An Audit Alternative: Measuring Employer Preferences and Beliefs without Deception

Judd B. Kessler, Corinne Low, and Colin D. Sullivan

July 23, 2018

What matters on the job market?

- Fundamental question in labor economics is how employers value different candidate characteristics, such as:
  - Human capital characteristics (education, field of study, experience, (e.g., Autor and Houseman [2010], Pallais [2014])
  - Gender and race (e.g., Altonji and Blank [1999])

- We need powerful tools to study these questions—obviously observational studies are insufficient

What matters on the job market?

- Fundamental question in labor economics is how employers value different candidate characteristics, such as:
  - Human capital characteristics (education, field of study, experience, (e.g., Autor and Houseman [2010], Pallais [2014])
  - Gender and race (e.g., Altonji and Blank [1999])

- We need powerful tools to study these questions—obviously observational studies are insufficient

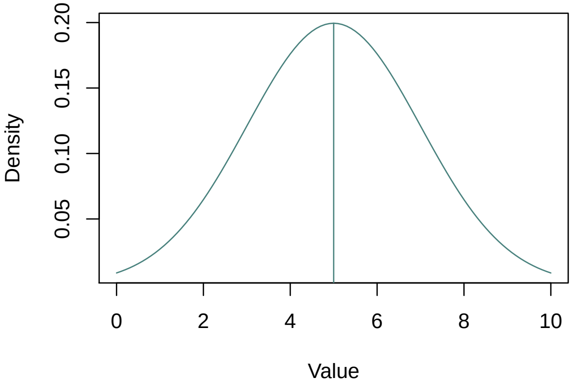
- Audit studies have been a workhorse in this literature
  - In-person (critiqued by Turner et al. [1991], Heckman and Siegelman [1992], Heckman [1998])
  - Correspondence and resume audits for discrimination (large literature launched by Bertrand and Mullainathan [2004])
  - Branched out into new areas (e.g., unemployment spells, Kroft et al. [2013], Eriksson and Rooth [2014], Nunley et al. [2017], value of for-profit-college degrees, Deming et al. [2016])

- Resume audit studies give you the difference in callback rates between groups

Callback indicates a candidate is above a threshold

- Imagine a distribution of employer i's expected productivity of candidate j with vector of characteristics X_j as in the below:

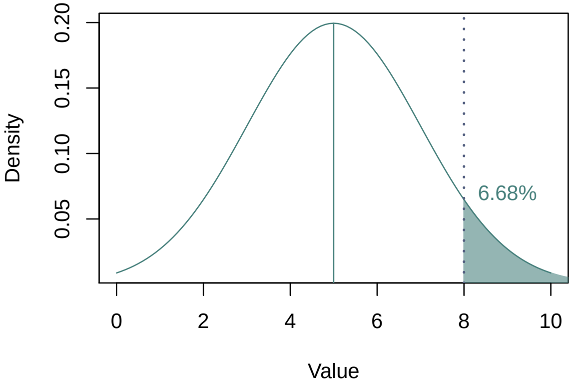
$$V_{ij} = \beta X_j + \xi_{ij},$$



Callback indicates a candidate is above a threshold

- What we observe in an audit study is an indicator for whether a candidate is called back:

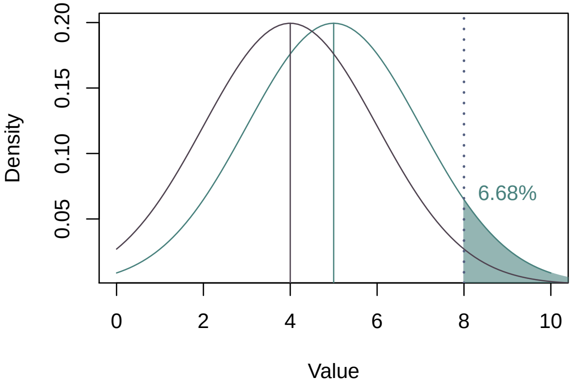
$$D_{ij} = \mathbb{1}[V_{ij} \geq V_i^*(c_i)]$$



Callback indicates a candidate is above a threshold

- We can compare callback rates of different groups. For simplicity, imagine a binary characteristic x_j:

