R^2	0.282	0.274	0.283	0.323	0.218	0.269	0.357
First Year ARC Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Degree School FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sample Gender:	All	Male	Female	Male	Female	Male	Female
Sample Nationality:	All	- All -		- English -		Non-English	

Notes: This table reports coefficient estimates from linear probability models as in equation (2) for the probability that a new PhD's first job is at a highly ranked university. The first column reproduces the model from Table 4 column (4) (see Table 4 note for details); it is estimated on the full sample of candidates whose degree schools adopted ARC within our sample period and produced PhD candidates in both the pre- and post-ARC years. The next two columns estimate the model separately for men (column 2) and women (column 3). The remaining columns show estimates from the same model for four mutually exclusive sub-samples: male candidates within English-sounding names (column 4); female candidates with English-sounding names (column 5); male candidates with non-English names (column 6); and female candidates with non-English names (column 7). Parentheses report bootstrapped standard errors clustered on degree school.
+ p<.10 * p<.05 ** p<.01.

A Online Appendix: Better in Person? The Effects of In-Person Screening on Hiring Outcomes

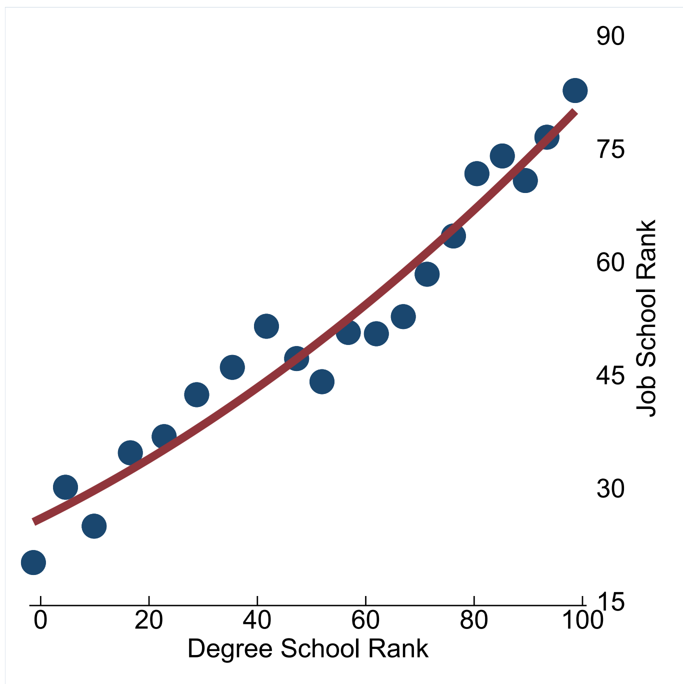
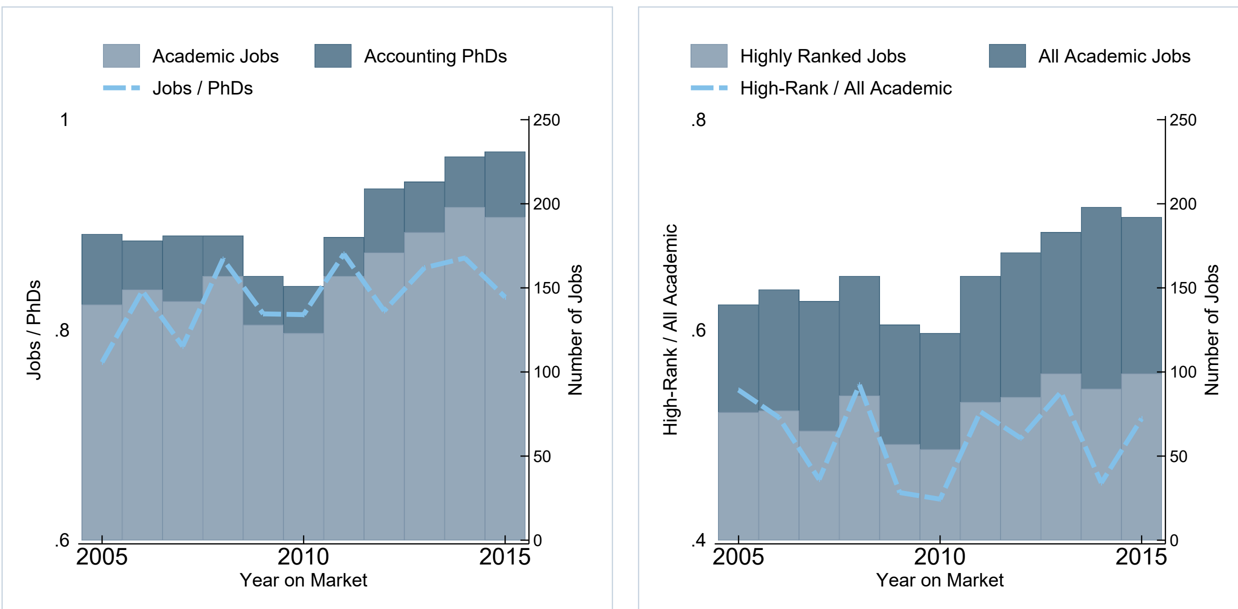


