Brodeur, Abel; Kattan, Lamis; Musumeci, Marco

Working Paper

Job Market Stars

GLO Discussion Paper, No. 1514

Provided in Cooperation with:
Global Labor Organization (GLO)

Job Market Stars

Abel Brodeur Lamis Kattan Marco Musumeci

October 25, 2024

Abstract

Graduating economics PhDs face intense competition when seeking faculty or research positions at universities and research institutions. We examine the relationship between statistically significant results, arguably used as indicators of research quality in a competitive academic market, and academic hiring outcomes. We start by investigating the determinants of academic success by analyzing 604 job market papers (JMPs) from 2018-2019 to 2020-2021. We then turn to the importance of statistical significance focusing on 150 empirical JMPs. We find evidence that 'marginally' significant results in JMPs are associated with higher academic placement likelihoods. During the COVID-19 pandemic, a tighter job market strengthened this relationship without altering the p-hacking behavior of PhD candidates, suggesting that our results reflect a recruitment bias by academic employers. We also find evidence of publication bias, suggesting that recruiters may use statistical significance to gauge candidates' potential for future publications, thus influencing recruitment decisions. Overall, our findings provide insights into the dynamics of the academic job market and highlight incentives that would potentially reward academics for questionable research practices.

KEYWORDS: Academic job market - p-Hacking - Publication bias - Research credibility

JEL CODES: B41, C13, C40, C93.

Authors: Brodeur: University of Ottawa, Global Labor Organization and Institute for Replication. E-mail: abrodeur@uottawa.ca. Kattan: Georgetown University in Qatar and Global Labor Organization. lk811@georgetown.edu. Musumeci: Erasmus University Rotterdam. musumeci@ese.eur.nl. We are grateful to Jonathan Créchet, Jishnu Das, Tomas Havránek, Zuzana Iršová, Thomas Lemieux, Louis-Philippe Morin, and seminar participants at the Erasmus School of Economics for very helpful advice and suggestions. Prajwal Bista and Olesia Savka provided excellent research assistance. The pre-analysis plan (https://osf.io/pe392/) for our study was written in March 2023. We had not collected the data nor conducted any empirical analysis at the time of pre-registration. Moreover, many Job Market papers had yet to be published. Nonetheless, our study should be regarded as exploratory since we cannot prove definitively that the PAP was posted prior to conducting any analyses. Kattan acknowledges financial support from Georgetown University in Qatar. Errors are ours.

1 Introduction

Aspiring economists often encounter fierce competition when pursuing faculty positions or research roles in universities and research institutions. This competitiveness potentially stems from several factors, including candidates' characteristics, the prestige of their PhD institution, and the limited availability of faculty positions relative to the number of qualified PhD graduates. Additionally, given the requirement to publish in reputable journals and maintain a record of impactful research, succeeding in the academic job market may require not only academic pedigree, networking, and effective communication but also a very strong Job Market Paper (JMP) that reflects exceptional research skills.[1]

We hypothesize that the competitiveness of the academic job market, along with the weight carried by the job market paper in reflecting a candidate's potential, can lead to a positive association between statistical significant results in the JMP –often marked by stars in estimates– and academic hiring. The underlying assumption is that significant results increase the likelihood of publishing in top journals (Brodeur et al. (2023); Chopra et al. (2024)). On the one hand, the pressure to produce "publishable" results and secure a competitive edge in the job market can incentivize some junior researchers to engage in questionable research practices, such as p-hacking in their JMP (i.e., manipulation and/or selective reporting of results' p-values) and/or desk-drawer effects (i.e., ceasing to pursue projects with insignificant results) as a mean of presenting more compelling findings.[2] On the other hand, academic institutions and journals may play a crucial role in fostering a research culture that prioritizes statistically significant results on the academic market.[3] For instance, academic institutions may be more likely to interview and offer a contract

[1]See Cawley (2023) and Coles et al. (2010) for more information on the academic job market for economists.

[2]Additionally, the significant impact of the first placement on future careers (see, for instance, Oyer (2006)) further heightens the pressure on PhD students, potentially encouraging practices that prioritize career advancement over methodological integrity.

[3]While p-hacking and publication bias pose significant problems for the integrity of the literature as a whole, they do not necessarily mean that the methodological integrity of a given paper is compromised.

to a job market candidate (JMC) whose JMP finds a statistically significant effect, thus leading to recruitment bias.

In this paper, we aim to empirically examine these hypotheses by investigating the relationship between 'marginally' significant results and academic job market placement.[4]

To gain a better understanding of the academic job market in economics, we start by investigating the determinants of academic job market success for 604 graduating PhDs seeking jobs in academia or research intensive institutions from 12 universities. We exploit the internet archive to retrieve the list of job market candidates of these universities for the academic years from 2018-2019 to 2020-2021. The internet archive allows us to navigate the webpage dedicated to the presentation of the job market candidates (JMCs) in certain dates in the past. In this way we can access the list of JMCs that were posted on the university websites in the Fall preceding the three job market sessions considered in our analysis. These 12 institutions also provide placement information for JMCs.[5]

We first investigate the relationship between JMCs and JMP characteristics and academic placement.[6] In our sample, the most common placement is an assistant professor (AP) position, secured by about 36% of JMCs.[7] Another 25% of the JMCs remains in academia with a postdoc position (i.e., non-tenure academic placement). Around 39% of the candidates leave academia, with most of them (23% of the entire sample) obtaining a job in the private sector. In terms of placement, the majority of

[4]In our analysis, we examine the mass of test statistics in a narrow band just above and below a statistical significance threshold. We define marginally significant estimates as those with z-statistics just over the statistical threshold.

[5]Our analysis focuses on PhD students listed on departmental websites dedicated to the academic job market rather than all PhD students graduating from these 12 universities. To verify whether PhD candidates are initially interested in academic jobs, we conducted a brief survey among all job market candidates from the Top 100 economics departments during the 2022–2023 academic year, requesting them to rank their preferences for future placements. The survey was reviewed and found exempt by the Georgetown University IRB for Human Participants (IRB ID#:STUDY00006050). Approximately 65% of respondents ranked academic jobs as their most preferred placement. Further details can be found in Appendix Figure A1.

[6]We rely on two measures for academic placement: the likelihood of securing an academic placement (extensive measure) and the ranking of the research institution, conditional on securing an academic placement (intensive measure). Academic placement includes postdocs, APs, or a placement at a research institution that is listed in the Ideas/RePEc ranking.

[7]By AP position, we refer to tenure-track junior faculty positions.

AP placements are outside of the top 100, with only a few top departments placing their JMCs in top 10 institutions.

Our findings suggest that, conditional on an academic placement, female and white JMCs obtain a placement ranking that is around 35% and 30% higher than their male and non-white counterparts, respectively.[8] However, we find that white candidates are 8 percentage points less likely to obtain an AP position compared to non-white candidate.[9] Using the supervisor's number of coauthors as a proxy for network, we find weak evidence that a larger network is associated with higher placement ranking. Surprisingly, perhaps, the advisor's citation counts and the number of coauthors are not significantly correlated with academic placement's measures, whether accounting for the PhD institution rank or not.

On the JMPs characteristics, the presence of a theoretical model is positively related to the likelihood of securing an assistant professor position. Furthermore, when combined with empirical analysis, it is associated with a higher placement ranking. Finally, the ranking of the PhD institution appears to matter for academic placement, as a 1% increase in PhD ranking is associated with an approximately 8% increase in placement ranking. In contrast, the ranking of the institution where a candidate has completed the PhD is not a strong predictor of the chances of securing an AP position.[10]

We then turn to a subset of JMCs whose JMP is empirical to examine the relationship between statistical significance and academic placement. For this analysis, we harvest and analyze the universe of hypothesis tests drawn from 150 JMPs. The analysis includes 2,708 test statistics.

We first plot and compare visually the distribution of test statistics. The distribution displays a hump shape with a maximum at around the 10% statistical

[8] Of note, our sample does not include a sufficient number of minority candidates to further subdivide the non-white category into multiple groups.

[9] Our analysis provides weak evidence suggesting that female candidates are less likely to secure tenure-track positions. This finding partially aligns with recent studies on the gender gap in the economics academic job market (see, for instance, McFall et al. (2024)).

[10] This finding is perhaps not surprising, given that our sample includes only candidates from 12 of the top 100 economic institutions.

threshold, suggesting misallocated z-statistics. This result could be explained by a file-drawer bias, where candidates either shift to ideas with more significant results, or by p-hacking, where candidates manipulate their models to achieve significant results. This is concerning as it implies that PhD students anticipate recruitment biases, which in turn shape their methodologies and research processes.[11]

We then use a series of methodologies to formally document the extent of misallocation across subsamples. We find that statistical significance appears to matter for academic placement, as the density of the marginally significant tests is higher around significance thresholds for candidates placed in academia. The larger hump around the 10% significance threshold is for JMCs who succeeded in securing an AP position. This finding highlights recruitment bias in academia and suggests that p-hacking could potentially be a predictor of faculty placement.

To formally investigate this relationship, we rely on caliper tests where we restrict our sample to test statistics within a narrow range around the significance thresholds. We find that, conditional on placement determinants, marginally significant tests at the 10% level are associated with an increase in the likelihood of academic and assistant professorship placement by approximately 10 and 14 percentage points, respectively. These estimates are also economically significant as this effect represents approximately 20% of the likelihood of academic and 50% of assistant professorship placement. While these results do not directly measure the causal effect of p-hacking on placement outcomes, they provide suggestive evidence that (i) recruiting bias is present in academic placement – potential preferences for candidates with significant results in the JMP or (ii) JMCs targeting academic careers selectively choose their results based on significance.[12]

[11] Another explanation could be that PhD students, even highly qualified ones, with insignificant results, self-select out of the job market. This is also problematic, as it suggests that field norms contribute to these non-optimal patterns.

[12] We also examine the correlations between marginally significant results and observable characteristics, including those related to the JMP, JMC, PhD institution, and PhD advisor. Our analysis reveals that none of these observables strongly predict statistical significance. Although the absence of strong correlations does not allow us to claim that p-hacking is exogenous, it lends greater credibility to our initial findings on the correlations between p-hacking and academic placement.

To get insights on whether the positive association between statistical significance and academic placement is driven by recruiter bias in academia or a preference among PhD candidates targeting academic careers for significant results, we use the COVID-19 pandemic as an unexpected event that disrupted the equilibrium of the academic labor market by decreasing demand for academic employment. The academic job market at this time was slack and highly competitive for JMCs, potentially leading academic employers to become more selective in their hiring processes. While our findings do not indicate that PhD candidates altered their ethical decision-making in response to increased competitiveness, we observe a stronger relationship between statistical significance at the 10% threshold and the likelihood of academic employment during the academic year 2020-2021. This phenomenon may be attributed to the unexpected nature of the pandemic's impact on the job market, coupled with the limited window for JMCs to adapt their research strategies once the competitiveness of the job market becomes apparent.[13] These results thus provide evidence that the positive relationship is (at least partially) driven by academic employers' preference for marginally significant results (i.e., recruitment bias).

Finally, if our assumption that academic employers rely on the statistical significance of estimates in JMPs as a gauge for future publication holds, we should find evidence of publication bias in JMPs. To empirically examine this, we explore the relationship between statistical significance and "future" publication of JMPs. We search all 150 researchers' CVs and webpages for information on whether their JMP is under revision (R&R) or published. Our estimates indicate that test statistics marginally significant at the 10% level are significantly associated with approximately a 10-percentage-point higher likelihood of publication, equating to roughly a 35% increase. This finding supports our hypothesis that perceived pressure and journal/research institution's behavior act as a primary mechanism for recruitment

[13]The relatively small decline in the labor supply of graduate students in the year following the pandemic further suggests that PhD candidates did not fully anticipate the significant impact the pandemic would have on the academic job market.

bias.

Our study contributes to existing research on p-hacking and publication bias in several ways. First, our findings indicate that specification searching and p-hacking undermine the credibility of research in the early stages of an academic career. This result adds to a growing literature documenting the determinants and the extent of p-hacking for published and unpublished research (Abadie (2020); Andrews and Kasy (2019); Askarov et al. (2023); Brodeur et al. (2016); Brodeur et al. (2020); Christensen and Miguel (2018); Doucouliagos and Stanley (2013); Furukawa (2024); Gerber and Malhotra (2008); Havránek (2015); Havránek et al. (Forthcoming); Kepes et al. (2022); Miguel (2021); O'Boyle Jr et al. (2017); Olsen et al. (2019); Stanley et al. (2024)).[14] Second, we rely on the COVID-19 pandemic as a demand shock on the number of positions to unpack the role of authors and research institutions in fueling p-hacking and publication bias. Our results add to Brodeur et al. (2023) who provide evidence that selective reporting of test statistics in published research cannot be fully attributed to a publication bias in peer review. Our findings also add and support DellaVigna and Linos (2022)'s conclusions that inflated effect sizes and publication bias are more prevalent for nudge trials published in academic journals in comparison to a set of trials ran by two large Nudge Units in government.

Our findings also relate to a large literature documenting the determinants of academic job market placement (e.g., Chen et al. (2013); Conley et al. (2016); Conti and Visentin (2015); Eberhardt et al. (2023); Fortin et al. (2021); Ge and Wu (forthcoming); Grogger and Hanson (2015); Hadlock and Pierce (2021); Hilmer and Hilmer (2012); Jones and Sloan (Forthcoming)).[15] We add to this literature by relying on departmental lists of graduating PhDs seeking jobs in academia as opposed to only economists that secured a position in academia.

[14] Our results also relate to a growing literature on reproducibility and replicability, and more generally research credibility in economics (see, for instance, Ankel-Peters et al. (2023), Camerer et al. (2016), Camerer et al. (2019) and Drazen et al. (2021)).

[15] Our results also relate to a literature documenting the determinants of productivity differences for early career economists (e.g., Conley and Önder (2014); García-Suaza et al. (2020)).

2 Data

We focus on leading economics departments. Our sample includes JMCs from 12 universities ranked among the top 100 economic institutions based on the IDEAS/ RePEc classification (August 2023).[16] We selected institutions for which the title of the previous job market papers remained available in the internet archive. These institutions also provide placement details of their PhD candidates. As stated in our pre-analysis plan (PAP), the selected economics departments are: Boston University, University of Chicago, University of Columbia, Cornell University, Harvard University, Michigan State University, Massachusetts Institute of Technology, New York University, Princeton University, Stanford University, University of California, Los Angeles and the University of Michigan. The median rank of these universities in the IDEAS/RePEc institution ranking is 20.5: 4 institutions are in the top 10, while the others are distributed between the 11th and 63rd rank.

We follow our PAP and keep three academic years in our final sample: 2018-2019; 2019-2020; and 2020-2021. These years were chosen as we could not find job market papers for most departments prior to the academic year 2018-2019. The JMPs are obtained searching the internet. In cases where only one version of the job market paper is available, we include this in our sample, regardless of when it was made accessible. Whenever newer versions of a JMP are available, we always select the one with the closest date to the Job Market, which we set in November prior to the market.[17]

For each JMP, we additionally record: whether it is published and in which journal; whether it is an empirical or theoretical paper; the number of authors; the names, the number of citations and the number of co-authors of the JMC's supervisors; PhD affiliation(s) of authors; and the method used (e.g., difference-in-differences, randomized control trials, instrumental variables). We also collect data on additional candidate characteristics, such as gender and race, by visually

[16] Available at https://ideas.repec.org/top/top.inst.all.html.

[17] When the JMP is found in the PhD candidate's thesis, the month of submission is sometimes unknown. We assume in these cases that the thesis was written in August of that year.

assessing the images available on their webpage.

We retrieve data on the placement of each candidate from the university webpages. When not available from this source, we exploit other online resources, such as the JMC website or LinkedIn. We classify placements as academic only if a candidate obtains a postdoctoral fellowship (postdoc) or an Assistant Professor (AP) position. Research positions at central banks or other institutions are not classified as academic. When a candidate's first placement is a postdoc, we check whether this is a one-year postdoc followed by an AP. In our main analyses, one-year postdocs followed by an AP position are classified as postdocs, but we check that our results are robust to considering these candidates as AP.

To select 150 JMPs for our p-hacking analysis, we proceed with a stratified random sampling , with institutions being our strata. For each stratum, we compute the share of JMCs and randomly select JMPs to form a representative sample. We restrict our selection to JMPs in applied micro/macro economics. We classify each paper of this sub-sample, in one of the following broad fields[18]: i) Education, Labor and Health ii) Macro, Finance and IO iii) Trade, Urban, Growth and Development iv) Public Economics and Political Economy v) Mixed, when the topic of a paper is at the intersection between two or more of the previous fields.

We present summary statistics for both the selected and full samples in columns (1)-(2) and (3)-(4) of Table 1, respectively. In the full sample, including 604 JMPs/JMCs, approximately 80% are empirical papers, with nearly half containing a theoretical model. Moreover, 28% of the candidates are female and 46% are white. Roughly 61% of the JMCs obtained a placement in academia, with a mean ranking of 285.

Comparing the characteristics of the full sample to the 150 candidates with an empirical JMP randomly selected for the p-hacking analysis, as shown in columns (5)–(6), reveals an interesting finding: JMCs with an empirical paper are more likely to be female candidates compared to the total population of candidates. However,

[18]Using broad fields allows us to have a sufficient number of observations in each of these categories.

other characteristics such as race, PhD characteristics, and placement characteristics do not exhibit significant differences across both samples.[19]

From the sub-sample of 150 empirical JMPs, we collect the test statistics that serves as our data for the analysis of p-hacking in the Job Market. Following Brodeur et al. (2020), we collect only coefficients of interest from result tables, excluding constant terms, balance and robustness checks, regression controls, and placebo tests. Importantly, we were blind to any outcomes of interest (placement of the JMC, publication status, etc.) when manually selecting and coding test statistics. Our final sample includes 2,708 test statistics (about 18 test statistics per article). For each test statistic, we record how it is reported (e.g., t-statistic versus coefficient and standard error). We treat coefficient and standard error ratios as if they follow an asymptotically standard normal distribution. When articles report t-statistics or p-values, we transform them into equivalent z-statistics. We also address some complications noted in Brodeur et al. (2016). We re-weight articles based on the number of test statistics per article, and adjust for the rounding of test statistics in the tables.

3 Determinants of Academic Job Market Success

In this section we analyse which characteristics of the JMCs and of their JMPs are the strongest predictors of a successful Job Market, both in terms of type and ranking of the placement.