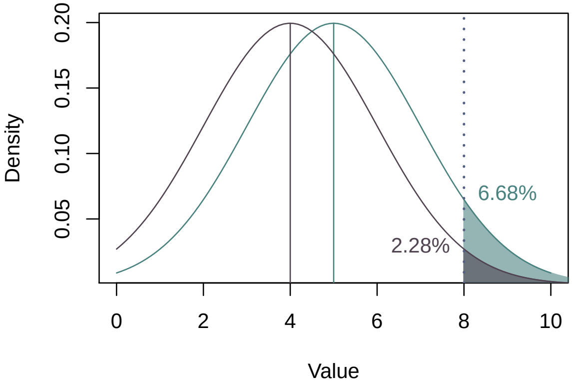
$$V_{ij} = \beta x_j + \xi_{ij},$$



Callback indicates a candidate is above a threshold

- Resume audit studies measure the impact of x_j on callback rate by estimating α as:

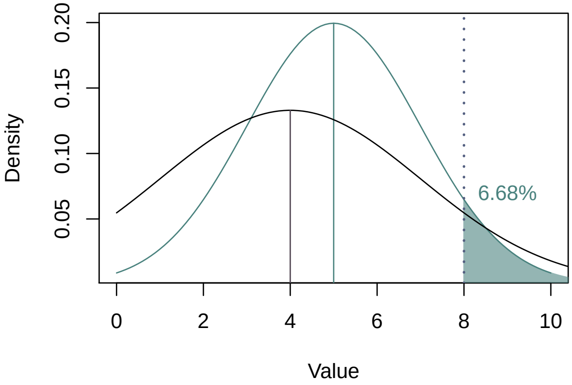
$$\alpha = E[D_{ij}|x_j = 1] - E[D_{ij}|x_j = 0]$$



Callback indicates a candidate is above a threshold

- Why might we be interested in richer information on V_{ij}?
- If the shape of the distribution depends on x_j, callback rates will not have consistent relationship across the distribution

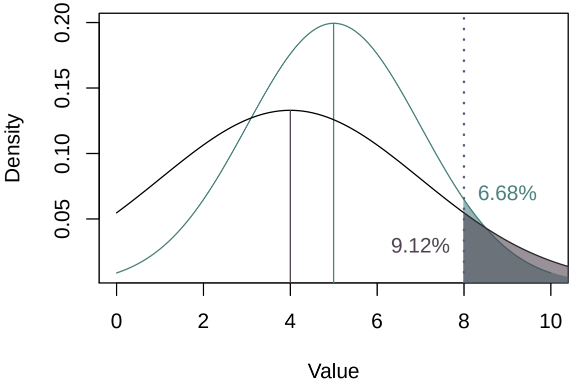
More



Callback indicates a candidate is above a threshold

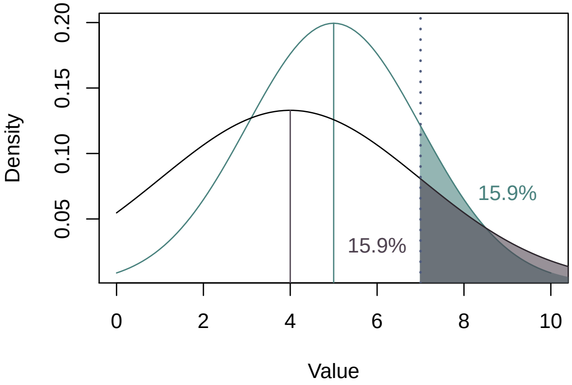
- Why might we be interested in richer information on V_{ij}?
- If the shape of the distribution depends on x_j, callback rates will not have consistent relationship across the distribution

More



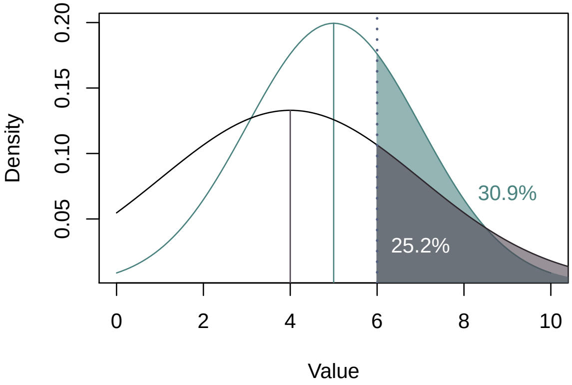
Callback indicates a candidate is above a threshold