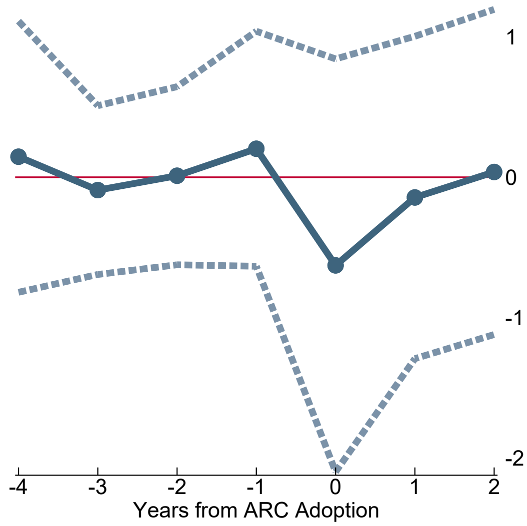
Figure A.1: This figure is a binned scatter plot of the rank of the hiring institution (Job School Rank) against the rank of the degree-granting school for all new accounting PhDs from 2005-2015 whose first job was at an academic institution (N=1,739). Bin means are adjusted for year fixed effects. The line shows the quadratic fit.



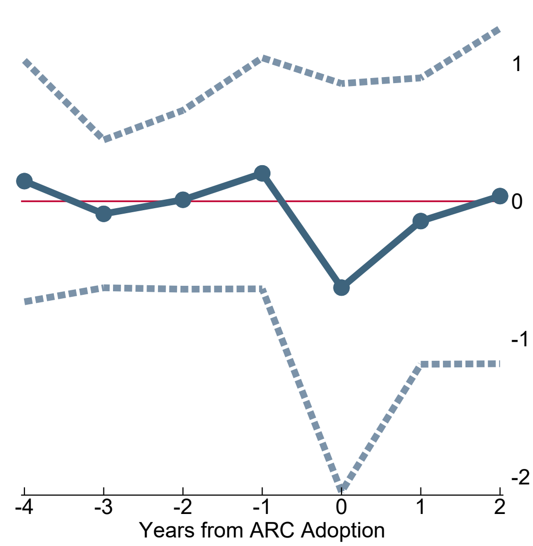
(a) Market for All Academic Jobs

(b) Market for Highly Ranked Jobs

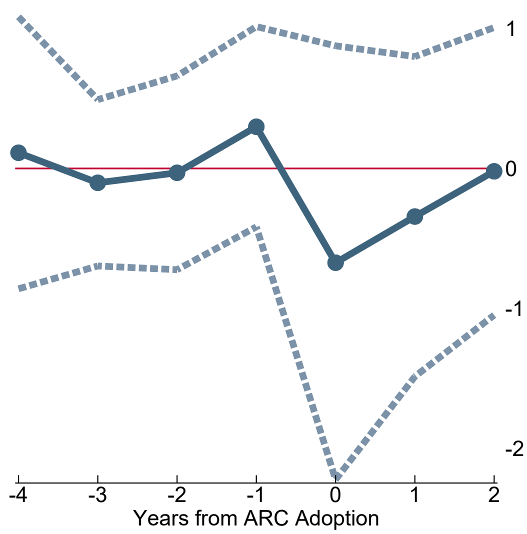
Figure A.2: This figure illustrates annual variation in labor market conditions in the market for accounting PhDs. Panel (a) compares the number of academic jobs to the total number of new PhDs for each year in the analysis sample. Panel (b) compares the number of jobs at highly-ranked schools ("highly ranked jobs") to the total number of academic jobs.