We first display in Figure 1 the type of placement obtained by the 604 JMCs in our sample. We focus on the first placement after the completion of the PhD, meaning that candidates who get a one year postdoc followed by an AP position are classified as postdocs. The most common placement is an assistant professor position, secured by 36.6% of the candidates. Another 24.5% of the JMCs remains in academia with a postdoc position. Around 39% of the candidates leave academia,

[19]Furthermore, given that our analysis primarily operates at the test statistic level, we include the descriptive statistics of our main variables at this level in Appendix Table A1. We rely on the means reported in this table to discuss the magnitude of our estimates.

with most of them (22.8% of the entire sample) obtaining a job in the private sector. A smaller part of the sample works at governmental institutions (6.5%), central banks (5.3%) or international agencies (4.3%).

In academia, it's often believed that obtaining a PhD from a prestigious institution increases the likelihood of securing a faculty position, especially at a top-ranked school. This is because prestigious programs tend to offer valuable resources, networks, and mentorship opportunities that enhance one's academic profile. To study the relationship between the ranking of the PhD institution and the ranking of the placement institution, we show in Figure 2 the scatter plot of these two variables, conditioning on a candidate obtaining an AP position. The figure shows that only few top departments can place their JM candidates in the top ten economics institutions. Overall, even if we focus on some of the best institutions worldwide, the majority of AP placements from these institutions are outside of the top 100.[20]

To examine more in details the link between JMCs and JMPs characteristics and placement ranking, we employ the following model:

$$\begin{aligned} PlacementRank_{it} = \beta_0 &+ \beta_1 RankPhD_i + \beta_2 Female_i + \beta_3 White_i \\ &+ \beta_4 Sup.Citations_i + \beta_5 Sup.Coauthors_i \\ &+ \beta_6 Theory_i + \beta_7 Theo.\&Emp._i + \beta_8 Authors_i + \gamma_t + \epsilon_{it} \end{aligned} \tag{1}$$

where $PlacementRank_{it}$ is the ranking of the placement obtained by candidate i in year t. We consider all JMCs that obtain an academic placement (postdocs and APs) or a placement at a research institution that is listed in the Ideas/RePEc ranking. $RankPhD_i$ is the ranking of the institution where candidate i has earned the PhD. $Female_i$ is a dummy equal to one if candidate i is female. $White_i$ is a dummy equal to one if candidate i is white. To examine the role of the PhD supervisor's network and productivity, we include the number of citations, $Sup.Citations_i$,

[20] In Appendix Figure A2, we include candidates who secured an AP position after a one-year postdoc are, and the scatter plot exhibits consistent trends. Employing a linear scale for the AP placement rank in Appendix Figure A3, and including the complete sample of academic placement institutions in Appendix Figure A4, suggest a consistent pattern of positive correlation between the rank of the PhD institution and the rank of the academic placement institution.

and the number of coauthors, $Sup.Coauthors_i$, of candidate's i supervisor. If candidate i has more than one supervisor, the mean of the citations and coauthors are used. $Theory_i$ is a dummy that equals one if candidate's i JMP is a theoretical paper. $Theo.\&Emp._i$ is a dummy taking value one if the paper has both theoretical and empirical sections. The reference category is therefore fully empirical JMPs.[21] $Authors_i$ represents the number of authors in candidate's i JMP. γ_t represents year fixed effects and accounts for market year specific characteristics.

To study the link between JMCs characteristics and the probability of obtaining an AP placement, we rely on a similar probit regression:

$$\begin{aligned} Pr(AP_{it} = 1) = \Phi(\ & \beta_0 + \beta_1 RankPhD_i + \beta_2 Female_i + \beta_3 White_i \\ & + \beta_4 Sup.Citations_i + \beta_5 Sup.Coauthors_i \\ & + \beta_6 Theory_i + \beta_7 Theo.\&Emp._i + \beta_8 Authors_i + \gamma_t + \epsilon_{it}) \end{aligned} \tag{2}$$

where AP_{it} is a dummy equal to one if candidate i obtains an AP placement in year t, and zero otherwise. The set of covariates remains the same as in equation 1.

We report the estimates for equation 1 and equation 2 in columns (1)–(3) and (4)–(6) of Table 2, respectively.[22]

Our findings in column (3) suggest that female Job Market Candidates (JMCs) and white JMCs tend to achieve between around 35% and 30% higher placement rankings compared to their male and non-white counterparts, respectively.[23] Turning to the probability of securing an AP position, columns (4)–(6) suggest that while white and female candidates are more likely to achieve higher placement rankings once employed in academia, they face lower initial likelihoods of obtaining academic placements, although the coefficient for women is not statistically significant.[24]

[21] Although some empirical papers incorporate a toy model or brief theoretical framework, we still categorize them as empirical studies. Moreover, controlling for whether a paper includes a toy model or not does not affect our findings.

[22] We restrict the sample to JMCs that obtained an Assistant Professor position in Appendix Table A2.

[23] The average placement ranking in this sample being around 285.

[24] Results on women candidates from our sample are in line with Fortin et al. (2021) who investigate the impact of gender differences in job placement among recent economics PhD candidates and find that women are underrepresented as assistant professors.

Estimates on $Sup.Coauthors_i$ in column (3) provide weak evidence that a larger network is associated with higher placement ranking. Surprisingly, perhaps, the advisor's citation counts and the number of coauthors are not significantly correlated with academic placement in columns (4)–(6). The JMPs characteristics seem to play a role in academic placement. While column (3) indicates that the presence of a theoretical model combined with an empirical analysis is associated with higher placement raking, column (6) suggests that a theoretical papers are likely to increase the likelihood of securing an assistant professor position by around 20-percentage-points. Additionally, a 10% increase in PhD ranking is associated with an approximately 8% increase in placement ranking. Although the PhD ranking is a robust predictor of placement ranking, it does not seem to influence whether a candidate secures an academic placement in columns (4)–(6). The coefficient on $RankPhD_i^2$ fails to provide additional evidence of any nonlinearity between the ranking of the PhD institution and the placement outcomes. Finally, candidates who were on the job market during the COVID-19 pandemic (i.e., year 2020–2021) exhibit around 15-percentage-point decrease of being placed in academia as an assistant professor. We further investigate this tighter market dynamic in a more detailed analysis in Section 5.

To sum up, our descriptive analysis of the placement determinants reveals a positive correlation between the ranking of placement and the supervisor's network. This relationship is robust and not absorbed by the strong influence of the PhD institution rank. Gender and race are also strong predictors of placement quality, with female and white candidates obtaining, ceteris paribus, a significantly better ranked placement. Market conditions (e.g., COVID) are also shown to impact the likelihood of academic placement. We now shift our focus to examining p-hacking and its implications for academic placement outcomes.

4 Statistical Significance and Academic Jobs

In this section, we delineate our approach to identifying and quantifying p-hacking through both graphical and formal analyses. Our methodology involves distinct methods that compare the distribution of test statistics across various subsamples, with supplementary tests detailed in the appendix. While none of these approaches provides indisputable evidence regarding the extent of p-hacking, when considered collectively, we assert that the consistent alignment of results across diverse methodologies should be compelling for the majority of readers.

We briefly note two limitations in our empirical analysis. First, p-hacking has been demonstrated to result in inflated effect sizes (Gelman and Carlin (2014); Ioannidis (2008)), a phenomenon not captured or quantified by our methods. It is also important to highlight our specific focus on marginal p-hacking, where the manipulation revolves around achieving or narrowly missing statistical significance. It is conceivable that non-marginal p-hacking, occurring significantly beyond the unit circle surrounding our statistical significance thresholds, may be taking place without detection through the methods employed in this study.

4.1 Graphical Analysis

Figure 3 presents the distribution of z-statistics in our full sample, weighted by the inverse of the number of z-statistics collected in each paper.[25] The distribution displays a hump shape with z-statistics between 1.8 and 2.2. The distribution exhibits a local minimum around 1.3 and a maximum at around the 10% statistical threshold, suggesting misallocated z-statistics. About 65, 57, and 42 percent of test statistics are significant at the 10, 5, and 1 percent levels, respectively. These figures are larger that what has been reported in Brodeur et al. (2020) who documented that the share of respectively significant papers were about 56, 48, and 34 percent in 25 top economics journals; which potentially highlights the important role that

[25] Appendix Figure A5 represents the raw distribution of z-statistics in our sample, while in Appendix Figure A6 we adjust for the rounding of test statistics. In both cases the shape of the distribution remains similar to the weighted distribution.

statistical significance plays in signaling potential on the job market.[26]

The hump around the statistical threshold suggests that PhD candidates may be engaging in questionable research practices. The most common form of this is p-hacking, where JMCs manipulate their results to achieve statistical significance. P-hacking typically involves selectively reporting only significant results or adjusting models to produce significance. However, it is not solely limited to such practices. It may also involve the "desk-drawer effect," where candidates abandon projects with insignificant results. Additionally, the hump may be influenced by selection bias, where candidates opt not to enter the job market if their results are insignificant, instead postponing for another year to improve their results or pursue ideas with more significant findings. While it is difficult to disentangle which factor is more dominant, the clear conclusion is that research integrity is being compromised. With some caution, we refer to this phenomenon as p-hacking throughout the remainder of this analysis.

To visualize the relationship between significant estimates and academic placement, we compare the weighted distribution of test statistics across four categories of placement: non-academic placement (panel A), postdoctoral placement (panel B), academic placement (panel C), and placement as assistant professor (panel D) in Figure 4.[27] First, statistical significance appear to matter for academic placement as density of the test statistic is higher around significance thresholds. The larger hump around the 10% significance threshold in AP also suggests that statistical significance could potentially be a predictor of faculty placement. For instance, 75% of test statistics are marginally significant at the 10% level (i.e., significant in 1.64 $\pm$

[26] We formally test for p-hacking across our subsamples in Section 4.2. In Appendix Figure A7, we rely on tests developed by Elliott et al. (2022) against a null hypothesis of no p-hacking. We use these tests to test for the presence of p-hacking in our main sample. We find that about half of these tests reject their null hypothesis. We do not rely on these tests in our main analysis as they are not designed to compare subsamples nor allow to control for covariates. Some of the tests, as operationalized by Elliott et al. (2022), are underpowered as they rely on very tight windows resulting in a small sample size.

[27] The same graphs are replicated for the unweighted distributions in Appendix Figure A8. We also replicate the graph showing the weighted distribution of test statistics for placement as assistant professor of Panel D and include candidates that obtained an AP position after a one year of postdoc in Appendix Figure A9.

0.5) for the sample of AP hires compared to 67% for the sample of academic hires, and 59% for the sample of non academics.

Appendix Table A3 confirms clear discontinuities around the 10% statistical threshold. We use binomial proportions test to compare the differences in test statistics mass just above and below conventional statistical thresholds by candidates' placement. This method operates on minimal assumptions: (1) the probability of being marginally above versus below any threshold should be equal, and (2) the likelihoods of falling marginally above or below the significance threshold should be comparable between test statistics of candidates placed in academia and those who are not. In column (1), we observe that the proportion of marginally significant results at the 10% level is consistently higher than the non-significant ones, suggesting potential manipulation in job market papers. In columns (4) and (7), we find a significant positive association between academic placement, especially AP, and the proportion of marginally significant test statistics at the 10% significance level. However, this relationship is not evident at the 5% level.

The significantly larger discontinuity in academic placement, even with a smaller sample size, supports earlier visual inspections and suggests that academic placement suffers from recruitment bias where candidates with significant results are more likely to land a job in academia.[28] Employers' preference for statistically significant results creates incentives for junior researchers to engage in questionable research practices and may ultimately result in the selection of candidates with lower scientific integrity.

Lastly, we examine heterogeneity across academic institution rankings. In Appendix Figure A11, we compare the weighted distribution of test statistics between above and below median placement in terms of institution ranking in Panel A and B, respectively. The larger hump around the conventionally significance thresh-

[28]It is empirically challenging to test at which stage statistical significance has its influence on hiring, primarily due to the confidential nature of the hiring process. In an exploratory analysis, we restrict our sample to JMPs accessed before the initial interview period (i.e., December of the academic year). The distributions in Appendix Figure A10 support our argument on the presence of recruitment bias and suggest that statistical significance may influence candidate selection in the early stages of the academic job market.

old in Panel A compared to Panel B suggests that p-hacking is potentially more "beneficial" in lower ranked institutions (i.e., above median). This finding suggests that the larger hump observed in academic placement is unlikely to be attributed to a compositional effect.[29] We also compare the distributions of test statistics by different PhD institution rankings in Appendix Figure A12. The heightened density around conventional statistical significance levels suggests a higher likelihood of p-hacking behavior in lower-ranked PhD institutions.

4.2 Caliper Tests

To further investigate the relationship between p-hacking and academic placement, we exploit caliper tests (Gerber and Malhotra, 2008).[30] Caliper tests rely on test statistics within a narrow range around the most commonly used significance thresholds to define marginal significance. If marginally insignificant results are as valuable as marginally significant ones, we would expect to find no differences in the likelihood of employment in a specific category. For example, if significance plays no role in placement for assistant professors, we should see no differences in likelihood of AP placement between candidates with marginally significant tests in their JMP and candidates with marginally insignificant ones.

It is important to stress that the results of the caliper tests do not measure the causal effect of p-hacking on the outcome considered. First, they jointly capture the effect of p-hacking and recruiting bias. Here recruiting bias refers to the fact that recruiters might prefer, ceteris paribus, candidates with marginally significant results in the JMP to candidates with marginally insignificant results. Second, JMPs might vary along other dimensions than p-hacking. We thus rely on caliper tests because we can control for observable characteristics of the JMPs and JMCs.

[29] The positive relationship between p-hacking and academic placement could be driven by a compositional effect if, for instance, top candidates (i.e., candidates from prestigious institutions) are more more prone to having marginally significant results and, simultaneously, more likely to secure academic positions. Graphical evidence rules out this explanation. Furthermore, we include institution fixed effects in our formal analysis to mitigate this potentially confounding factor.

[30] A distinctive characteristic of the caliper tests that we implement in this paper is that the dummy indicating the significance of a test is the independent variable in our models, instead of the dependent one.

To partially address potential endogeneity issues in our model, we examine the correlations between p-hacking and observable characteristics, including those related to the JMP, JMC, PhD institution, and PhD advisor, as shown in Appendix Table A4. Our estimates indicate that none of these observables strongly predict p-hacking. Notably, the presence of a theoretical model is the only factor consistently associated with a reduced likelihood of p-hacking across all significance levels (Brodeur et al. (2016)).[31] Although the absence of strong correlations does not allow us to assert that p-hacking is exogenous, it strengthens the credibility of our caliper test specification.

Consistently with our PAP, we focus on the 5% and 10% significance thresholds, and show estimates for the 1% threshold in the appendix. For the 10% threshold we restrict the sample to the following interval:

$$R_{-,h} = [1.65 - h, 1.65]; R_{+,h} = [1.65, 1.65 + h] \tag{3}$$

Where h represents the parameter that we set to change the width of the interval considered.

We estimate the following (probit) regression:

$$Pr(Y_j = 1) = \Phi(\alpha + \gamma_t + \tau_u + X_j'\beta + \delta Sig_i) \tag{4}$$

Y_j is a dummy equal to one if a certain outcome of interest occurred for JMC j (for example, if candidate j obtained an academic placement). X_j denotes a set of covariates including dummy variables for the method used in the JMP of candidate j (e.g., DiD or IV), dummy variables for the field of the paper, the number of authors of the paper, a dummy equal to one if candidate j is female, a dummy equal to one if candidate j is white, the number of citations and the number of coauthors of the candidate's supervisor. If a candidate has more than one supervisor, the mean of the citations and coauthors are used. Sig_i is a dummy taking value one if test

[31] Interestingly, female candidates are less likely, while white candidates are more likely, to report marginally significant results at the 5% level in their JMPs.

i is statistically significant at the 10% level (or 5% and 1%, when these different threshold are considered). γ_t denotes year fixed effects. τ_u indicates institution fixed effects. To prevent tables with numerous test statistics from being overweighted, we weight observations using the inverse of the number of tests presented in the same article. Standard errors are clustered at the article level.

This specification allows us to examine the relationship between market outcomes and discontinuities in the probability of a test statistic appearing just above or below a conventional statistical threshold tests statistics. We thus report the difference in local mean (average) market outcomes between test statistics assigned to the "treatment" (i.e., marginally significant) and those assigned to the control. Since this discontinuity reflects bunching near statistical thresholds, "treatment" in this context refers to p-hacking.

The results are presented in Table 3 for academic placements (Panel A) and assistant professor positions (Panel B) at the 10% threshold. Columns (1) to (4) examine a window of half-width $h = 0.5$ around the absolute value of the one-star significance threshold (i.e., $\mid t \mid = 1.65$). Our preferred specification in column (4) includes the full set of controls and fixed effects. Coefficients indicate that marginally significant tests are associated with a significant increase in the likelihood of placement in academia and as an assistant professor by approximately 10 and 14 percentage points, respectively. Our estimates are also economically significant as this effect represents approximately 20% of the likelihood of academic and 50% of assistant professorship placement. In columns (5) and (6), we show that our caliper findings for the 1.65 cutoff are robust to alternative windows: 1.96 ± 0.35 and $\pm$ 0.20. The estimates in column (6) remain of similar magnitude but are statistically insignificant.[32]

We replicate the analysis at the 5% and 1% significance levels in Table 4 and Appendix Table A8, respectively. While results at the 1% significance level follow the same pattern, estimates at the 5% are close to zero and not statistically signif-

[32]We present the unweighted results in Appendix Table A5. In Appendix Tables A6 and A7, we adjust for the rounding of test statistics.

icant suggesting that marginally significant estimates at the 5% threshold are not associated with higher probability of getting an academic or AP placement. These findings are in line with the suggestive evidence presented in Figure 4 indicating that the samples of academic placement and AP do not exhibit a higher density in the distribution of test statistics for the 5% threshold.

In Appendix Table A9, we assess the robustness of our findings by incorporating one-year postdocs as APs in Panel A, and our results remain consistent. Recognizing potential differences in characteristics between academia and non-academic placements, we narrow the sample to candidates who exclusively remained in academia. In Panel B, we observe a significant positive relationship between marginal significance and the likelihood of securing a placement as an AP. Lastly, in Panel C, we find weak evidence that marginal significance is correlated with higher placement as measured by the ranking of the academic institution.

To sum up, our formal investigation using a caliper test aligns with our graphical analysis, showing that p-hacking is rewarded in academia and suggesting the presence of recruitment bias.

5 Insights from COVID-19 Disruption

The current association between statistical significance and the likelihood of academic placement aligns with (at least) two hypotheses: (1) PhD candidates with significant results may prefer academic jobs, perceiving such findings as a signal enhancing their job market potential, and (2) recruiting bias.

To get insights on whether the positive association between statistical significance and academic placement is driven by a recruiter bias in academia, we leverage the COVID-19 pandemic as a disruption to the equilibrium of the academic labor market.

We rely on data from AEA report Cawley (2023) to illustrate the economics PhD market during our analysis period in Appendix Figure A13. We rely on the number of full-time academic positions listed on Job Openings for Economists (JOE, AEA) and the number of PhD students who applied to at least one job through JOE as proxies for the demand and supply for new PhD economists, respectively.

Trends suggest that the supply of new PhD economists remained relatively constant during the pandemic (i.e., year 2020).[33] On the demand side, job openings for full-time academic positions decreased by 42.9% compared to its value in 2019. We thus rely on the COVID-19 pandemic as an exogenous shock disrupting the labor market equilibrium through a decrease in the demand for academic positions (including AP) and an increase in competitiveness.

The increased competition for available positions is potentially characterized with academic employers becoming more selective in their hiring processes. If employers rely on statistical significance as a signal for potential (i.e., recruitment bias), then the relationship between statistically significant and the likelihood of academic employment would be stronger during the academic year 2020-2021.[34]

We start by comparing the distribution of test statistics before and after the pandemic respectively in Panels A and B of Figure 6. The hump around the 10% significance threshold is slightly larger pre-pandemic. However, the density in the distribution of test statistics is higher for lower significance levels. There is thus no strong evidence that PhD candidates are altering their p-hacking behavior in response to a sudden increase in competitiveness. One potential explanation is that students did not anticipate the large negative impact that Covid can have on demand of academic positions.

In Table 5, we introduce an interaction term between our dummy variable for marginal significance and *Covid*, a dummy variable that takes the value of one during the academic year 2020–2021 and zero otherwise.[35]

[33]This finding is consistent with data on openings from EJM and the AEA ad hoc Committee report on the economists job market. It is also consistent with data from our full sample on 12 PhD institutions.

[34]The supply and demand dynamics of new PhD economists during this academic year are depicted by the statistics from the year 2020 in Appendix Figure A13, given that the majority of postings and applications occur toward the last trimester of 2020.

[35]Although Covid began spreading in March 2020, its impact on the demand for academic positions was primarily felt during the 2020-2021 academic year. Therefore, we include the academic year 2019-2020 in our control group. Repeating the analysis without the 2019-2020 sample does not affect our results.

The results in Panel A suggest that p-hacking led to better placement during the Covid period.[36] Estimates on the interaction term indicates that during the Covid period, marginally significant tests were associated with approximately a 15-percentage-point higher likelihood of academic employment compared to the pre-Covid period. Results in Panel B follow a similar pattern but are not statistically significant.[37]

If employers are more selective when their constraint to recruitment is higher, these findings suggest that they are more likely to rely on statistical significance as signal for potential. Thus, the positive relationship between p-hacking and academic employment appear to be (at least partly) driven by academic employers' preference for marginally significant results. These results align with DellaVigna and Linos (2022)'s findings, which demonstrate that exaggerated effect sizes and publication bias are more common in academic journals compared to those conducted by major Nudge Units within the government.

6 Publication Bias

We now turn our attention to publication bias in our sample. Publication bias reflects a potential preference among editors and reviewers for results that display statistical significance (Havránek (2015); Stanley et al. (2024)). In this context, it serves to highlight a potential mechanism whereby recruiters rely on the statistical significance of estimates in JMPs as an indicator of future publication and, consequently, candidates' qualifications. We thus investigate the degree of publication bias by documenting the relationship between marginal significance in job market papers (JMPs) and subsequent publication.

For this analysis, we search all 150 researchers' CVs and webpages for information on whether their JMP is under revision (R&R) or published as of May, 2024.

[36] In Appendix Table A10, we adjust for the rounding of test statistics.

[37] We also examine the impact of COVID on the relationship between p-hacking and likelihood of AP placement in Appendix Table A11. Estimates suggest weak evidence that p-hacking at the 10% level was more rewarding in COVID for AP jobs.

At this date, 20% of JMPs are published while around 11% are under revision.[38]

Figure 5 illustrates the distribution of z-statistics based on the publication status of the JMP. The top panel displays the distribution for JMPs that are neither published nor R&R, the central panel for published JMPs, and the bottom panel includes both published papers and those under R&R. The hump in the sample of published JMPs (panel B) appear to be more pronounced than those of unpublished JMPs (panel A). When papers under revision are included in Panel C, the hump diminishes, but this is primarily due to a higher density of the distribution for higher levels of significance.[39]

The employment status in academia could potentially act as a confounding factor in this context, particularly because individuals who do not secure an academic appointment may not be as motivated to continue working on their job market paper and pursue publication. To address this issue, we narrow down our sample to include only candidates who received an academic placement, enabling us to conduct a more focused graphical analysis. Appendix Figure A14 shows that the distribution of test statistics remains consistent.

Next, we undertake a more formal investigation to determine whether the presence of marginally significant results influence the likelihood of publication. As discussed earlier, the potential for publication could serve as a key factor in the connection between statistical significance and p-hacking. A positive relationship would suggest that p-hacking contributes to publication success, which could potentially be a driver of the effect observed on academic employment.

Appendix Table A12 presents our results from estimating Equation 4 on the likelihood of publication, focusing on the 10% and 5% significance thresholds in Panels A and B, respectively. Consistent with our main findings, the estimates suggest that

[38]Of note, JMPs go through changes after the market. In this analysis, we are documenting the relationship between statistical significance in JMPs and future publication using the version of the JMP at the time of the market rather than the submitted version of the paper.

[39]We replicate this graph using the unweighted sample in Appendix Figure A15. We also conduct an analysis excluding the most recent year in our sample (i.e., 2020–2021) in Appendix Figure A16 to ensure the robustness of our results. This allows us to confirm that excluding the most recent JMPs, which may require more time to be published, does not affect our findings.

test statistics that are marginally significant at the 10% level are significantly associated with around a 10-percentage-points higher likelihood of publication, equivalent to around 35% increase in the baseline likelihood. Results for the 5% threshold follow the same direction, although with a smaller effect size and lacking statistical significance.[40]

Given that statistical significance can have both direct and indirect effects on publication (through academic placement), we add to our specification a dummy variable indicating whether the candidate obtained an academic placement or not. The results are presented in Appendix Table A14. Both academic placement and marginal significance are positively and significantly associated with higher publication rates, and neither effect is absorbed by the other.

The evidence of publication bias in JMPs supports the claim that recruiters may use the statistical significance of estimates in JMPs as a gauge for future publication and, consequently, candidates' qualifications, leading to recruitment bias. It also shows that the null results penalty imposed by economic journals begin at an early stage of a researcher's career, and wrong incentives from publications are going all the way down to PhD candidates (Chopra et al. (2024)).

7 Conclusion

Aspiring economists often face stiff competition when aiming for faculty positions or research roles at universities and research centers. This competitive environment was exacerbated during the COVID-19 pandemic.

In our study, we explore what influences success in the academic job market by examining 604 PhD graduates seeking academic or research-focused positions from 12 universities. We first document a positive relationship between the ranking of the PhD institution and the placement ranking of JMCs. We then turn to demographic characteristics and document that female and white candidates obtain a significantly better ranked placement.

[40]We replicate this table by adding papers under revision as published papers in Appendix Table A13. Estimates are smaller in magnitude and most lose statistical significance.

We then shift our focus to examining p-hacking and its implications for academic placement outcomes. We find that marginally significant tests are associated with a significant increase in the likelihood of placement in academia and as an assistant professor. Furthermore, we provide evidence that marginally significant test statistics are significantly associated with a higher likelihood of publication.

We rely on the COVID-19 pandemic as an exogenous shock to demand. This allows us to document the role of market tightness in explaining p-hacking and ultimately publication bias in academia. The academic job market was characterized by a relatively small number of job vacancies. In this context, we find that JMPs did not suffer from additional p-hacking, but that job market candidates with relatively more p-hacked JMPs were more likely to obtain an academic position. This result suggests that researchers more inclined to selective reporting are more likely to stay in academia, especially when the academic market is slack. These results align with DellaVigna and Linos (2022)'s findings, which demonstrate that exaggerated effect sizes and publication bias are more common in academic journals compared to those conducted by major Nudge Units within the government.

References

Abadie, A.: 2020, Statistical Nonsignificance in Empirical Economics, *American Economic Review: Insights* **2**(2), 193–208.

Andrews, I. and Kasy, M.: 2019, Identification of and Correction for Publication Bias, *American Economic Review* **109**(8), 2766–94.

Ankel-Peters, J., Fiala, N. and Neubauer, F.: 2023, Do Economists Replicate?, *Journal of Economic Behavior & Organization* **212**, 219–232.

Askarov, Z., Doucouliagos, A., Doucouliagos, H. and Stanley, T.: 2023, The Significance of Data-Sharing Policy, *Journal of the European Economic Association* **21**(3), 1191–1226.

Brodeur, A., Carrell, S., Figlio, D. and Lusher, L.: 2023, Unpacking p-Hacking and Publication Bias, *American Economic Review* **113**(11), 2974–3002.

Brodeur, A., Cook, N. and Heyes, A.: 2020, Methods Matter: P-Hacking and Publication Bias in Causal Analysis in Economics, *American Economic Review* **110**(11), 3634–60.

Brodeur, A., Lé, M., Sangnier, M. and Zylberberg, Y.: 2016, Star Wars: The Empirics Strike Back, *American Economic Journal: Applied Economics* **8**(1), 1–32.

Camerer, C. F., Dreber, A., Forsell, E., Ho, T.-H., Huber, J., Johannesson, M., Kirchler, M., Almenberg, J., Altmejd, A., Chan, T. et al.: 2016, Evaluating Replicability of Laboratory Experiments in Economics, *Science* **351**(6280), 1433–1436.

Camerer, C. F., Dreber, A. and Johannesson, M.: 2019, Replication and Other Practices for Improving Scientific Quality in Experimental Economics, *Handbook of Research Methods and Applications in Experimental Economics*, Edward Elgar Publishing, pp. 83–102.

Cawley, J.: 2023, Report of the Committee on the Job Market, *AEA Papers and Proceedings* **113**, 789–91.

Chen, J., Liu, Q. and Billger, S.: 2013, Where Do New Ph.D. Economists Go? Recent Evidence from Initial Labor Market, *Journal of Labor Research* **34**, 312–338.

Chopra, F., Haaland, I., Roth, C. and Stegmann, A.: 2024, The Null Result Penalty, *Economic Journal* **134**(657), 193–219.

Christensen, G. and Miguel, E.: 2018, Transparency, Reproducibility, and the Credibility of Economics Research, *Journal of Economic Literature* **56**(3), 920–80.

Coles, P., Cawley, J., Levine, P. B., Niederle, M., Roth, A. E. and Siegfried, J. J.: 2010, The Job Market for New Economists: A Market Design Perspective, *Journal of Economic Perspectives* **24**(4), 187–206.

Conley, J. P. and Önder, A. S.: 2014, The Research Productivity of New PhDs in Economics: The Surprisingly High Non-Success of the Successful, *Journal of Economic Perspectives* **28**(3), 205–216.

Conley, J. P., Önder, A. S. and Torgler, B.: 2016, Are All Economics Graduate Cohorts Created Equal? Gender, Job Openings, and Research Productivity, *Scientometrics* **108**, 937–958.

Conti, A. and Visentin, F.: 2015, A Revealed Preference Analysis of PhD Students' Choices Over Employment Outcomes, *Research Policy* **44**(10), 1931–1947.

DellaVigna, S. and Linos, E.: 2022, RCTs to Scale: Comprehensive Evidence from Two Nudge Units, *Econometrica* **90**(1), 81–116.

Doucouliagos, C. and Stanley, T. D.: 2013, Are All Economic Facts Greatly Exaggerated? Theory Competition and Selectivity, *Journal of Economic Surveys* **27**(2), 316–339.

Drazen, A., Dreber, A., Ozbay, E. Y. and Snowberg, E.: 2021, Journal-Based Replication of Experiments: An Application to "Being Chosen to Lead", *Journal of Public Economics* **202**, 104482.

Eberhardt, M., Facchini, G. and Rueda, V.: 2023, Gender Differences in Reference Letters: Evidence From the Economics Job Market, *Economic Journal* **133**(655), 2676–2708.

Elliott, G., Kudrin, N. and Wüthrich, K.: 2022, Detecting p-Hacking, *Econometrica* **90**(2), 887–906.

Fortin, N., Lemieux, T. and Rehavi, M.: 2021, Gender Differences in Fields of Specialization and Placement Outcomes Among PHDs in Economics, *AEA Papers and Proceedings*, Vol. 111, pp. 74–79.

Furukawa, C.: 2024, Publication Bias under Aggregation Frictions: From Communication Model to New Correction Method. Yokohama National University Mimeo.

García-Suaza, A., Otero, J. and Winkelmann, R.: 2020, Predicting Early Career Productivity of PhD Economists: Does Advisor-Match Matter?, *Scientometrics* **122**(1), 429–449.

Ge, Q. and Wu, S.: forthcoming, How Do You Say Your Name? Difficult-To-Pronounce Names and Labor Market Outcomes, *American Economic Journal: Economic Policy* .

Gelman, A. and Carlin, J.: 2014, Beyond Power Calculations: Assessing Type S (Sign) and Type M (Magnitude) Errors, *Perspectives on Psychological Science* **9**(6), 641–651.

Gerber, A. and Malhotra, N.: 2008, Do Statistical Reporting Standards Affect what is Published? Publication Bias in Two Leading Political Science Journals, *Quarterly Journal of Political Science* **3**(3), 313–326.

Grogger, J. and Hanson, G. H.: 2015, Attracting Talent: Location Choices of Foreign-Born PhDs in the United States, *Journal of Labor Economics* **33**(S1), S5–S38.

Hadlock, C. J. and Pierce, J. R.: 2021, Hiring your Friends: Evidence from the Market for Financial Economists, *ILR Review* **74**(4), 977–1007.

Havránek, T.: 2015, Measuring Intertemporal Substitution: The Importance of Method Choices and Selective Reporting, *Journal of the European Economic Association* **13**(6), 1180–1204.

Havránek, T., Irsova, Z., Laslopova, L. and Zeynalova, O.: Forthcoming, Publication and Attenuation Biases in Measuring Skill Substitution, *Review of Economics and Statistics* .

Hilmer, M. and Hilmer, C. E.: 2012, On the Search for Talent in Academic Labor Markets: Easily Observable Later-Graduate Study Outcomes as Predictors of Early-Career Publishing, Placement, and Tenure, *Economic Inquiry* **50**(1), 232–247.

Ioannidis, J. P.: 2008, Why Most Discovered True Associations Are Inflated, *Epidemiology* pp. 640–648.

Jones, T. and Sloan, A.: Forthcoming, The Academic Origins of Economics Faculty, *Journal of Economic Education* .

Kepes, S., Keener, S. K., McDaniel, M. A. and Hartman, N. S.: 2022, Questionable Research Practices Among Researchers in the Most Research-Productive Management Programs, *Journal of Organizational Behavior* **43**(7), 1190–1208.

McFall, B. H., Parolin, E. D. and Zafar, B.: 2024, Career expectations and outcomes: Evidence (on gender gaps) from the economics job market, *Technical report*, National Bureau of Economic Research.

Miguel, E.: 2021, Evidence on Research Transparency in Economics, *Journal of Economic Perspectives* **35**(3), 193–214.

Olsen, J., Mosen, J., Voracek, M. and Kirchler, E.: 2019, Research Practices and Statistical Reporting Quality in 250 Economic Psychology Master's Theses: A Meta-Research Investigation, *Royal Society Open Science* **6**(12), 190738.

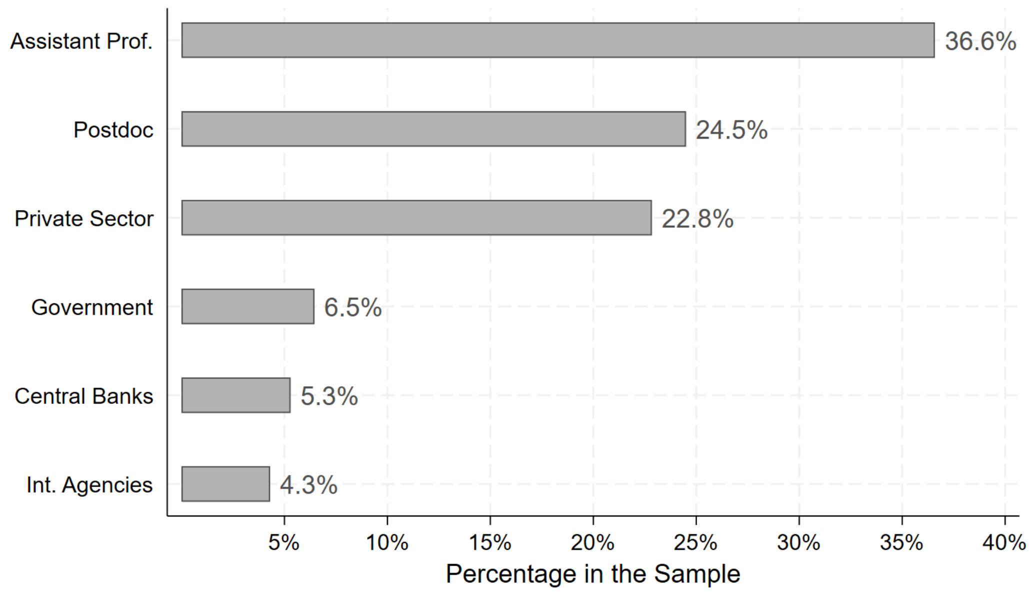
Oyer, P.: 2006, Initial Labor Market Conditions and Long-Term Outcomes for Economists, *Journal of Economic Perspectives* **20**(3), 143–160.

O'Boyle Jr, E. H., Banks, G. C. and Gonzalez-Mulé, E.: 2017, The Chrysalis Effect: How Ugly Initial Results Metamorphosize Into Beautiful Articles, *Journal of Management* **43**(2), 376–399.

Stanley, T. D., Doucouliagos, H. and Havránek, T.: 2024, Meta-Analyses of Partial Correlations Are Biased: Detection and Solutions, *Research Synthesis Methods* **15**(2), 313–325.