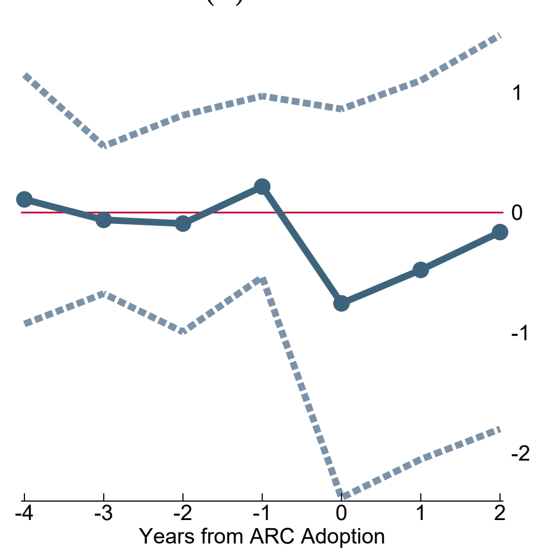
(a) Model 1



(b) Model 2

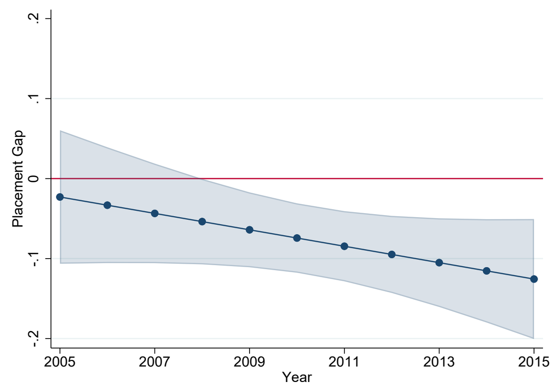


(c) Model 3

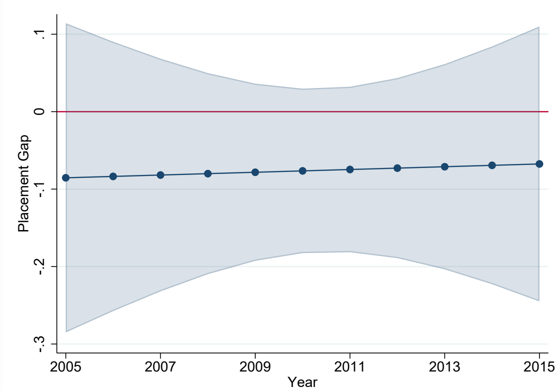


(d) Model 4

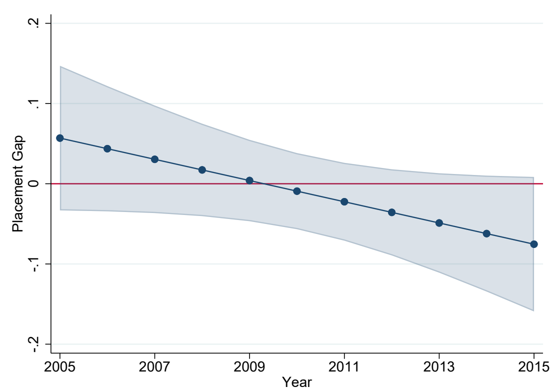
Figure A.3: This figure plots the coefficient estimates and 95% confidence intervals from models for the effects of ARC adoption by recruiters. The models are similar to those shown in Figure 4, which illustrate how the relationship between *Recruiter Reputation* and the degree-school rank of new hires evolves over time, except that here, the models are estimated using only the first treated cohort (i.e., recruiters who adopted ARC in 2011) and the pre-adoption years of the last treated cohort (i.e., recruiters who adopted in 2015). See note to figure 4 for details.



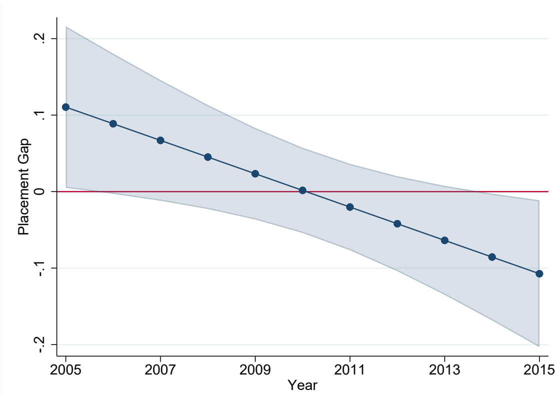
(a) Female (vs. Male)



(b) Black/Hispanic (vs. White)



(c) Asian (vs. White)



(d) Chinese (vs. English)

Figure A.4: This figure graphs estimated gaps in the probability of obtaining a first job at a highly-ranked university among new PhD's who graduated from similarly ranked degree schools. The estimates are from linear probability models that control for the rank of a candidate's degree school as well as a dummy variable for the indicated group, a linear time trend, and the interaction of the trend with the group dummy. Panel (a) is estimated using the full sample (N=1,740); Panel (b) uses only candidates whose name predicts Black race or Hispanic ethnicity and White candidates (N=1,225); Panel (c) uses only candidates whose race is predicted to be Asian or White (N=1,654); and Panel (d) uses only candidates whose national origin is predicted to be either Chinese or English (N=1,162). The shaded band in each panel shows the 95% confidence intervals for the estimated placement gaps as they evolve over time.

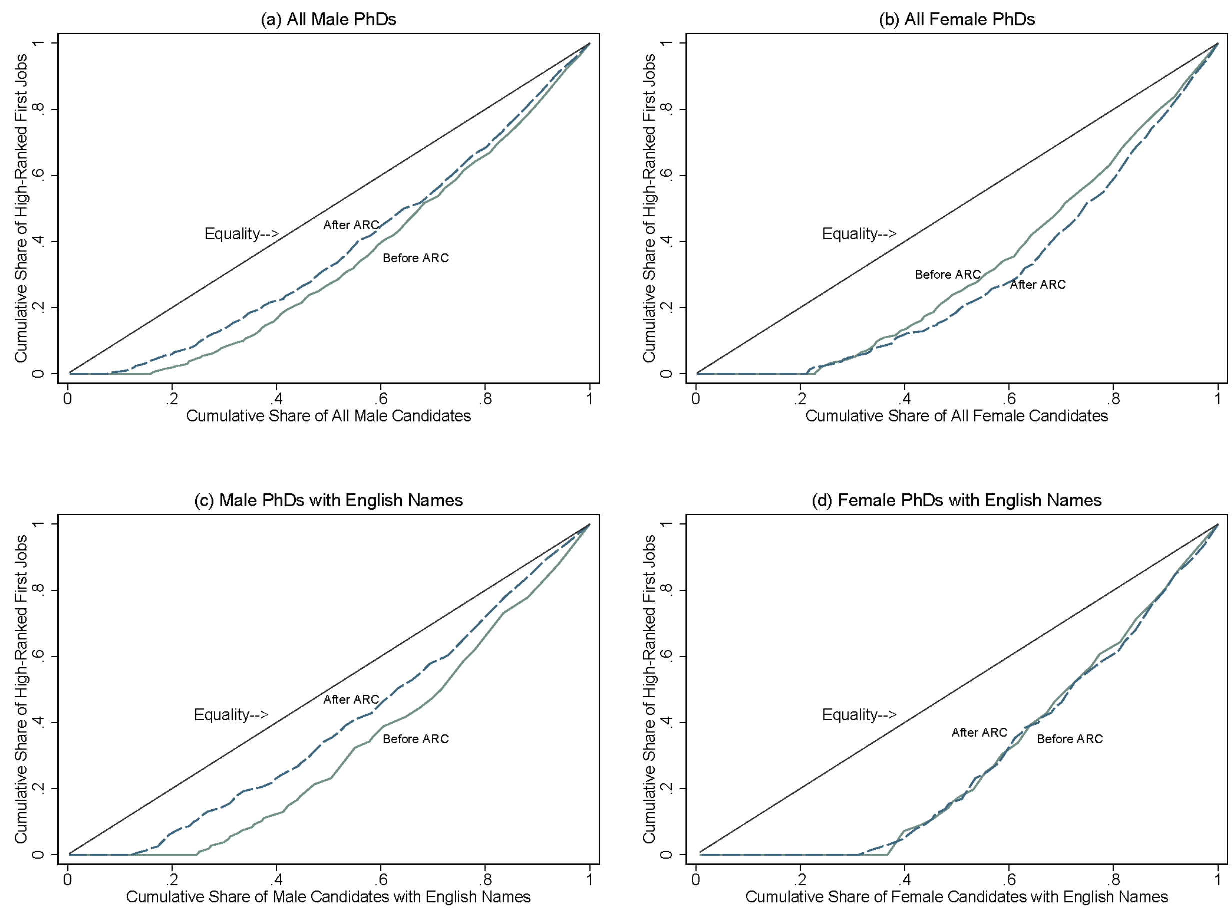
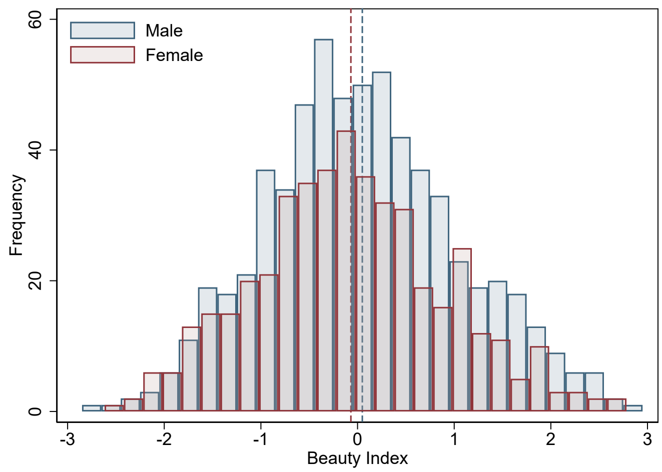
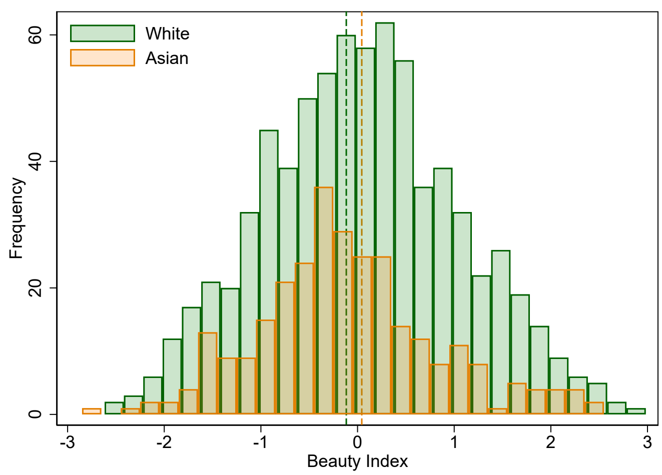