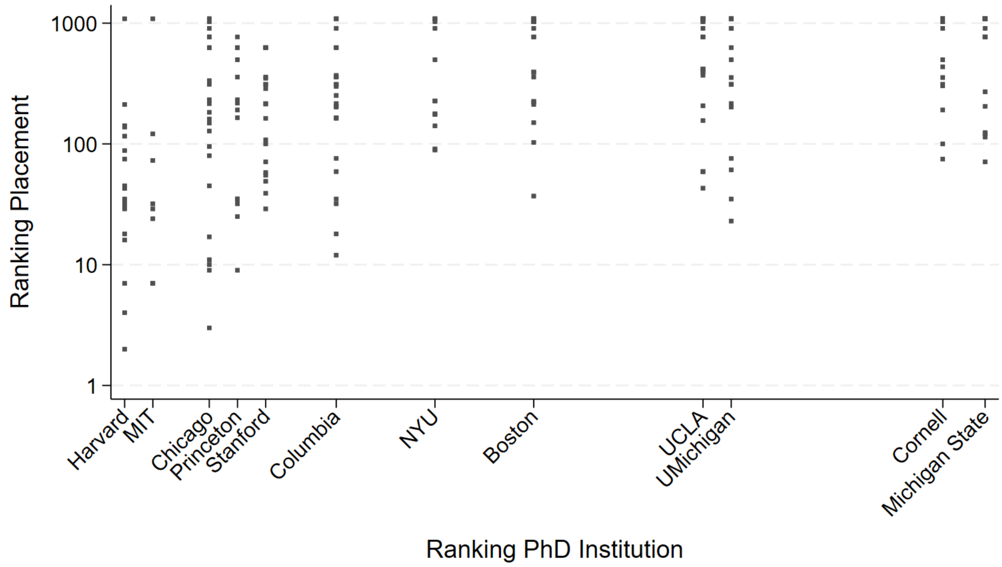
8 Figures

Figure 1: Placement of Job Market Candidates



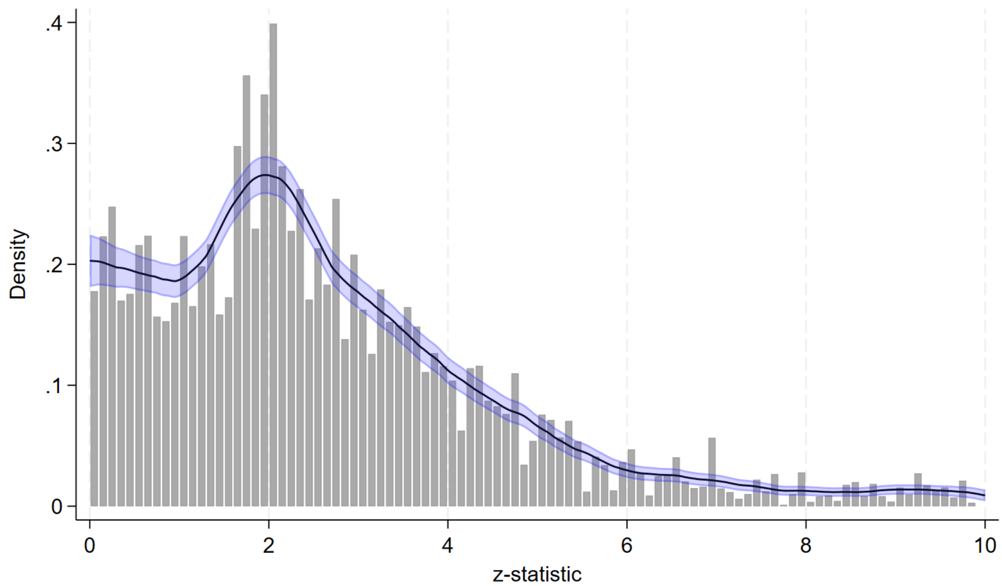
Notes: This figure shows the initial placement obtained by 604 job market candidates from 12 universities ranked among the top 100 economic institutions based on the IDEAS/RePEc classification (August 2023), pooled for three academic years: 2018–2019, 2019–2020, and 2020–2021.

Figure 2: Assistant Professor Placement Ranking



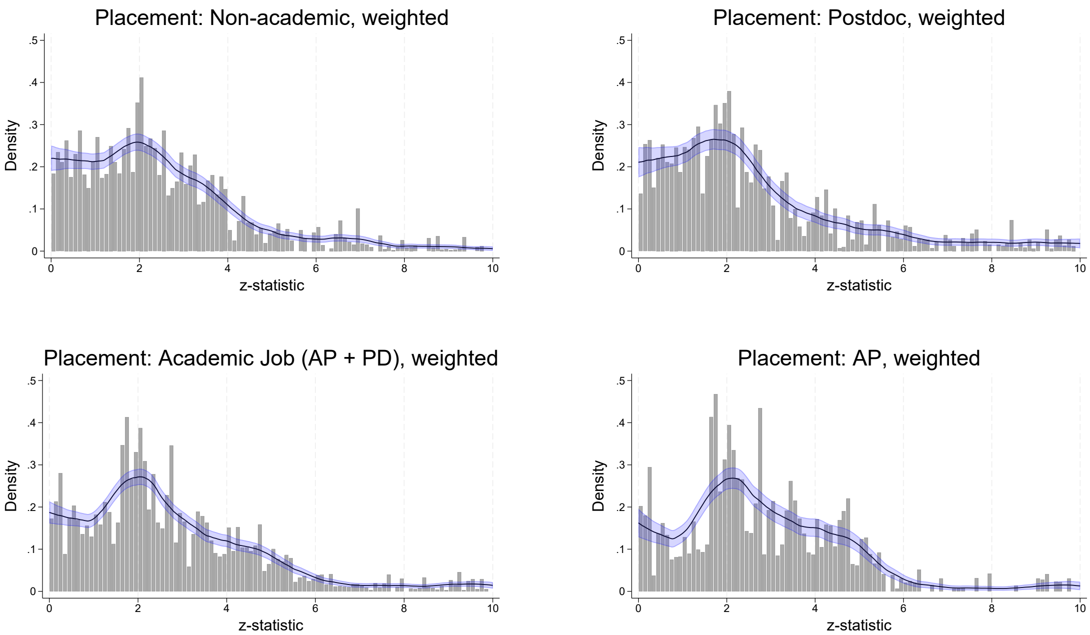
Notes: This figure displays the scatter plot of placement ranking and PhD institution ranking for 221 job market candidates who obtained an assistant professor position after graduation. We include job market candidates from 12 universities ranked among the top 100 economic institutions based on the IDEAS/RePEc classification (August 2023), pooled for three academic years: 2018–2019, 2019–2020, and 2020–2021. Universities without a ranking are assigned the lowest ranking (1088). We employ a logarithmic scale for the y axis.

Figure 3: Distribution of Tests Statistics



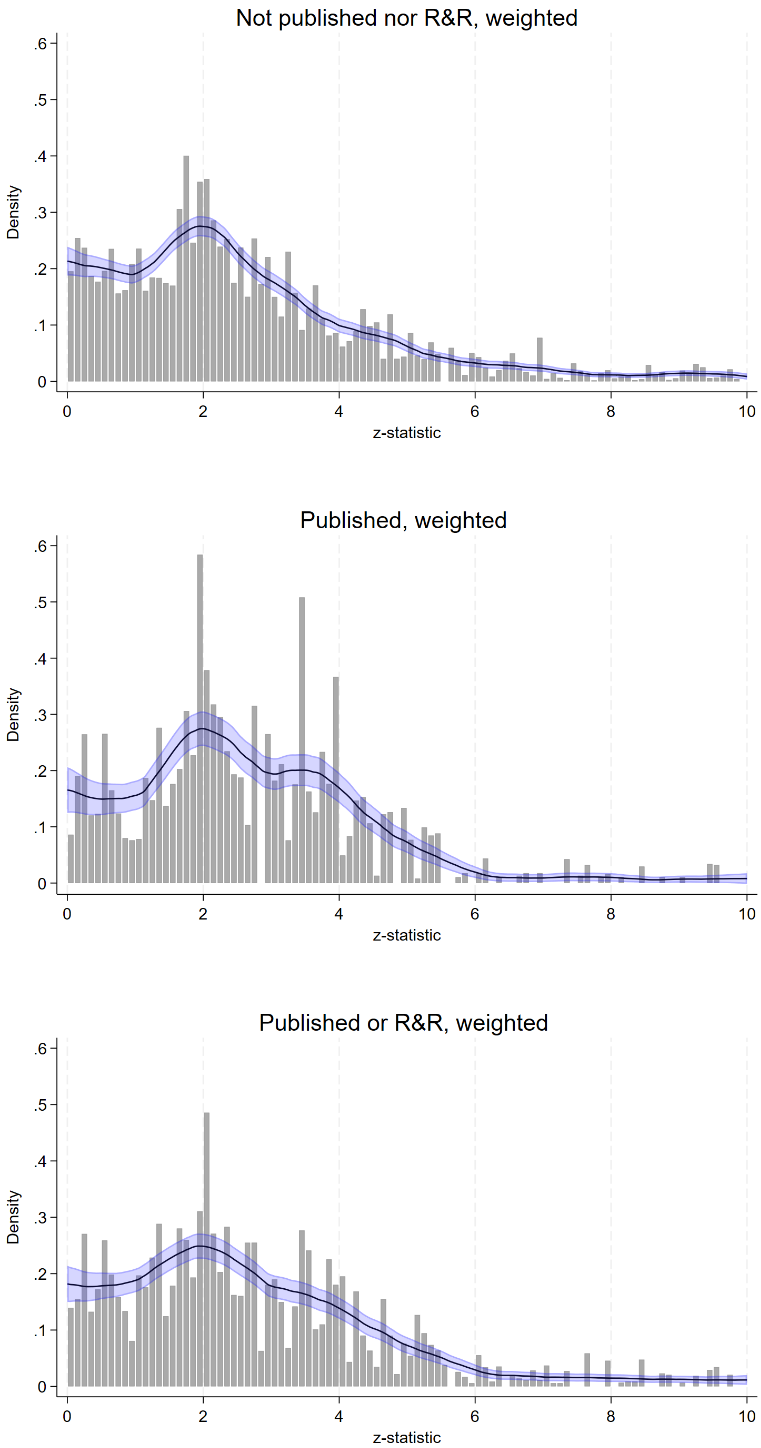
Notes: This figure shows the distribution of 2,708 test statistics for $z \in [0, 10]$ from the 150 empirical job market papers considered in our sample. Bins are 0.1 wide and we superimpose an Epanechnikov kernel. Observations are weighted by the inverse of the number of tests in the paper.

Figure 4: Test Statistics Distribution by Placement



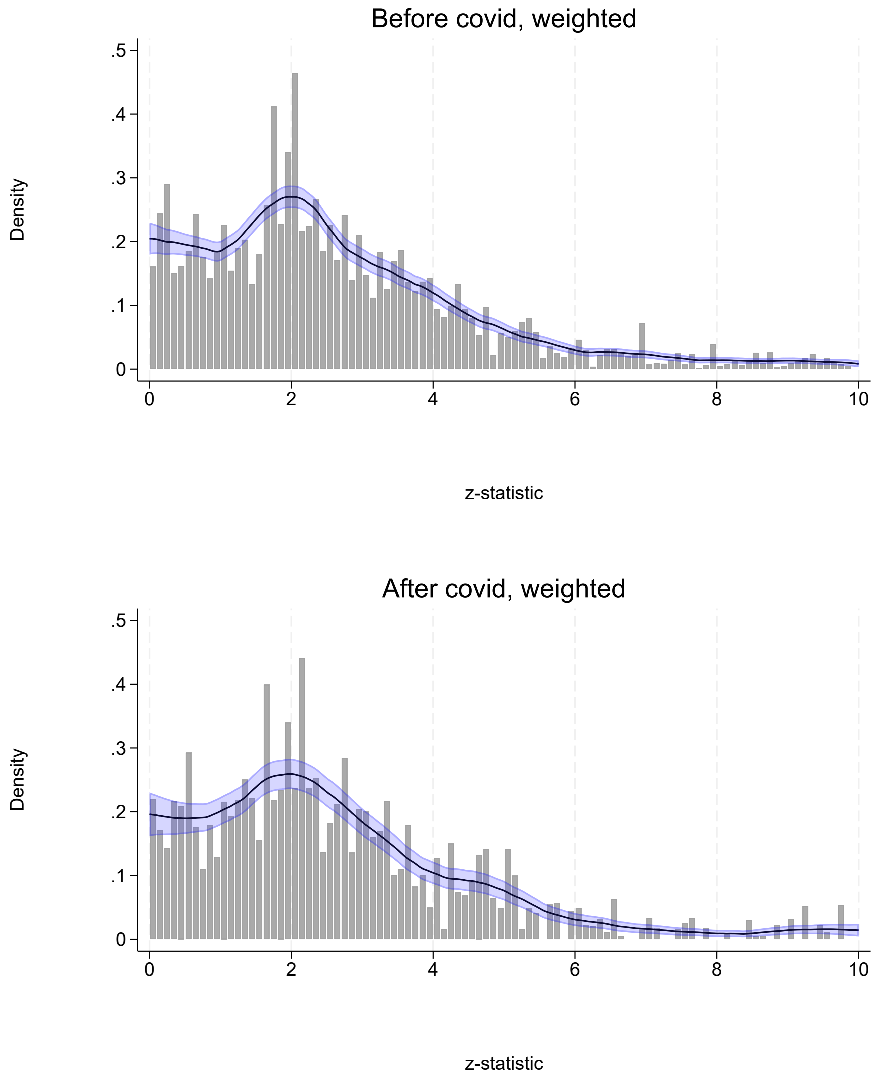
Notes: This figure shows the distribution of test statistics for $z \in [0, 10]$ from 150 job market papers by placement of the job market candidates. The top left panel displays the t-statistics in the job market papers that resulted in a non-academic placement (private sector, government, central banks or international agencies). The top right panel shows the t-statistics for candidates that obtained a postdoc. The bottom left panel shows the distribution of JMPs that resulted in an academic placement (assistant professor or postdoc). The bottom right panel focuses on assistant professor placements only. Candidates obtaining a 1 year postdoc followed by an AP position are classified as postdocs. Bins are 0.1 wide and we superimpose an Epanechnikov kernel. Observations are weighted by the inverse of the number of tests in the paper.

Figure 5: Test Statistics Distribution by Publication Outcome



Notes: This figure shows the distribution of test statistics for $z \in [0, 10]$ from 150 job market papers by publication outcome. The top panel shows the distribution of test statistics for articles that are not published nor R&R. The middle panel displays the test statistics for published manuscripts. The bottom panel shows papers that are published or under revision (R&R). Bins are 0.1 wide and we superimpose an Epanechnikov kernel. Observations are weighted by the inverse of the number of tests in the paper.

Figure 6: Test Statistic Distribution by Job Market Year



Notes: This figure shows the distribution of test statistics for $z \in [0, 10]$ from 150 job market papers by job market (JM) year. The figure in the top panel shows z-statistics for the papers in the job markets that occurred before covid (2018 and 2019 JMs). The figure in the bottom panel displays z-statistics from the job market that occurred after covid (2020). Bins are 0.1 wide and we superimpose an Epanechnikov kernel. Observations are weighted by the inverse of the number of tests in the paper.

9 Tables

Table 1: Summary Statistics – Full and Selected Samples

	P-Hacking Sample		Full Sample		Difference	P-Value
	Mean	SD	Mean	SD		
N. Authors JMP	1.208	0.497	1.194	0.512	0.014	0.763
Female	0.420	0.495	0.283	0.451	0.137***	0.002
White	0.433	0.497	0.459	0.499	-0.025	0.578
Theory Paper	0.000	0.000	0.204	0.403	-0.204***	0.000
Theory & Empirical	0.067	0.250	0.411	0.492	-0.344***	0.000
Rank PhD	25.233	20.261	24.677	19.949	0.556	0.763
Placement Ranking	280.347	350.582	285.251	360.663	-4.904	0.900
Supervisor Citations	27215.174	34157.850	27355.739	40715.751	-140.565	0.966
Supervisor Coauthors	59.989	69.287	48.782	58.630	11.207*	0.070
Academic Placement	0.540	0.500	0.611	0.488	-0.071	0.120
AP Placement	0.280	0.451	0.366	0.482	-0.086**	0.040
Observations	150		604			

Notes: This table provides a comparison between the characteristics of the selected 150 job market candidates (JMC) for the p-hacking analysis and the broader sample (604 JMCs) included in the study of the determinants of academic job market success. *Theory Paper* is a dummy equal to one if the job market paper is fully theoretical. *Theory & Emp. Paper* is a dummy taking value one if the JMP has both theoretical and empirical sections. *Rank PhD Institution* is the ranking of the institution at which the JMC has earned the PhD. The variables *Supervisor Coauthors* and *Supervisor Citations* are determined by averaging the number of coauthors and the number of citations of the candidate's supervisors, should there be more than one. *Academic Placement* is a dummy equal to one if the candidate obtains an academic placement. *AP placement* is a dummy taking value one if the candidate secures an assistant professor position.

Table 2: Determinants of Academic Placement

	(1)	(2)	(3)	(4)	(5)	(6)
		Placement Rank			Placement AP	
Year 2019-2020	-29.561	-26.706	-28.760	-0.078*	-0.073	-0.073
	(41.265)	(39.095)	(39.129)	(0.047)	(0.047)	(0.047)
Year 2020-2021	-38.805	-42.134	-46.510	-0.140***	-0.137***	-0.137***
	(42.836)	(40.584)	(40.766)	(0.049)	(0.049)	(0.049)
Female	-76.023*	-102.788***	-98.280***	-0.020	-0.028	-0.028
	(39.427)	(37.547)	(37.759)	(0.045)	(0.045)	(0.045)
White	-110.640***	-89.033***	-84.686**	-0.076*	-0.071*	-0.071*
	(34.783)	(33.097)	(33.321)	(0.040)	(0.040)	(0.040)
Supervisor Coauthors	-0.539	-0.632*	-0.598*	-0.000	-0.000	-0.000
	(0.350)	(0.331)	(0.333)	(0.000)	(0.000)	(0.000)
Supervisor Citations	-0.001	0.000	0.000	0.000	0.000	0.000
	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)
Theoretical	-27.765	5.326	6.234	0.181***	0.191***	0.191***
	(47.561)	(45.306)	(45.302)	(0.053)	(0.053)	(0.053)
Theo. & Emp.	-82.607**	-68.421*	-69.532*	0.023	0.026	0.026
	(40.227)	(38.163)	(38.167)	(0.046)	(0.045)	(0.046)
N. Authors JMP	-32.002	-19.998	-15.115	0.010	0.015	0.015
	(34.762)	(32.977)	(33.263)	(0.040)	(0.040)	(0.040)
Rank PhD		5.935***	9.880***		0.002**	0.002
		(0.849)	(3.673)		(0.001)	(0.004)
Rank PhD^2			-0.062			-0.000
			(0.057)			(0.000)
N	429	429	429	578	578	578

Notes: Each observation represents a Job Market Candidate (JMC) from 12 universities ranked among the top 100 economic institutions based on the IDEAS/RePEc classification (August 2023), pooled for three academic years: 2018–2019, 2019–2020, and 2020–2021. In columns (1)–(3), the outcome variable is the ranking of the placement obtained by candidate i in academic year t. The sample includes all JMCs that obtained an academic placement (including assistant professor and post-doctoral positions) or a placement at a research institution that is listed in the IDEAS/RePEc ranking. In columns (4)–(6), the outcome variable is a binary variable equal to one if candidate i obtains an AP placement in academic year t, and zero otherwise. We include the full sample and rely and present the average marginal effects from probit models in columns (4)–(6). *Rank PhD* is the ranking of the institution at which the JMC has earned the PhD. *Theory Paper* is a dummy equal to one if the job market paper (JMP) is fully theoretical. *Theory & Emp. Paper* is a dummy taking value one if the JMP has both theoretical and empirical sections. The variables *Supervisor Coauthors* and *Supervisor Citations* are determined by averaging the number of coauthors and the number of citations of the candidate's supervisors, should there be more than one.