Figure A.5: Distribution of Highly Ranked Jobs, Pre- and Post- ARC

Figure A.6: This figure xxx.



(a) Beauty Scores by Gender



(b) Beauty Scores by Predicted Race

Figure A.7: This figure shows the frequency distribution of beauty scores, by gender (a) and predicted race (b), for the candidates in our sample for whom a photograph was available. The sample in figure (a) includes N=xxx candidates for whom the candidates gender could be identified; the sample if figure (b) includes N=xxx whose race was predicted to be either White or Asian. Vertical dashed lines show the sample means by sub-group. Photographs were collected from individuals' personal academic websites or LinkedIn profiles, or from their online faculty profiles posted on their universities' websites. Beauty scores were determined using XXXX.

Table A.1: Degree-School Based Models for ARC's Impact on Placements: Alternate Outcome Measures

	(1) High-Rank Job	(2) Hiring School Rank	(3) Hiring School is ARC Adopter
Post ARC Adoption (By Degree School)	0.04	1.85	0.12+
	(0.05)	(3.86)	(0.06)
Degree-School Reputation × Post ARC	-0.23**	-0.15*	0.49**
	(0.08)	(0.07)	(0.09)
Observations	1,658	1,658	1,658
Adjusted R^2	0.282	0.323	0.333
Sample	Ever ARC & Pre-Post	Ever ARC & Pre-Post	Ever ARC & Pre-Post
First Year ARC Controls	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Degree School FE	Yes	Yes	Yes

Notes: This table reproduces the degree-school based models of ARC's impact on job placements (equation 2 and Table 4) using alternate outcome measures. Column (1) replicates Table 4, column 4 in which the outcome is a binary measure of placement at a highly ranked university. In column (2), the dependent variable is the percentile rank of the hiring school. In column (3), it is an indicator for whether the hiring school participated at the rookie camp. All models include year and degree-school fixed effects as in column (4) of Table 4. Parentheses report bootstrapped standard errors clustered on degree school. + p<.10 * p<.05 ** p<.01.

Table A.2: ARC's Effect on the Return to Pre-Market Publications

	(1)	(2)	(3)
Post ARC Adoption (By Degree School)	0.03	0.05	0.05
	(0.05)	(0.06)	(0.06)
Candidate Has Any Pre-Market Pub	0.07*	0.08+	0.08+
	(0.03)	(0.05)	(0.05)
Candidate Has Top-tier Pub	0.15**	0.18**	0.16*
	(0.04)	(0.06)	(0.06)
Candidate Has Pre-Market Pub × Post ARC		-0.04	-0.04
		(0.08)	(0.08)
Candidate Has Top-tier Pub × Post ARC		-0.01	0.02
		(0.08)	(0.08)
Degree-School Reputation × Post ARC			-0.21*
			(0.08)
Observations	1,658	1,658	1,658
Adjusted R^2	0.303	0.303	0.304
Sample	Ever ARC & Pre-Post	Ever ARC & Pre-Post	Ever ARC & Pre-Post
First Year ARC Controls	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Degree School FE	Yes	Yes	Yes

Notes: This table reports coefficient estimates from degree-school based models of ARC's impact on the probability that a new PhD's first job is at a highly-ranked university (see equation 2 and Table 4) using specifications with degree-school fixed effects as in column (4) of Table 4. Column (1) shows the average association between the outcome and two measures of a candidate's publication output as of their graduation year: (i) a dummy for having at least one publication in any peer-reviewed journal, and (ii) a dummy for having at least one publication in a top-tier accounting journal. Column (2) adds the interactions of each of the publication variables with a the dummy indicating whether the candidate's degree school had previously adopted ARC. Finally, column (3) controls for the interaction of the ARC adoption indicator with the candidate's Degree-School Reputation (see text and Table 4 note). Parentheses report bootstrapped standard errors clustered on degree school. + p<.10 * p<.05 ** p<.01.