Table 3: Statistical Significance and Academic Placement – 10% Significance Level

	(1)	(2)	(3)	(4)	(5)	(6)
			Panel A: Academic Placement			
Significant	0.077	0.108*	0.105**	0.102**	0.103**	0.110*
	(0.064)	(0.055)	(0.049)	(0.049)	(0.050)	(0.057)
Observations	658	658	653	653	466	252
			Panel B: Assistant Professor			
Significant	0.138**	0.144***	0.144***	0.145***	0.154***	0.071
	(0.061)	(0.051)	(0.045)	(0.045)	(0.047)	(0.050)
Observations	658	658	653	653	434	238
Window	[1.65±0.50]	[1.65±0.50]	[1.65±0.50]	[1.65±0.50]	[1.65±0.35]	[1.65±0.20]
JM Year FE		Y	Y	Y	Y	Y
Institution FE		Y	Y	Y	Y	Y
JMP & JMC Info			Y	Y	Y	Y
Advisor Info				Y	Y	Y

Notes: This table reports the estimates of the coefficient δ obtained in equation (4). The coefficients are shown as average marginal effects from the probit model. Each observation is a test statistic. In panel (a), the outcome variable is a dummy equal to one if the job market candidate (JMC) obtains an academic placement (postdoc or AP). In panel (b) the outcome variable takes value one only for AP placements. $Significant$ is a dummy for whether a test statistic is significant at the 10% level. "JMP & JMC Info" includes control variables for the methodology used in the paper (e.g., DiD or IV), the field of the article, the number of authors, gender and race of the job market candidate. "Advisor Info" includes (the average of) the number of citations and number of coauthors of the JMC's supervisor(s). Observations are weighted by the inverse of the number of tests in the paper. Standard errors are clustered at the paper level.

Table 4: Statistical Significance and Academic Placement – 5% Significance Level

	(1)	(2)	(3)	(4)	(5)	(6)
			Panel A: Academic Placement			
Significant	-0.046	-0.014	-0.016	-0.033	-0.076	-0.009
	(0.061)	(0.053)	(0.047)	(0.047)	(0.054)	(0.059)
Observations	677	677	669	669	499	316
			Panel B: Assistant Professor			
Significant	-0.017	0.001	0.025	0.012	-0.011	0.057
	(0.068)	(0.054)	(0.044)	(0.043)	(0.048)	(0.046)
Observations	677	677	669	669	499	316
Window	[1.96±0.50]	[1.96±0.50]	[1.96±0.50]	[1.96±0.50]	[1.96±0.35]	[1.96±0.20]
JM Year FE		Y	Y	Y	Y	Y
Institution FE		Y	Y	Y	Y	Y
JMP & JMC Info			Y	Y	Y	Y
Advisor Info				Y	Y	Y

Notes: This table reports the estimates of the coefficient δ obtained in equation (4). The coefficients are shown as average marginal effects from the probit model. Each observation is a test statistic. In panel (a), the outcome variable is a dummy equal to one if the job market candidate (JMC) obtains an academic placement (postdoc or AP). In panel (b) the outcome variable takes value one only for AP placements. *Significant* is a dummy for whether a test statistic is significant at the 5% level. "JMP & JMC Info" includes control variables for the methodology used in the paper (e.g., DiD or IV), the field of the article, the number of authors, gender and race of the job market candidate. "Advisor Info" includes (the average of) the number of citations and number of coauthors of the JMC's supervisor(s). Observations are weighted by the inverse of the number of tests in the paper. Standard errors are clustered at the paper level.

Table 5: Statistical Significance, Academic Placement, and COVID

	Academic Placement					
	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A: 10% significance level					
Significant	0.027	0.045	0.050	0.052	0.063	0.108
	(0.079)	(0.071)	(0.062)	(0.062)	(0.063)	(0.075)
COVID	-0.194	-0.184	-0.216*	-0.198*	-0.191*	-0.175
	(0.134)	(0.123)	(0.112)	(0.107)	(0.110)	(0.119)
COVID * Significant	0.132	0.173	0.158*	0.149	0.113	-0.000
	(0.127)	(0.115)	(0.093)	(0.093)	(0.097)	(0.106)
Observations	658	658	653	653	466	252
Window	[1.65±0.50]	[1.65±0.50]	[1.65±0.50]	[1.65±0.50]	[1.65±0.35]	[1.65±0.20]
	Panel B: 5% significance level					
Significant	-0.053	-0.035	-0.032	-0.046	-0.105	-0.027
	(0.074)	(0.070)	(0.061)	(0.061)	(0.068)	(0.075)
COVID	-0.126	-0.086	-0.114	-0.104	-0.088	-0.139
	(0.118)	(0.117)	(0.113)	(0.107)	(0.122)	(0.144)
COVID * Significant	0.029	0.034	0.017	0.001	0.056	0.068
	(0.120)	(0.108)	(0.095)	(0.092)	(0.100)	(0.109)
Observations	677	677	669	669	499	316
Window	[1.96±0.50]	[1.96±0.50]	[1.96±0.50]	[1.96±0.50]	[1.96±0.35]	[1.96±0.20]
JM Year FE		Y	Y	Y	Y	Y
Institution FE		Y	Y	Y	Y	Y
JMP & JMC Info			Y	Y	Y	Y
Advisor Info				Y	Y	Y

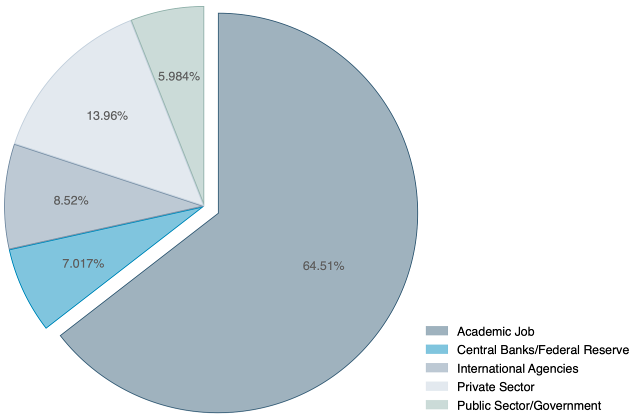
Notes: This table reports the estimates of the coefficient δ obtained from equation (4) with the addition of an interaction term between a dummy variable for marginal significance and *covid*, a dummy variable equal to one during the academic year 2020–2021. The coefficients are shown as average marginal effects from the probit model. Each observation is a test statistic. The outcome variable is a dummy equal to one if the job market candidate(JMC) obtains an academic placement (postdoc or AP). In panel (a), $Significant$ is a dummy for whether a test statistic is significant at the 10% level. In panel (b), $Significant$ is a dummy for whether a test statistic is significant at the 5% level. "JMP & JMC Info" includes control variables for the methodology used in the paper (e.g., DiD or IV), the field of the article, the number of authors, gender and race of the job market candidate. "Advisor Info" includes (the average of) the number of citations and number of coauthors of the JMC's supervisor(s). Observations are weighted by the inverse of the number of tests in the paper. Standard errors are clustered at the paper level.

10 ONLINE APPENDIX

10.1 Methods for p-Hacking

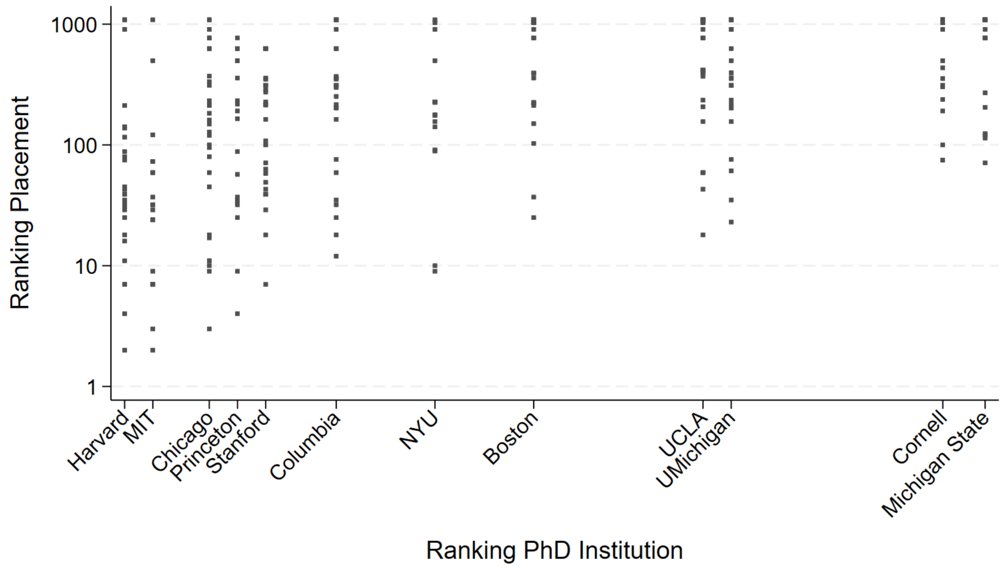
11 Appendix Figures

Figure A1: Job Preferences – PhD candidates



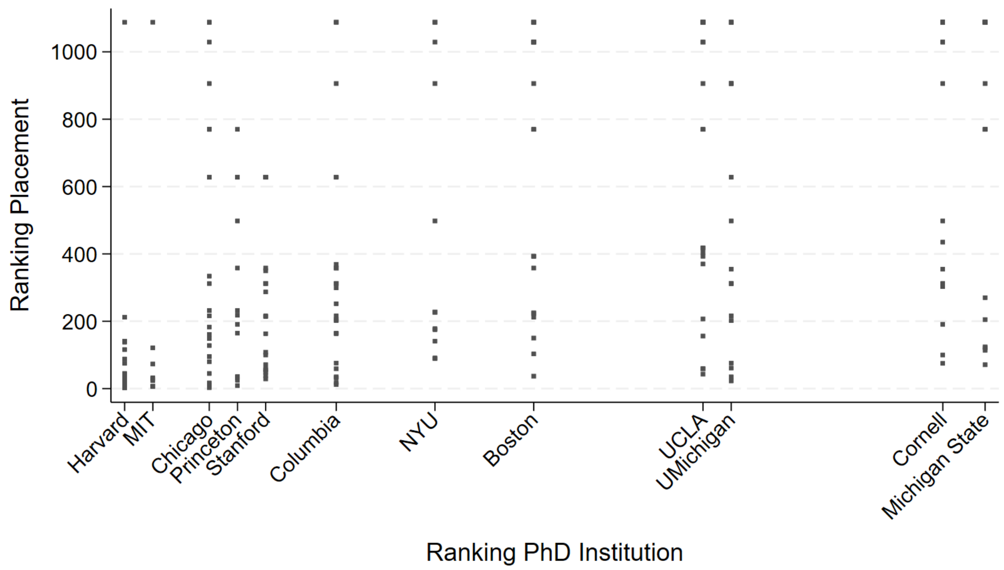
Notes: This figure illustrates the job preferences of PhD candidates from the top 100 economics institutions during the academic year 2022-2023. The data is derived from an online survey sent to 797 recipients, with a response rate of 25%. Private sector placement includes consulting firms, think tanks, and private banks. International organizations include the World Bank, IMF, OECD.

Figure A2: Assistant Professor Placement Ranking (Incl. 1Y PD + AP)



Notes: This figure displays the scatter plot of placement ranking and PhD institution ranking for 300 job market candidates who obtained an assistant professor position after graduation or a 1 year postdoc followed by an AP position. We include job market candidates from 12 universities ranked among the top 100 economic institutions based on the IDEAS/RePEc classification (August 2023), pooled for three academic years: 2018–2019, 2019–2020, and 2020–2021. Universities without a ranking are assigned the lowest ranking (1088). We employ a logarithmic scale for the y axis.

Figure A3: Assistant Professor Placement Ranking

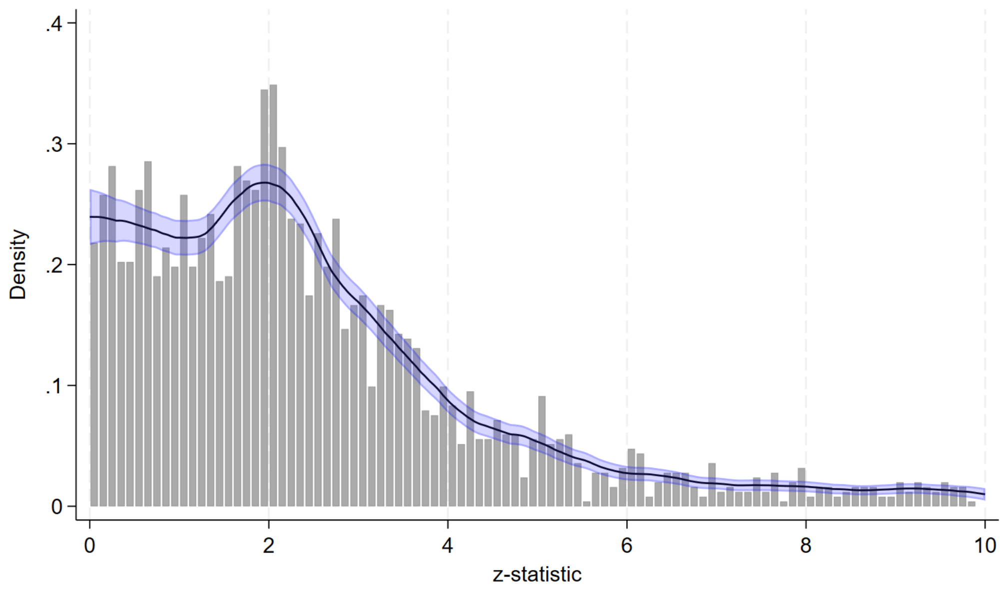


Notes: This figure displays the scatter plot of placement ranking and PhD institution ranking for 221 job market candidates who obtained an assistant professor position after graduation. We include job market candidates from 12 universities ranked among the top 100 economic institutions based on the IDEAS/RePEc classification (August 2023), pooled for three academic years: 2018–2019, 2019–2020, and 2020–2021. Universities without a ranking are assigned the lowest ranking (1088). We employ a linear scale for the y axis.

Figure A4: Placement Ranking (Any Placement)

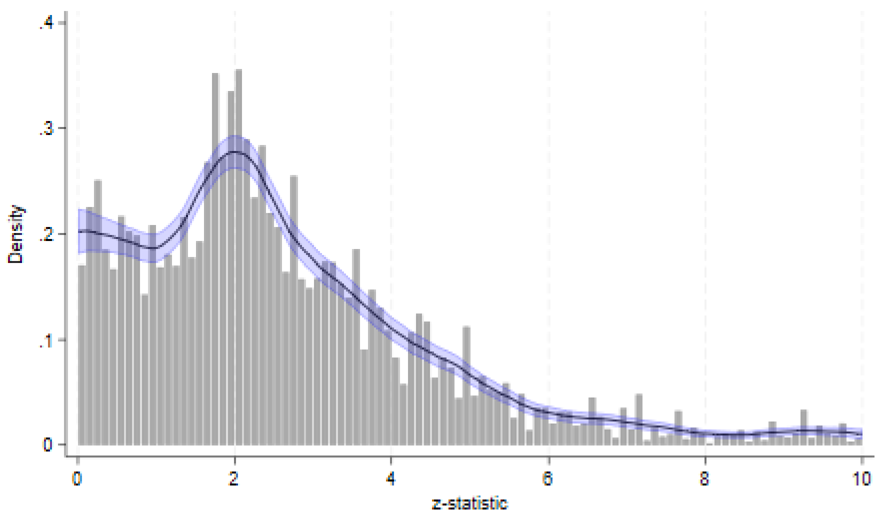
Notes: This figure shows the scatter plot of placement ranking and PhD institution ranking for 443 job market candidates. We include job market candidates from 12 universities ranked among the top 100 economic institutions based on the IDEAS/RePEc classification (August 2023), pooled for three academic years: 2018–2019, 2019–2020, and 2020–2021. We consider all academic placements, together with non-academic institutions that have an IDEAS/RePEc ranking. Universities without a ranking are assigned the lowest ranking (1088). Non-academic institutions are included only if they have a ranking. We employ a logarithmic scale for the y axis.

Figure A5: Unweighted Distribution of Test Statistics



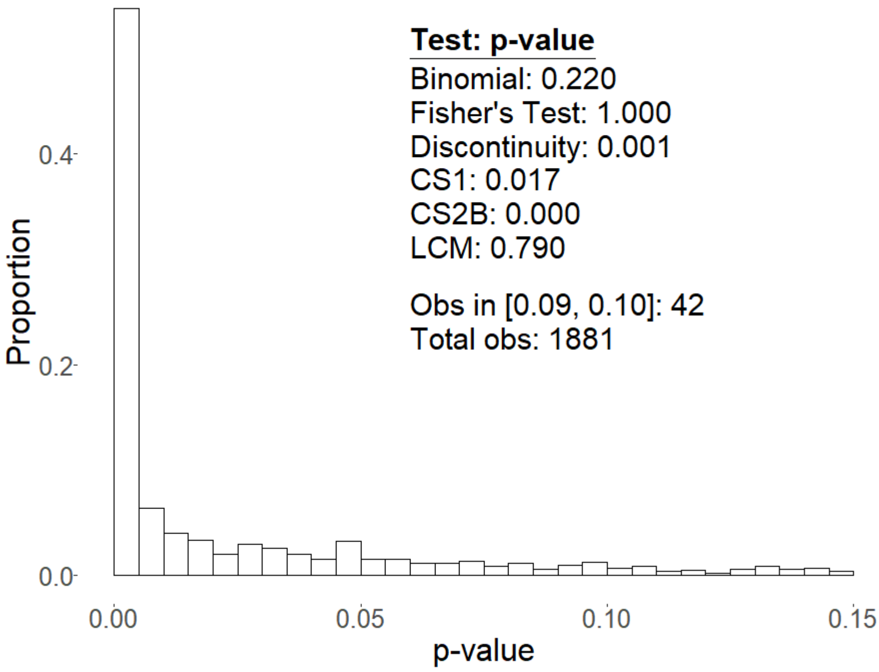
Notes: This figure shows the distribution of 2,708 test statistics for $z \in [0, 10]$ from the 150 empirical job market papers considered in our sample. Bins are 0.1 wide and we superimpose an Epanechnikov kernel. No weighting applied.

Figure A6: De-rounded Distribution of Test Statistics



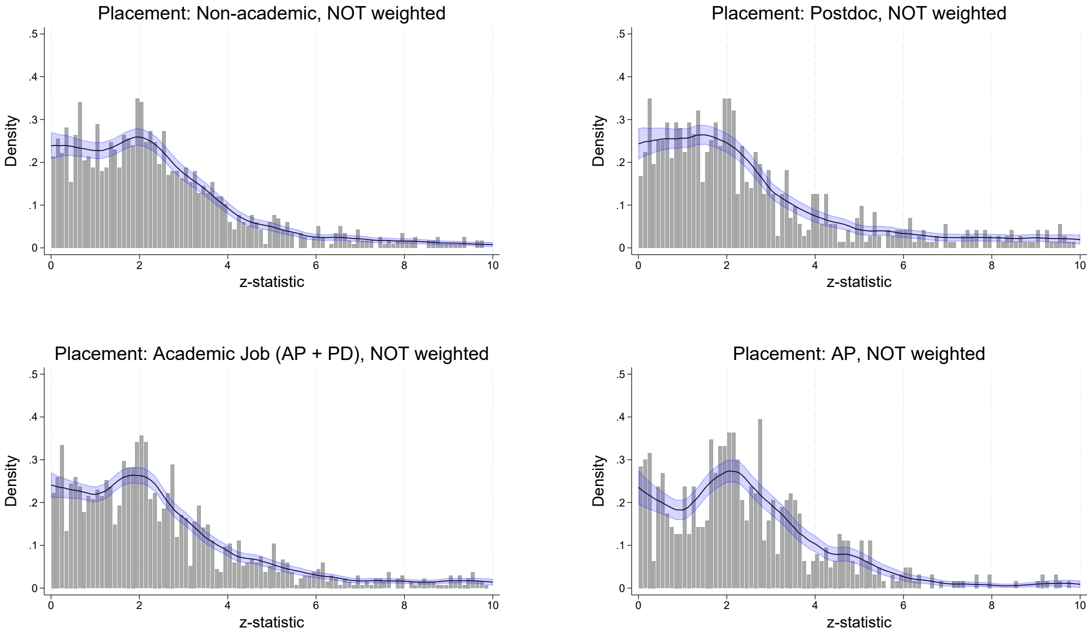
Notes: This figure shows the distribution of 2,708 test statistics for $z \in [0, 10]$ from the 150 empirical job market papers considered in our sample. We adjust for the rounding of test statistics as in Brodeur et al. (2016). Bins are 0.1 wide and we superimpose an Epanechnikov kernel. Observations are weighted by the inverse of the number of tests in the paper.

Figure A7: Elliott et al. (2022) Tests



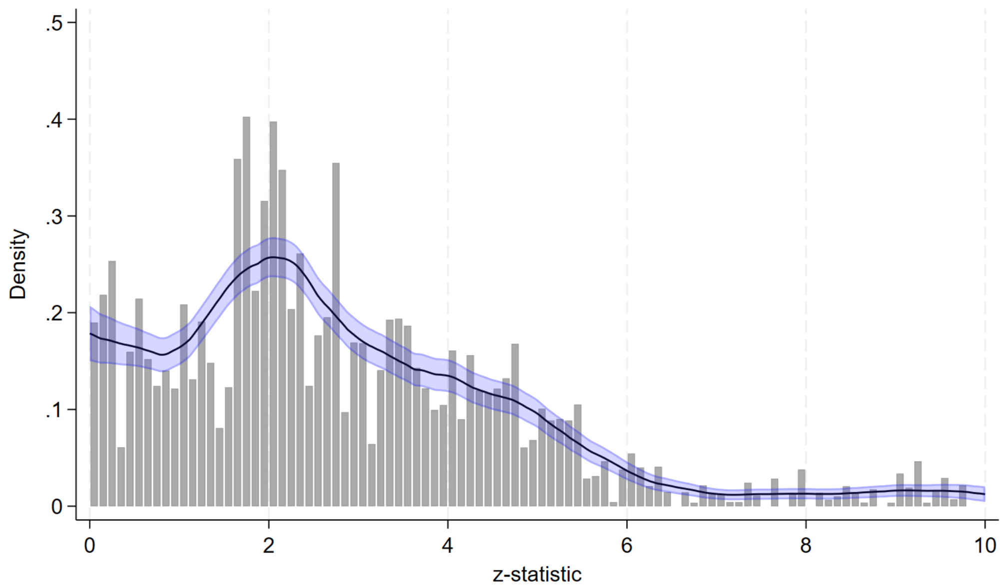
Notes: This figure shows the results of Elliott et al. (2022)'s p-hacking detecting tests. The binomial test compares the mass of test statistics that are just statistically significant to those that are just slightly more statistically significant. For the discontinuity test, under the null hypothesis the estimated density above and below the threshold should be equal. The CS1 (non-increasingness) and CS2B (bounds on the p-curve and its first two derivatives) tests are both histogram-based tests. The LCM test attempts to reject the null that the CDF of the p-curve is concave. See Elliott et al. (2022) for more details.

Figure A8: Unweighted Test Statistics Distribution by Placement



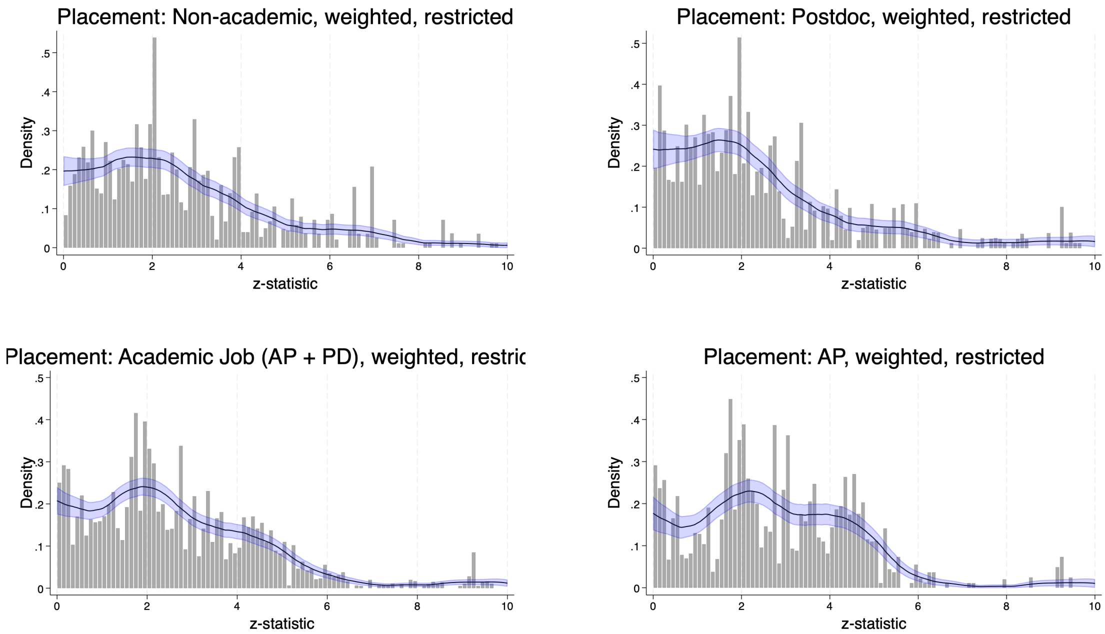
Notes: This figure shows the distribution of test statistics for $z \in [0, 10]$ from 150 job market papers by placement of the job market candidates. The top left panel displays the t-statistics in the job market papers that resulted in a non-academic placement (private sector, government, central banks or international agencies). The top right panel shows the t-statistics for candidates that obtained a postdoc. The bottom left panel shows the distribution of JMPs that resulted in an academic placement (assistant professor or postdoc). The bottom right panel focuses on assistant professor placements only. Candidates obtaining a 1 year postdoc followed by an AP position are classified as postdocs. Bins are 0.1 wide and we superimpose an Epanechnikov kernel. No weighting applied.

Figure A9: Distribution of Test Statistics for AP Placements Including 1 Year Postdocs



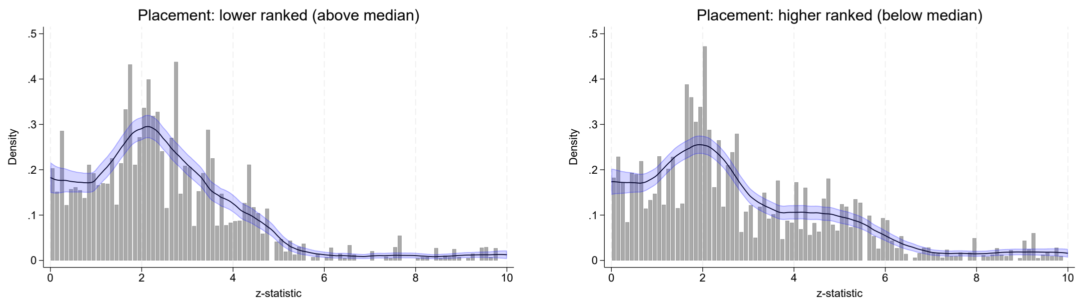
Notes: This figure shows the distribution of test statistics for $z \in [0, 10]$ from the 61 job market papers that resulted in an assistant professor placement or a 1-year postdoc followed by an AP placement. Bins are 0.1 wide and we superimpose an Epanechnikov kernel. Observations are weighted by the inverse of the number of tests in the paper.

Figure A10: Test Statistics Distribution by Placement – Before Initial Interviews



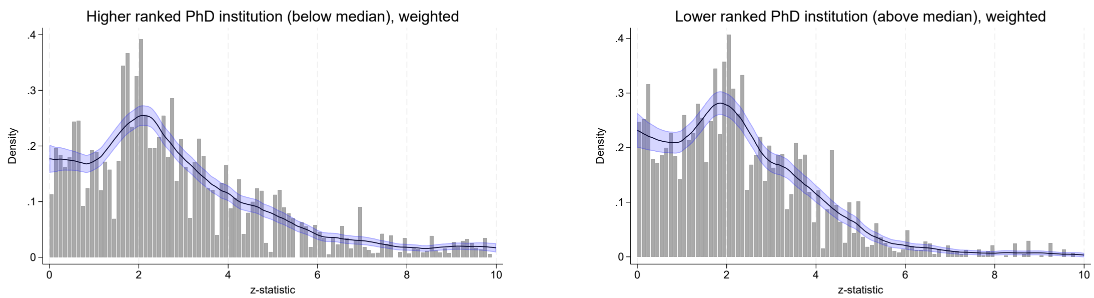
Notes: This figure shows the distribution of test statistics for $z \in [0, 10]$ from 71 job market papers, accessed before December 31st of the academic year, by placement of the job market candidates. The top left panel displays the t-statistics in the job market papers that resulted in a non-academic placement (private sector, government, central banks or international agencies). The top right panel shows the t-statistics for candidates that obtained a postdoc. The bottom left panel shows the distribution of JMPs that resulted in an academic placement (assistant professor or postdoc). The bottom right panel focuses on assistant professor placements only. Candidates obtaining a 1 year postdoc followed by an AP position are classified as postdocs. Bins are 0.1 wide and we superimpose an Epanechnikov kernel. Observations are weighted by the inverse of the number of tests in the paper.

Figure A11: Test Statistics by Placement Ranking



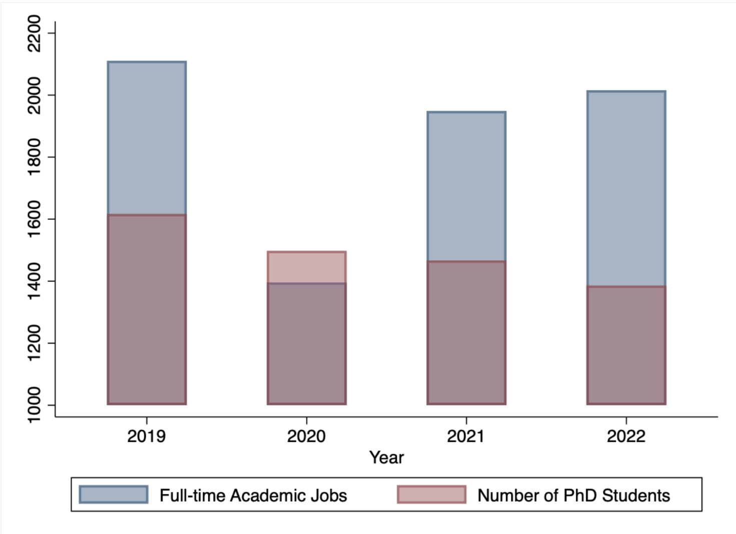
Notes: This figure shows the distribution of test statistics for $z \in [0, 10]$ from 95 job market papers by placement ranking. The left panel shows z-statistics from job market papers that resulted in an above median placement in terms of institution ranking. The right panel displays z-statistics from Job Market Papers that resulted in a below median placement in terms of institution ranking. All universities are included. Universities without a ranking are assigned the lowest ranking (1080). Bins are 0.1 wide and we superimpose an Epanechnikov kernel. Observations are weighted by the inverse of the number of tests in the paper.

Figure A12: Test Statistics by PhD Institution Ranking



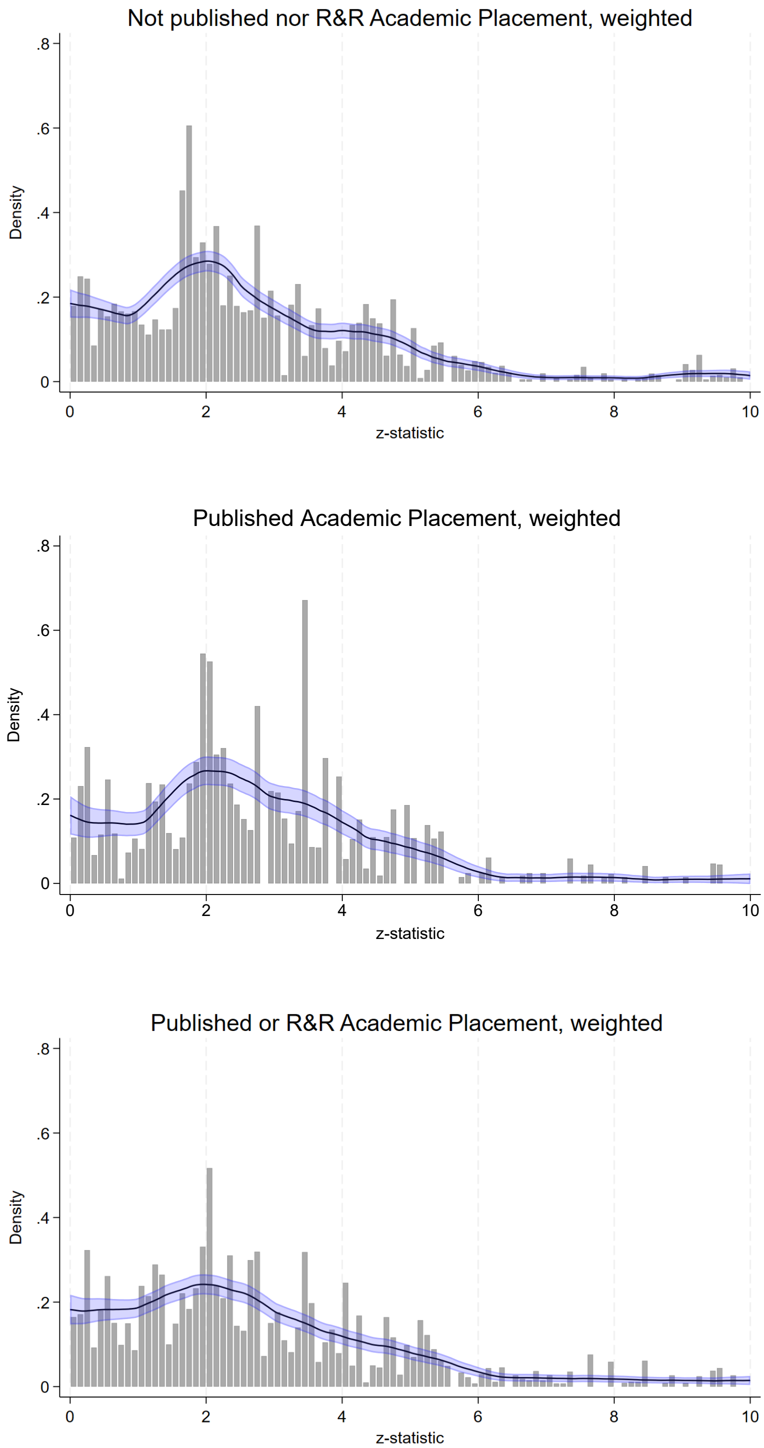
Notes: This figure shows the distribution of test statistics for $z \in [0, 10]$ from 150 job market papers by ranking of the institution of graduation of the job market candidates. The left panel shows the distribution of test statistics for job market papers from higher ranked PhD institutions. The right panel shows papers from lower ranked PhD institutions. Bins are 0.1 wide and we superimpose an Epanechnikov kernel. Observations are weighted by the inverse of the number of tests in the paper.