Table A.3: ARC's Impact on Placement Gaps by Race & Nationality

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Group Indicator	-0.16*	-0.15*	-0.03	-0.04	0.01	0.00	-0.01	-0.03
	(0.07)	(0.07)	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)	(0.03)
Group Indicator × Post-ARC	0.09	0.08	-0.10*	-0.09+	-0.10*	-0.09+	-0.16**	-0.14**
	(0.13)	(0.13)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)
Observations	1,161	1,161	1,578	1,578	1,658	1,658	1,110	1,110
Adjusted R^2	0.277	0.300	0.286	0.308	0.283	0.305	0.296	0.316
First Year ARC Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pre-Market Pub Controls	No	Yes	No	Yes	No	Yes	No	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Degree School FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Group Indicator = 1 if:	Black/Hispanic (vs. White)	Black/Hispanic (vs. White)	Asian (vs. White)	Asian (vs. White)	Non-English (vs. English)	Non-English (vs. English)	Chinese (vs. English)	Chinese (vs. English)

Notes: This table reports coefficient estimates from degree-school based models of ARC's impact on the probability that a new PhD's first job is at a highly-ranked university (see equation 2 and Table 4) using specifications with degree-school fixed effects as in column (4) of Table 4. All specifications control for the interaction of the candidate's degree-school reputation with the post-ARC dummy, and hence measure ARC's impact on placement gaps among candidates from similarly ranked degree schools. Odd-numbered columns show the estimates used to construct the pre- and post-ARC placement gaps by demographic group shown in Figure 5. Even-numbered columns add controls for pre-market publication dummies (see Table 6) and the interactions of these dummies with the post-ARC adoption dummy. Parentheses report bootstrapped standard errors clustered on degree school. + p<.10 * p<.05 ** p<.01.

Table A.4: Predicted Nationalities & Language Distances from English

Predicted Nationality (Based on Name)	Language Distance from English	Undergrad Instruction in English (%)	Frequency
English	0	94.9	767
Nordic	26.7	89.5	20
Dutch	27.2	91.7	13
African	27.5	63.6	16
German	30.8	88.3	85
Italian	47.8	68.8	18
French	48.7	80.0	52
Romanian	54.9	0.0	4
Hispanic	57	61.0	64
Slav	60.3	67.4	49
Indian	65.2	89.4	59
Greek	69.9	45.5	11
Chinese	82.4	13.8	395
Arab	83.6	54.8	41
Hungarian	87.9	100.0	4
Indonesian	87.9	100.0	1
Japanese	88.3	60.0	6
Korean	90	20.0	76
Israeli	91.1	34.8	28
Turkish	92	13.0	24
Thai	92.9	16.7	7

Notes:This table reports the various predicted nationalities obtained via candidates' last names and the distance of their predicted language from English. The language distance is based on the ease with which an English speaker can learn the language. We also report the percent of the candidates of each of the various predicted nationalities that attended an undergrad institution taught in English and the frequency of each predicted nationality in the sample.

Table A.5: ARC's Impact on Placement Gaps by Language Distance from English

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Language Variable	0.04	0.03	-0.01	-0.01	0.00	-0.01	0.00	-0.01
	(0.04)	(0.04)	(0.08)	(0.08)	(0.04)	(0.03)	(0.04)	(0.04)
Language Variable×Post-ARC	-0.04	-0.04	-0.05	-0.02	-0.19**	-0.16**	-0.16**	-0.13*
	(0.07)	(0.07)	(0.11)	(0.11)	(0.05)	(0.05)	(0.06)	(0.06)
Observations	1,099	1,099	825	825	1,190	1,190	1,658	1,658
Adjusted R^2	0.266	0.283	0.260	0.278	0.311	0.329	0.285	0.306
First Year ARC Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pre-Market Pub Controls	No	Yes	No	Yes	No	Yes	No	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Degree School FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Language Variable:	= 1 if Closer to English (vs. English Native)		= 1 if Further from Engl. & English UG		= 1 if Further from Engl. & Non-English UG		Distance from English (min=0; max=1)	

Notes: This table reports coefficient estimates from degree-school based models of ARC's impact on the probability that a new PhD's first job is at a highly-ranked university (see equation 2 and Table 4) using specifications with degree-school fixed effects as in column (4) of Table 4. All specifications control for the interaction of the candidate's degree-school reputation with the post-ARC dummy, and hence measure ARC's impact on placement gaps among candidates from similarly ranked degree schools. Odd-numbered columns show the estimates used to construct the language-based placement gaps shown in Figure 6. See figure 6 note for definitions of language variables. Even-numbered columns add controls for pre-market publication dummies (see Table 6) and the interactions of these dummies with the post-ARC adoption dummy. Parentheses report bootstrapped standard errors clustered on degree school. + p<.10 * p<.05 ** p<.01.