Figure A13: Supply of PhD Economists and Demand for Academic Positions



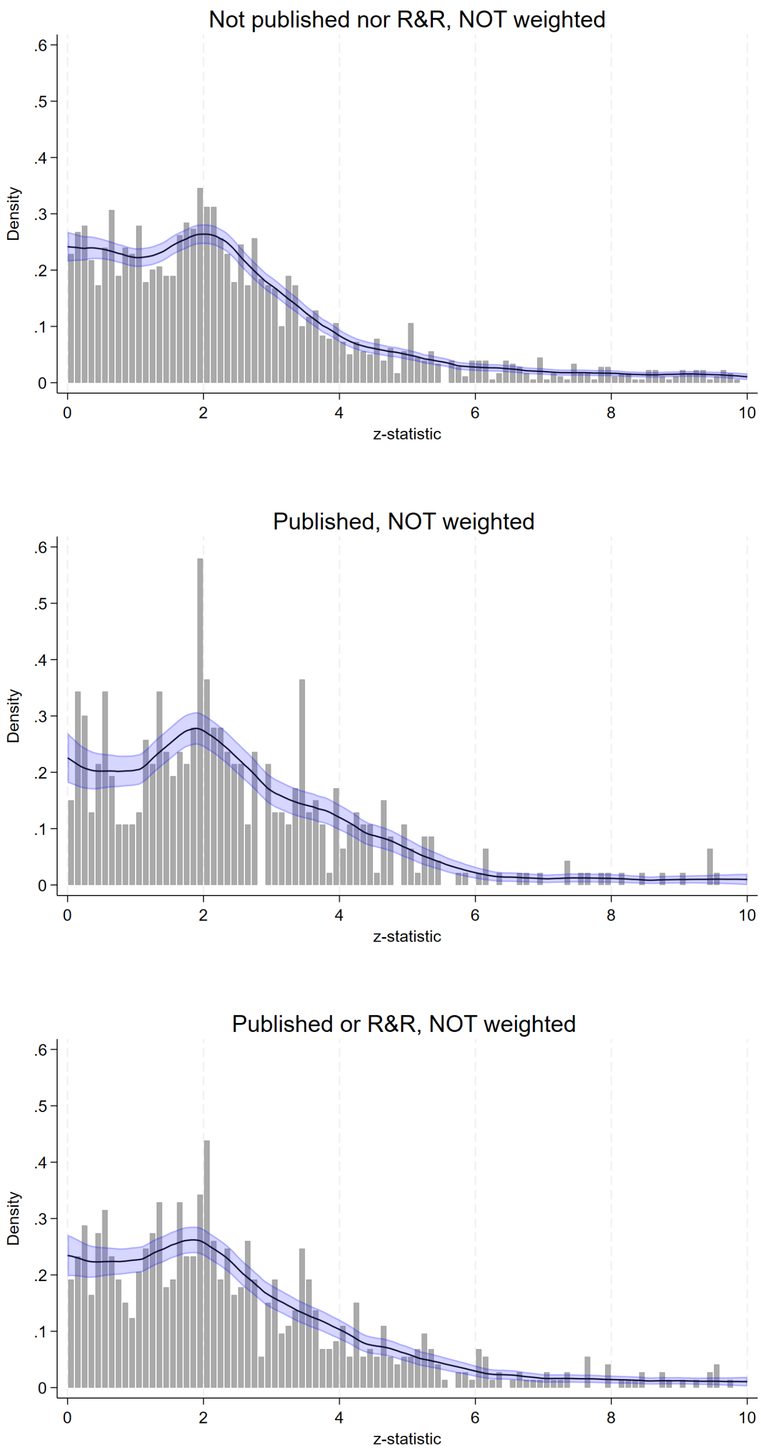
Notes: This figure displays trends in the supply of new PhD economists and demand for full-time academic positions. Data is from AEA Committee on the Job Market, based on data from JOE Cawley (2023). Full-time academic jobs include positions listed on Job Openings for Economists (JOE, AEA) in the US and outside the US. The number of PhD students is proxied by those who applied to at least one job through JOE from August to December of a given year.

Figure A14: Weighted Distribution of Test Statistics by Publication Outcome



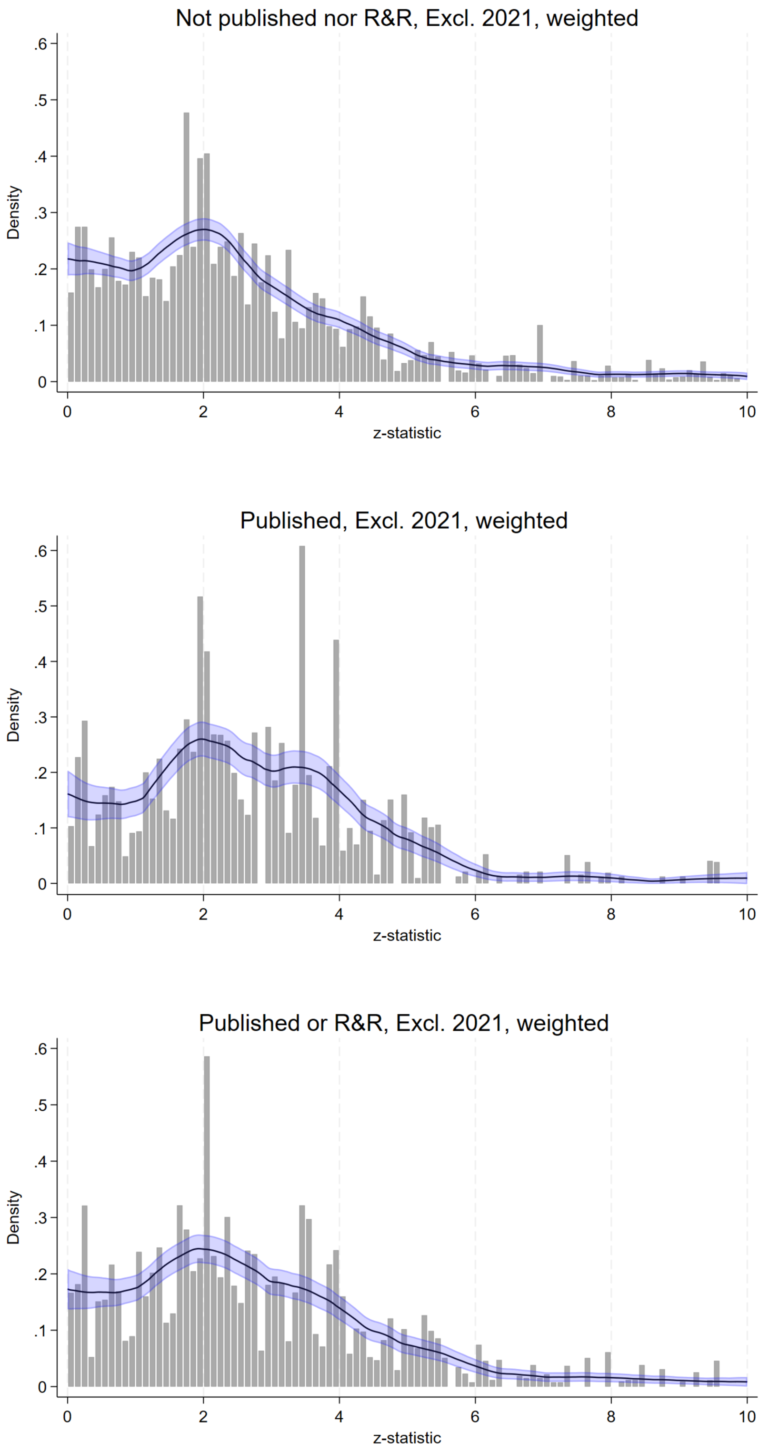
Note: This figure shows the distribution of test statistics for $z \in [0, 10]$ from 79 job market papers by publication outcome. We consider only job market papers from candidates who secured an academic job. The top panel shows the distribution of test statistics for articles that are not published nor R&R. The middle panel panel shows papers that are published. The bottom panel panel shows papers that are published or under revision (R&R). Bins are 0.1 wide and we superimpose an Epanechnikov kernel. Observations are weighted by the inverse of the number of tests in the paper.

Figure A15: Unweighted Distribution of Test Statistics by Publication Outcome



Note: This figure shows the distribution of test statistics for $z \in [0, 10]$ from 150 job market papers by publication outcome. The top panel shows the distribution of test statistics for articles that are not published nor R&R. The middle panel displays the test statistics for published manuscripts. The bottom panel shows papers that are published or under revision (R&R). Bins are 0.1 wide and we superimpose an Epanechnikov kernel. No weighting applied.

Figure A16: Test Statistics by Publication Outcome, excluding AY 2020–2021



Note: This figure shows the distribution of test statistics for $z \in [0, 10]$ from 106 job market papers by publication outcome. We exclude the job market papers of the academic year 2020-2021. The top panel shows the distribution of test statistics for articles that are not published nor R&R. The middle panel displays the test statistics for published manuscripts. The bottom panel shows papers that are published or under revision (R&R). Bins are 0.1 wide and we superimpose an Epanechnikov kernel. Observations are weighted by the inverse of the number of tests in the paper.

12 Appendix Tables

Table A1: Summary Statistics – P-Hacking Analysis – Test Statistics

	Mean	SD	Min	Max
N. Authors JMP	1.245	0.531	1	4
Female	0.433	0.496	0	1
White	0.402	0.490	0	1
Supervisor Citations	28,321	35,267	791	168,567
Supervisor Coauthors	68.097	86.644	4	488
Rank PhD Institution	25.515	20.921	2	63
Rank Placement	243.368	305.731	1	1,088
Academic Placement	0.524	0.499	0	1
AP Placement	0.239	0.426	0	1
Published	0.177	0.382	0	1
Published or R&R	0.282	0.450	0	1

Notes: This table provides an overview of the the distribution of test statistics from 150 selected job market candidates (JMC) and their job market papers (JMP). Each observation is a test statistics. *Theory Paper* is a dummy equal to one if the job market paper is fully theoretical. *Theory & Emp. Paper* is a dummy taking value one if the JMP has both theoretical and empirical sections. *Rank PhD Institution* is the ranking of the institution at which the JMC has earned the PhD. The variables *Supervisor Coauthors* and *Supervisor Citations* are determined by averaging the number of coauthors and the number of citations of the candidate's supervisors, should there be more than one. *Academic Placement* is a dummy equal to one if the candidate obtains an academic placement. *AP placement* is a dummy taking value one if the candidate secures an assistant professor position. *Published* is a dummy equal to one if the JMP is published as of May 2024. *Published or R&R* is a dummy taking value one if the JMP is published or under revision as of May 2024.

Table A2: Placement Determinants, AP positions only

	(1) Placement rank	(2) Placement rank	(3) Placement rank
Year 2019-2020	-12.971	-13.854	-19.778
	(62.427)	(57.177)	(57.191)
Year 2020-2021	2.708	-29.401	-51.237
	(69.887)	(64.209)	(65.893)
Female	-45.511	-100.544*	-92.529
	(63.349)	(58.667)	(58.799)
White	-131.860**	-136.693***	-120.051**
	(56.487)	(51.742)	(52.944)
Supervisor Coauthors	-0.611	-0.428	-0.337
	(0.644)	(0.591)	(0.593)
Supervisor Citations	-0.000	0.000	0.001
	(0.001)	(0.001)	(0.001)
Theory Paper	-90.044	-59.945	-56.063
	(70.294)	(64.557)	(64.458)
Theory & Emp. Paper	-68.578	-78.019	-79.174
	(65.334)	(59.858)	(59.717)
N. Authors JMP	-75.796	-37.673	-16.966
	(61.906)	(57.017)	(58.738)
Rank PhD		8.020***	16.430***
		(1.265)	(6.087)
Rank PhD^2			-0.127
			(0.090)
Observations	214	214	214

Notes: Each observation represents a Job Market Candidate (JMC) from 12 universities ranked among the top 100 economic institutions based on the IDEAS/RePEc classification (August 2023), pooled for three academic years: 2018–2019, 2019–2020, and 2020–2021. The outcome variable is the ranking of the placement obtained by candidate i in academic year t. The sample includes only JMCs that obtained an Assistant Professor position. *Rank PhD* is the ranking of the institution at which the JMC has earned the PhD. *Theory Paper* is a dummy equal to one if the job market paper (JMP) is fully theoretical. *Theory & Emp. Paper* is a dummy taking value one if the JMP has both theoretical and empirical sections. The *Supervisor Coauthors* and *Supervisor Citations* are determined by averaging the number of coauthors and the number of citations of the candidate's supervisors, should there be more than one.

Table A3: Randomization Tests, 10 and 5 percent significance thresholds

	Full Sample	Academic	Non Academic	Difference	AP	Non AP	Difference
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: 10% Significance Threshold							
Proportion significant in 1.65 ± 0.5	0.639	0.672	0.599	0.072	0.748	0.596	0.152
Number of tests in 1.65 ± 0.5	658	364	294		165	493	
One sided p-value				[0.028]			[0.0001]
Proportion significant in 1.65 ± 0.35	0.650	0.704	0.584	0.119	0.773	0.600	0.173
Number of tests in 1.65 ± 0.35	471	256	215		120	351	
One sided p-value				[0.004]			[0.0001]
Proportion significant in 1.65 ± 0.2	0.640	0.695	0.563	0.132	0.736	0.594	0.142
Number of tests in 1.65 ± 0.2	254	141	113		67	187	
One sided p-value				[0.016]			[0.015]
Panel B: 5% Significance Threshold							
Proportion significant in 1.96 ± 0.5	0.511	0.490	0.537	-0.046	0.497	0.517	-0.019
Number of tests in 1.96 ± 0.5	677	362	315		180	497	
One sided p-value				[0.888]			[0.672]
Proportion significant in 1.96 ± 0.35	0.502	0.464	0.551	-0.087	0.467	0.519	-0.053
Number of tests in 1.96 ± 0.35	504	280	224		143	361	
One sided p-value				[0.973]			[0.858]
Proportion significant in 1.96 ± 0.2	0.552	0.547	0.557	-0.010	0.619	0.525	0.093
Number of tests in 1.96 ± 0.2	319	180	2139		86	233	
One sided p-value				[0.567]			[0.067]

Notes: This table presents the results of binomial proportion tests, where a success is defined as a statistically significant observation at the 10% and 5% threshold level in Panels A and B, respectively. For each threshold level $z*$, we consider observations where $z = [z * 0.5]$, $z = [z * 0.35]$, and $z = [z * 0.2]$. Column (1) displays the share of marginally significant test statistics for the full sample of our analysis. Columns (2) and (3) show the proportion of academic versus non-academic subsamples (i.e., test statistics associated with JMPs of JMCs who obtained an academic placement). In columns (4), we test whether the proportions in academic placement are statistically greater than the non-academic placement ones. Columns (5) and (6) display the proportion of AP versus non-AP (i.e., test statistics associated with JMPs of JMCs who obtained an AP placement). In columns (7), we test whether the proportions in AP placement are statistically greater than the non-AP placement ones. The associated p-values in columns (4) and (7) are reported. Observations are weighted by the inverse of the number of tests in the paper.

Table A4: Determinants of Marginal Significance

	10% level (1)	5% level (2)	1% level (3)
Year 2019-2020	0.062	0.104	0.099
	(0.071)	(0.069)	(0.071)
Year 2020-2021	-0.073	0.049	0.057
	(0.065)	(0.067)	(0.075)
Female	0.051	-0.128**	0.024
	(0.056)	(0.058)	(0.058)
White	0.017	0.129**	-0.010
	(0.058)	(0.058)	(0.058)
Supervisor Coauthors	-0.000	0.001	0.000
	(0.000)	(0.000)	(0.000)
Supervisor Citations	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)
Theoretical Model	-0.191**	-0.125*	-0.195*
	(0.091)	(0.074)	(0.104)
N. Authors JMP	0.026	-0.012	0.009
	(0.069)	(0.050)	(0.064)
Rank PhD	0.005	-0.000	0.001
	(0.006)	(0.006)	(0.007)
Rank PhD^2	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)
N	653	669	528
Window	[1.65±0.50]	[1.96±0.50]	[2.58±0.50]
Weighted	Yes	Yes	Yes
Cluster	Paper	Paper	Paper

Notes: Each observation represents a test statistic. The outcome variable in columns (1)–(3) is a dummy for whether a test statistic is significant at the 10%, 5%, and 1%, respectively. The coefficients are shown as average marginal effects from the probit model. *Rank PhD* is the ranking of the institution at which the JMC has earned the PhD. *Theoretical Model* is a dummy equal to one if the job market paper (JMP) includes a formal theoretical sections. The variables *Supervisor Coauthors* and *Supervisor Citations* are determined by averaging the number of coauthors and the number of citations of the candidate's supervisors, should there be more than one. Observations are weighted by the inverse of the number of tests in the paper. Standard errors are clustered at the paper level.

Table A5: Statistical Significance and Academic Placement – 10% Significance Level – Unweighted

	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A: Academic Placement					
Significant	0.011	0.052	0.059	0.055	0.073*	0.067
	(0.059)	(0.045)	(0.038)	(0.038)	(0.040)	(0.048)
Observations	658	658	653	653	466	252
	Panel B: Assistant Professor					
Significant	0.060	0.083**	0.083***	0.082**	0.109***	0.027
	(0.049)	(0.036)	(0.032)	(0.032)	(0.034)	(0.040)
Observations	658	658	653	653	434	238
Window	[1.65±0.50]	[1.65±0.50]	[1.65±0.50]	[1.65±0.50]	[1.65±0.35]	[1.65±0.20]
JM Year FE		Y	Y	Y	Y	Y
Institution FE		Y	Y	Y	Y	Y
JMP & JMC Info			Y	Y	Y	Y
Advisor info				Y	Y	Y

Notes: This table reports the estimates of the coefficient δ obtained in equation (4). The coefficients are shown as average marginal effects from the probit model. Each observation is a test statistic. In panel (a), the outcome variable is a dummy equal to one if the job market candidate (JMC) obtains an academic placement (postdoc or AP). In panel (b) the outcome variable takes value one only for AP placements. *Significant* is a dummy for whether a test statistic is significant at the 10% level. "JMP & JMC Info" includes control variables for the methodology used in the paper (e.g., DiD or IV), the field of the article, the number of authors, gender and race of the job market candidate. "Advisor Info" includes (the average of) the number of citations and number of coauthors of the JMC's supervisor(s). No weighting applied. Standard errors are clustered at the paper level.

Table A6: Statistical Significance and Academic Placement – 10% Significance Level – De-Rounded

	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A: Academic Placement					
Significant	0.060	0.088	0.088*	0.088*	0.066	0.070
	(0.064)	(0.056)	(0.049)	(0.049)	(0.053)	(0.065)
Observations	654	654	649	649	452	246
	Panel B: Assistant Professor					
Significant	0.115*	0.116**	0.110**	0.111**	0.075*	-0.055
	(0.060)	(0.049)	(0.045)	(0.044)	(0.044)	(0.052)
Observations	654	654	649	649	421	233
Window	[1.65±0.50]	[1.65±0.50]	[1.65±0.50]	[1.65±0.50]	[1.65±0.35]	[1.65±0.20]
JM Year FE		Y	Y	Y	Y	Y
Institution FE		Y	Y	Y	Y	Y
JMP & JMC Info			Y	Y	Y	Y
Advisor Info				Y	Y	Y

Notes: This table reports the estimates of the coefficient δ obtained in equation (4). The coefficients are shown as average marginal effects from the probit model. Each observation is a test statistic. Test statistics are de-rounded as in Brodeur et al. (2016). In panel (a), the outcome variable is a dummy equal to one if the job market candidate (JMC) obtains an academic placement (postdoc or AP). In panel (b) the outcome variable takes value one only for AP placements. $Significant$ is a dummy for whether a test statistic is significant at the 10% level. "JMP & JMC Info" includes control variables for the methodology used in the paper (e.g., DiD or IV), the field of the article, the number of authors, gender and race of the job market candidate. "Advisor Info" includes (the average of) the number of citations and number of coauthors of the JMC's supervisor(s). Observations are weighted by the inverse of the number of tests in the paper. Standard errors are clustered at the paper level.