Table A.6: ARC's Impact on Placement Gaps by Language Distance, Alternative Outcomes

	Highly Ranked First Job		Hiring School Percentile Rank		Highly Ranked & English-speaking		Highly Ranked & Non-English	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Language Distance	0.00	-0.01	-0.03	-0.04	-0.06	-0.08+	0.07**	0.07**
	(0.04)	(0.03)	(0.03)	(0.02)	(0.04)	(0.04)	(0.02)	(0.02)
Language Distance×Post-ARC	-0.16**	-0.13*	-0.11**	-0.10*	-0.19**	-0.17*	0.03	0.03
	(0.05)	(0.06)	(0.04)	(0.04)	(0.06)	(0.07)	(0.03)	(0.03)
Observations	1,658	1,658	1,658	1,658	1,658	1,658	1,658	1,658
Adjusted R^2	0.285	0.306	0.328	0.347	0.260	0.282	0.056	0.055
First Year ARC Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pre-Market Pub Controls	No	Yes	No	Yes	No	Yes	No	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Degree School FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table reports coefficient estimates from degree-school based models of ARC's impact on the quality of a candidate's first job. This table reproduces the degree-school based models of ARC's impact on job placements, and its moderating effect on the role of language, as in Table A.6, using alternate outcome measures. Columns (1) and (2) replicate Table A.6 columns 7 and 8, in which the outcome is a binary measure of placement at a highly ranked university. All models include year and degree-school fixed effects as in column (4) of Table 4. Even-numbered columns add controls for pre-market publication dummies (see Table 6) and the interactions of these dummies with the post-ARC adoption dummy. Parentheses report bootstrapped standard errors clustered on degree school. + p<.10 * p<.05 ** p<.01.

Table A.7: Gaps in Writing Quality by Language Distance from English

	(1) Language complexity	(2) Grammar mistakes	(3) Rate of typos
Language Distance	0.12	-0.01	-0.10
	(0.10)	(0.12)	(0.09)
Language Distance×Post-ARC	0.01	-0.23	-0.02
	(0.16)	(0.19)	(0.04)
Observations	1,235	1,235	1,235
Adjusted R^2	0.156	0.121	0.046
First Year ARC Controls	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Degree School FE	Yes	Yes	Yes

Notes: This table reports coefficient estimates from models similar to those in columns 7 & 8 of Appendix Table A.6, except that the dependent variable is a measure of the quality of written English as measured by textual analysis of the candidate's PhD thesis. The sample includes 1,235 candidate's for whom the thesis was published in ProQuest or otherwise available online. The coefficients represent differential quality of written English associated with a one-unit increase in the Distance from English of the candidate's predicted native language, and the change in this relationship after the degree school began participating in ARC. Parentheses report bootstrapped standard errors clustered on degree school. + p<.10 * p<.05 ** p<.01.

Table A.8: Role of English Language Signals for Candidates with Chinese Names

	(1)	(2)	(3)	(4)	(5)	(6)
Language Signal	-0.08	-0.11	0.02	0.00	-0.01	-0.02
	(0.09)	(0.09)	(0.06)	(0.06)	(0.05)	(0.05)
Language Signal×Post-ARC	0.13	0.18	-0.37**	-0.34**	-0.12*	-0.09+
	(0.12)	(0.12)	(0.10)	(0.11)	(0.06)	(0.05)
Observations	753	753	789	789	945	945
Adjusted R^2	0.257	0.275	0.283	0.298	0.290	0.307
First Year ARC Controls	Yes	Yes	Yes	Yes	Yes	Yes
Pre-Market Pub Controls	No	Yes	No	Yes	No	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Degree School FE	Yes	Yes	Yes	Yes	Yes	Yes
Language Signal:	English-speaking Undergrad School		English-sounding First Name		No English Language signal	

Notes: This table reports coefficient estimates from degree-school based models of ARC's impact on the probability that a new PhD's first job is at a highly-ranked university (see equation 2 and Table 4) using specifications with degree-school fixed effects as in column (4) of Table 4. All specifications control for the interaction of the candidate's degree-school reputation with the post-ARC dummy, and hence measure ARC's impact on placement gaps among candidates from similarly ranked degree schools. Odd-numbered columns show the estimates used to construct the pre- and post-ARC placement gaps by demographic group shown in Figure 7. Even-numbered columns add controls for pre-market publication dummies (see Table 6) and the interactions of these dummies with the post-ARC adoption dummy. Parentheses report bootstrapped standard errors clustered on degree school. + p<.10 * p<.05 ** p<.01.

Table A.9: ARC's Impact on Placement Gaps by Gender

	(1)	(2)	(3)	(4)	(5)	(6)
Group Indicator	-0.04	-0.04	0.02	0.01	-0.02	-0.02
	(0.05)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
Group Indicator×Post-ARC	-0.12	-0.09	-0.14*	-0.12+	-0.13+	-0.12
	(0.09)	(0.08)	(0.06)	(0.07)	(0.07)	(0.08)
Observations	727	727	964	964	873	873
Adjusted R^2	0.278	0.291	0.278	0.298	0.332	0.338
First Year ARC Controls	Yes	Yes	Yes	Yes	Yes	Yes
Pre-Market Pub Controls	No	Yes	No	Yes	No	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Degree School FE	Yes	Yes	Yes	Yes	Yes	Yes
Group Indicator = 1 if:	Female with English Name (vs. Male, Engl Name)		Male with Non-English Name (vs. Male, Engl Name)		Female with Non-English Name (vs. Male, Engl Name)	

Notes: This table reports coefficient estimates from degree-school based models of ARC's impact on the probability that a new PhD's first job is at a highly-ranked university (see equation 2 and Table 4) using specifications with degree-school fixed effects as in column (4) of Table 4. All specifications control for the interaction of the candidate's degree-school reputation with the post-ARC dummy, and hence measure ARC's impact on placement gaps among candidates from similarly ranked degree schools. Odd-numbered columns show the estimates used to construct the pre- and post-ARC placement gaps by demographic group shown in Figure 8. Even-numbered columns add controls for pre-market publication dummies (see Table 6) and the interactions of these dummies with the post-ARC adoption dummy. Parentheses report bootstrapped standard errors clustered on degree school. + p<.10 * p<.05 ** p<.01.

Table A.10: The Effects of ARC on New Hire Demographics

Dependent Variable:	Male with English name (vs. other)	Female (vs. Male)	Asian (vs. White)	Chinese name (vs. English name)	Chinese last & Engl. first name (vs. Engl. only)	Beauty Score (Std Dev)
	(1)	(2)	(3)	(4)	(5)	(6)
Post ARC Adoption (by Recruiter)	0.158*	-0.095	-0.068	-0.113	-0.177**	0.174
	(0.061)	(0.074)	(0.058)	(0.070)	(0.057)	(0.166)
Post ARC×Recruiter Reputation	0.002	-0.002	-0.002	-0.002	0.000	-0.002
	(0.003)	(0.003)	(0.003)	(0.004)	(0.003)	(0.007)
Observations	956	956	913	623	456	662
Adjusted R-squared	0.069	0.011	0.142	0.191	0.256	0.078
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Recruiter FE	Yes	Yes	Yes	Yes	Yes	Yes
Mean of dependent variable	0.302	0.384	0.290	0.313	0.107	0.00
ARC Impacts by Recruiter Reputation:						
50th pctile	0.149*	-0.082	-0.056	-0.101	-0.178**	0.187
	(0.068)	(0.078)	(0.065)	(0.081)	(0.068)	(0.187)
75th pctile	0.189**	-0.138+	-0.111+	-0.154*	-0.173**	0.126
	(0.068)	(0.084)	(0.060)	(0.078)	(0.063)	(0.146)
90th pctile	0.213*	-0.172	-0.144+	-0.185	-0.170+	0.089
	(0.099)	(0.111)	(0.087)	(0.120)	(0.102)	(0.196)

Notes: The table reports coefficient estimates from models as in equation 1 for the effect of recruiter participation in ARC on the quality of new hires. All models include recruiter and year fixed effects as in Table 3, column (4), and are estimated on the sample of recruiters that adopted ARC and hired new PhDs in both the pre- and post-ARC periods. In columns (1)-(5), the dependent variables are dummy variables equal to 1 if the new hire: (1) is male with an English name (and zero otherwise); (2) is female (and zero if male); (3) is predicted to be Asian (and zero if their predicted race is White); (4) has a Chinese name (and zero if their name is English); (5) has a Chinese last name but an English first name or nickname (and zero if their name is English only). In column (6), the dependent variable is the candidate's beauty index (see xxxx). The *Recruiter Reputation* index is demeaned before interacting with the *Post ARC* adoption dummy; so that the coefficients on *Post ARC Adoption* in the first row represent effects for recruiters ranked at the full sample mean, or roughly 55th percentile of *Recruiter Reputation.* The bottom panel calculates ARC impacts for recruiters ranked at the 50th, 75th, and 90th percentiles. Parentheses report bootstrapped standard errors clustered on recruiter.
+ p<.10 * p<.05 ** p<.01.