Table A7: Statistical Significance and Academic Placement – 5% Significance Level – De-Rounded

	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A: Academic Placement					
Significant	-0.048	-0.007	-0.005	-0.022	-0.039	-0.053
	(0.060)	(0.052)	(0.047)	(0.047)	(0.052)	(0.055)
Observations	689	689	680	680	507	324
	Panel B: Assistant Professor					
Significant	-0.029	0.000	0.016	0.002	-0.004	-0.009
	(0.066)	(0.050)	(0.042)	(0.040)	(0.043)	(0.047)
Observations	689	689	680	680	507	324
Window	[1.96±0.50]	[1.96±0.50]	[1.96±0.50]	[1.96±0.50]	[1.96±0.35]	[1.96±0.20]
JM Year FE		Y	Y	Y	Y	Y
Institution FE		Y	Y	Y	Y	Y
JMP & JMC Info			Y	Y	Y	Y
Advisor Info				Y	Y	Y

Notes: This table reports the estimates of the coefficient δ obtained in equation (4). The coefficients are shown as average marginal effects from the probit model. Each observation is a test statistic. Test statistics are de-rounded as in Brodeur et al. (2016). In panel (a), the outcome variable is a dummy equal to one if the job market candidate (JMC) obtains an academic placement (postdoc or AP). In panel (b) the outcome variable takes value one only for AP placements. $Significant$ is a dummy for whether a test statistic is significant at the 5% level. "JMP & JMC Info" includes control variables for the methodology used in the paper (e.g., DiD or IV), the field of the article, the number of authors, gender and race of the job market candidate. "Advisor Info" includes (the average of) the number of citations and number of coauthors of the JMC's supervisor(s). Observations are weighted by the inverse of the number of tests in the paper. Standard errors are clustered at the paper level.

Table A8: Statistical Significance and Academic Placement – 1% Significance Level

	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A: Academic Placement					
Significant	0.073	0.060	0.071	0.070	0.089*	0.156**
	(0.056)	(0.052)	(0.046)	(0.047)	(0.048)	(0.070)
Observations	537	537	528	528	347	210
	Panel B: Assistant Professor					
Significant	0.057	0.057	0.084**	0.082*	0.071	0.224***
	(0.055)	(0.051)	(0.043)	(0.043)	(0.045)	(0.054)
Observations	537	537	528	528	347	210
Window	[2.58±0.50]	[2.58±0.50]	[2.58±0.50]	[2.58±0.50]	[2.58±0.35]	[2.58±0.20]
JM Year FE		Y	Y	Y	Y	Y
Institution FE		Y	Y	Y	Y	Y
JMP & JMC Info			Y	Y	Y	Y
Advisor Info				Y	Y	Y

Notes: This table reports the estimates of the coefficient δ obtained in equation (4). The coefficients are shown as average marginal effects from the probit model. Each observation is a test statistic. In panel (a), the outcome variable is a dummy equal to one if the job market candidate (JMC) obtains an academic placement (postdoc or AP). In panel (b) the outcome variable takes value one only for AP placements. $Significant$ is a dummy for whether a test statistic is significant at the 1% level. "JMP & JMC Info" includes control variables for the methodology used in the paper (e.g., DiD or IV), the field of the article, the number of authors, gender and race of the job market candidate. "Advisor Info" includes (the average of) the number of citations and number of coauthors of the JMC's supervisor(s). Observations are weighted by the inverse of the number of tests in the paper. Standard errors are clustered at the paper level.

Table A9: Statistical Significance and Academic Placement – Other Outcomes – 10% Significance Level

	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A: Assistant Professor including 1 year postdocs					
Significant	0.114*	0.126**	0.126***	0.125***	0.139***	0.079
	(0.062)	(0.056)	(0.047)	(0.047)	(0.048)	(0.056)
Observations	537	537	528	528	347	210
	Panel B: Assistant Professor Conditional on Academic Placement					
Significant	0.176**	0.152**	0.148***	0.152***	0.147***	-0.007
	(0.087)	(0.071)	(0.051)	(0.049)	(0.053)	(0.052)
Observations	364	364	364	364	243	130
	Panel C: Academic Placement Rank (log)					
Significant	-0.081	-0.087	-0.057	-0.030	-0.071	-0.200**
	(0.167)	(0.087)	(0.066)	(0.061)	(0.078)	(0.090)
Observations	432	432	432	432	308	171
Window	[1.65±0.50]	[1.65±0.50]	[1.65±0.50]	[1.65±0.50]	[1.65±0.35]	[1.65±0.20]
JM Year FE		Y	Y	Y	Y	Y
Institution FE		Y	Y	Y	Y	Y
JMP & JMC Info			Y	Y	Y	Y
Advisor Info				Y	Y	Y

Notes: This table reports the estimates of the coefficient δ obtained in equation (4). The coefficients are shown as average marginal effects from the probit model. Each observation is a test statistic. In panel (a), the outcome variable is a dummy equal to one if the job market candidate (JMC) obtains an assistant professor placement or a 1 year postdoc followed by an AP position. In panel (b), the outcome variable takes value one only for AP placements. We restrict the sample to candidates obtaining an academic placement. In panel (c), the outcome variable is the log ranking of the placement obtained by the candidate. We restrict the sample to all JMCs that obtained an academic placement (including assistant professor and post-doctoral positions) or a placement at a research institution that is listed in the Ideas/RePEc ranking. *Significant* is a dummy for whether a test statistic is significant at the 10% level. "JMP & JMC Info" includes control variables for the methodology used in the paper (e.g., DiD or IV), the field of the article, the number of authors, gender and race of the job market candidate. "Advisor Info" includes (the average of) the number of citations and number of coauthors of the JMC's supervisor(s). Observations are weighted by the inverse of the number of tests in the paper. Standard errors are clustered at the paper level.

Table A10: Statistical Significance, Academic Placement, and COVID – De-Rounded

	Academic Placement					
	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A: 10% significance level					
Significant	0.003	0.017	0.025	0.028	0.017	0.029
	(0.074)	(0.068)	(0.059)	(0.059)	(0.063)	(0.080)
COVID	-0.229*	-0.228*	-0.259**	-0.239**	-0.214*	-0.232*
	(0.133)	(0.123)	(0.114)	(0.109)	(0.115)	(0.132)
COVID * Significant	0.179	0.230**	0.213**	0.204**	0.176*	0.144
	(0.122)	(0.109)	(0.090)	(0.091)	(0.106)	(0.126)
Observations	654	654	649	649	452	246
Window	[1.65±0.50]	[1.65±0.50]	[1.65±0.50]	[1.65±0.50]	[1.65±0.35]	[1.65±0.20]
	Panel B: 5% significance level					
Significant	-0.051	-0.029	-0.023	-0.036	-0.077	-0.065
	(0.069)	(0.065)	(0.056)	(0.056)	(0.061)	(0.065)
COVID	-0.129	-0.087	-0.118	-0.104	-0.119	-0.114
	(0.115)	(0.114)	(0.112)	(0.107)	(0.115)	(0.127)
COVID * Significant	0.040	0.041	0.022	0.003	0.086	0.042
	(0.109)	(0.097)	(0.087)	(0.083)	(0.083)	(0.104)
Observations	689	689	680	680	507	324
Window	[1.96±0.50]	[1.96±0.50]	[1.96±0.50]	[1.96±0.50]	[1.96±0.35]	[1.96±0.20]
JM Year FE		Y	Y	Y	Y	Y
Institution FE		Y	Y	Y	Y	Y
JMP & JMC Info			Y	Y	Y	Y
Advisor Info				Y	Y	Y

Notes: This table reports the estimates of the coefficient δ obtained from equation (4) with the addition of an interaction term between a dummy variable for marginal significance and *covid*, a dummy variable equal to one during the academic year 2020–2021. The coefficients are shown as average marginal effects from the probit model. Each observation is a test statistic. Test statistics are de-rounded as in Brodeur et al. (2016). The outcome variable is a dummy equal to one if the job market candidate(JMC) obtains an academic placement (postdoc or AP). In panel (a), $Significant$ is a dummy for whether a test statistic is significant at the 10% level. In panel (b), $Significant$ is a dummy for whether a test statistic is significant at the 5% level. "JMP & JMC Info" includes control variables for the methodology used in the paper (e.g., DiD or IV), the field of the article, the number of authors, gender and race of the job market candidate. "Advisor Info" includes (the average of) the number of citations and number of coauthors of the JMC's supervisor(s). Observations are weighted by the inverse of the number of tests in the paper. Standard errors are clustered at the paper level.

Table A11: Statistical Significance, AP, and COVID

	Assistant Professor					
	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A: 10% significance level					
Significant	0.134*	0.133**	0.131**	0.134**	0.131**	0.081
	(0.079)	(0.068)	(0.060)	(0.059)	(0.067)	(0.072)
COVID	-0.055	-0.038	-0.038	-0.027	0.029	0.110
	(0.118)	(0.117)	(0.114)	(0.110)	(0.119)	(0.120)
COVID * Significant	-0.001	0.015	0.035	0.024	0.061	0.007
	(0.105)	(0.100)	(0.096)	(0.092)	(0.104)	(0.110)
Observations	658	658	653	653	434	238
Window	[1.65±0.50]	[1.65±0.50]	[1.65±0.50]	[1.65±0.50]	[1.65±0.35]	[1.65±0.20]
	Panel B: 5% significance level					
Significant	-0.009	0.005	0.038	0.024	-0.005	0.040
	(0.084)	(0.073)	(0.057)	(0.057)	(0.064)	(0.059)
COVID	-0.059	-0.029	-0.016	-0.010	0.003	-0.099
	(0.119)	(0.109)	(0.105)	(0.104)	(0.112)	(0.109)
COVID * Significant	-0.026	-0.032	-0.051	-0.055	-0.030	0.059
	(0.128)	(0.121)	(0.105)	(0.107)	(0.115)	(0.077)
Observations	677	677	669	669	499	316
Window	[1.96±0.50]	[1.96±0.50]	[1.96±0.50]	[1.96±0.50]	[1.96±0.35]	[1.96±0.20]
JM Year FE		Y	Y	Y	Y	Y
Institution FE		Y	Y	Y	Y	Y
JMP & JMC Info			Y	Y	Y	Y
Advisor Info				Y	Y	Y

Notes: This table reports the estimates of the coefficient δ obtained from equation (4) with the addition of an interaction term between our dummy variable for marginal significance and *covid*, a dummy variable that takes the value of one during the academic year 2020–2021. The coefficients are shown as average marginal effects from the probit model. Each observation is a test statistic. The outcome variable is a dummy equal to one if the JMC obtains an assistant professor placement. In panel (a), $Significant$ is a dummy for whether a test statistic is significant at the 10% level. In panel (b), $Significant$ is a dummy for whether a test statistic is significant at the 5% level. "JMP & JMC Info" includes control variables for the methodology used in the paper (e.g., DiD or IV), the field of the article, the number of authors, gender and race of the job market candidate. "Advisor Info" includes (the average of) the number of citations and number of coauthors of the JMC's supervisor(s). Observations are weighted by the inverse of the number of tests in the paper. Standard errors are clustered at the paper level.

Table A12: Statistical Significance and Publication – Published Only

	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A: 10% Significance Threshold					
Significant	0.044	0.082*	0.090**	0.090**	0.112***	0.131***
	(0.051)	(0.045)	(0.038)	(0.037)	(0.040)	(0.038)
Observations	658	549	544	544	385	190
Window	[1.65±0.50]	[1.65±0.50]	[1.65±0.50]	[1.65±0.50]	[1.65±0.35]	[1.65±0.20]
	Panel B: 5% Significance Threshold					
Significant	0.060	0.038	0.049	0.050	0.029	0.043
	(0.047)	(0.034)	(0.035)	(0.034)	(0.035)	(0.041)
Observations	677	559	551	551	405	252
Window	[1.96±0.50]	[1.96±0.50]	[1.96±0.50]	[1.96±0.50]	[1.96±0.35]	[1.96±0.20]
JM Year FE		Y	Y	Y	Y	Y
Institution FE		Y	Y	Y	Y	Y
JMP & JMC Info			Y	Y	Y	Y
Advisor Info				Y	Y	Y

Notes: This table reports the estimates of the coefficient δ obtained in equation (4). The coefficients are shown as average marginal effects from the probit model. Each observation is a test statistic. The outcome variable is a dummy equal to one if the job market paper is published as of May 2024. In panel (a), $Significant$ is a dummy for whether a test statistic is significant at the 10% level. In panel (b), $Significant$ is a dummy for whether a test statistic is significant at the 5% level. "JMP & JMC Info" includes control variables for the methodology used in the paper (e.g., DiD or IV), the field of the article, the number of authors, gender and race of the job market candidate. "Advisor Info" includes (the average of) the number of citations and number of coauthors of the JMC's supervisor(s). Observations are weighted by the inverse of the number of tests in the paper. Standard errors are clustered at the paper level.

Table A13: Statistical Significance and Publication – Published or R&R

	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A: 10% Significance Threshold					
Significant	-0.039	0.027	0.039	0.039	0.051	0.101**
	(0.060)	(0.050)	(0.040)	(0.041)	(0.042)	(0.045)
Observations	658	613	608	608	435	219
Window	[1.65±0.50]	[1.65±0.50]	[1.65±0.50]	[1.65±0.50]	[1.65±0.35]	[1.65±0.20]
	Panel B: 5% Significance Threshold					
Significant	0.057	0.042	0.039	0.047	0.031	0.096**
	(0.051)	(0.040)	(0.035)	(0.036)	(0.037)	(0.044)
Observations	677	621	613	613	455	285
Window	[1.96±0.50]	[1.96±0.50]	[1.96±0.50]	[1.96±0.50]	[1.96±0.35]	[1.96±0.20]
JM Year FE		Y	Y	Y	Y	Y
Institution FE		Y	Y	Y	Y	Y
JMP & JMC Info			Y	Y	Y	Y
Advisor Info				Y	Y	Y

Notes: This table reports the estimates of the coefficient δ obtained in equation (4). The coefficients are shown as average marginal effects from the probit model. Each observation is a test statistic. The outcome variable is a dummy equal to one if the job market paper is published or under revision as of May 2024. In panel (a), $Significant$ is a dummy for whether a test statistic is significant at the 10% level. In panel (b), $Significant$ is a dummy for whether a test statistic is significant at the 5% level. "JMP & JMC Info" includes control variables for the methodology used in the paper (e.g., DiD or IV), the field of the article, the number of authors, gender and race of the job market candidate. "Advisor Info" includes (the average of) the number of citations and number of coauthors of the JMC's supervisor(s). Observations are weighted by the inverse of the number of tests in the paper. Standard errors are clustered at the paper level.

Table A14: Statistical Significance, Academic Placement, and Publication – Published – 10% Significance Level

	Published					
	(1)	(2)	(3)	(4)	(5)	(6)
Significant	0.034	0.070	0.086**	0.086**	0.107***	0.127***
	(0.050)	(0.045)	(0.039)	(0.038)	(0.041)	(0.041)
Academia	0.128	0.117	0.139**	0.141**	0.099	0.067
	(0.087)	(0.079)	(0.069)	(0.070)	(0.078)	(0.079)
Observations	658	549	544	544	385	190
Window	[1.65±0.50]	[1.65±0.50]	[1.65±0.50]	[1.65±0.50]	[1.65±0.35]	[1.65±0.20]
JM Year FE		Y	Y	Y	Y	Y
Institution FE		Y	Y	Y	Y	Y
JMP & JMC Info			Y	Y	Y	Y
Advisor Info				Y	Y	Y

Notes: This table reports the estimates of the coefficient δ obtained in equation (4), with the addition of the control variable *Academia*, that is a dummy variable equal to one if the job market candidates obtain an academic placement. The coefficients are shown as average marginal effects from the probit model. Each observation is a test statistic. The outcome variable is a dummy equal to one if the job market paper is published or under revision as of May 2024. *Significant* is a dummy for whether a test statistic is significant at the 10% level. "JMP & JMC Info" includes control variables for the methodology used in the paper (e.g., DiD or IV), the field of the article, the number of authors, gender and race of the job market candidate. "Advisor Info" includes (the average of) the number of citations and number of coauthors of the JMC's supervisor(s). Observations are weighted by the inverse of the number of tests in the paper. Standard errors are clustered at the paper level.

13 Deviations from the Pre-Analysis Plan

In this section we list the deviations from the pre-analysis plan (PAP)[41] of this research project:

1. The analysis of the determinants of academic job market success was not pre-registered.

2. In the pre-analysis plan (PAP), we included an analysis of the potential link between p-hacking by supervisors and PhD candidates. However, we did not include this analysis in the final paper because the z-statistics available in the existing literature did not provide a sufficient sample size to draw reliable conclusions.

3. We initially planned to investigate whether placements in the top 25, 50, and 100 institutions are associated with higher levels of p-hacking. However, our sample did not include a sufficient number of such placements to conduct a well-powered analysis. Therefore, we shifted our focus to comparing placements above and below the median.

4. We initially pre-registered to use the IDEAS/RePEc Economics Departments ranking. However, we decided to use the IDEAS/RePEc Economic Institutions ranking instead, as it encompasses a significantly larger number of universities and research institutions.

5. We deviated from the PAP in the categories used to classify the fields of the job market paper topics. This adjustment was made to ensure that each category had a sufficient number of observations.

6. In the main regression, we included the following control variables that were not specified in the pre-analysis plan: a dummy variable indicating whether a candidate is white, the number of citations of the candidate's supervisor, and

[41] Available here: https://osf.io/pe392/.

the number of coauthors of the candidate's supervisor. However, in all our tables, we show that excluding these control variables does not impact our findings.

7. While our PAP includes conducting an analysis on how the COVID-19 pandemic impacted the level of p-hacking in job market papers, it does not mention that this shock would also be used to examine potential channels leading to recruitment bias in academia.