- Why might we be interested in richer information on V_{ij}?
- If you change thresholds (e.g., selective hiring to economic expansion), the sign of α can flip



Callback indicates a candidate is above a threshold

- Why might we be interested in richer information on V_{ij}?
- If you change thresholds (e.g., selective hiring to economic expansion), the sign of α can flip



A new approach: Incentivized Resume Rating (IRR)

- Much richer information by being able to directly measure preferences
  - Parallel to buy / no-buy versus tracing demand curve (e.g., BDM)

A new approach: Incentivized Resume Rating (IRR)

- Much richer information by being able to directly measure preferences
  - Parallel to buy / no-buy versus tracing demand curve (e.g., BDM)
- How to do it in the hiring domain with incentives?
  - Employers rate **hypothetical resumes** with randomly assigned characteristics
  - They are matched with **real job seekers** according to their reported preferences
- Similar-in-spirit to design applied to dating markets in Low [2017]

A new approach: Incentivized Resume Rating (IRR)

- Much richer information by being able to directly measure preferences
  - Parallel to buy / no-buy versus tracing demand curve (e.g., BDM)
- How to do it in the hiring domain with incentives?
  - Employers rate **hypothetical resumes** with randomly assigned characteristics
  - They are matched with **real job seekers** according to their reported preferences
- Similar-in-spirit to design applied to dating markets in Low [2017]
- This offers the control of a laboratory experiment with the "stakes" of a field experiment
  - Independently randomize many characteristics
  - Get continuous measures of employer preferences
  - Have each employer rate multiple resumes
- Experimental paradigm is very flexible, and can be used to measure many different traits with different pools of employers and candidates

We study employer preferences for college students

- Traditionally hard to investigate preferences of "elite" employers because they do not accept cold resumes
- How they value human capital investments
  - College students spend three months a year outside of school; we explore the impact of their HC accumulation in those months
  - Investigate impact of quality (e.g., more prestigious internship) and quantity (e.g., an additional experience) of summer employment
  - Can compare these to impact of GPA, which we treat as a numeraire
- How they respond to demographics
  - On-campus recruiters may have different race and gender preferences than firms traditionally targeted in resume audit studies
  - We measure—for the first time—employers' beliefs about demographic groups' likelihood of job acceptance

Sample resume of graduating senior

Nathan Stewart

[illegible] • [illegible] • [illegible] • [illegible]

EDUCATION

University of Pennsylvania, College of Arts and Sciences **Philadelphia, PA**
BA in Economics **Expected May 2017**
Cumulative GPA: 3.82/4.00

WORK EXPERIENCE

Bank of America Merrill Lynch, New York, NY
Investment Banking Summer Analyst - Healthcare Finance **June - August 2016**

- Advised hospitals and healthcare systems on strategic financing options, new project initiatives, and M&A opportunities
- Prepared client pitchbooks and presentation materials for investor roadshows
- Conducted due diligence and filed reports on 103 deals (41% of group's total deals) as part of an SEC initiative (MCDC)

P.F. Chang's, Mclean, VA
Server **June - August 2015**

- Memorized entire menu and completed server training in five days
- Worked diligently under stressful conditions to deliver high quality service to customers
- Communicated and worked with servers, hosts, and bar staff to operate restaurant smoothly and uphold P.F. Chang's core values and principles

---

LEADERSHIP EXPERIENCE

MUSE - Undergraduate Marketing Club, Philadelphia, PA
Executive Board Member **2013-2015**

- Assisted in organizing speaker conferences, alumni panels, and networking sessions, with past

Incentivized Resume Rating: our design

- We partner with University of Pennsylvania Career Services
  - Collect hundreds of real Penn resumes to cull components
  - Use real Penn seniors interested in being matched as candidate pool
- Career Services offers employers the opportunity to try a new pilot tool designed by Wharton professors
  - Framed and marketed as a way to help employers find candidates
  - Only participation incentive is to be matched with Penn seniors

Incentivized Resume Rating: our design

- We partner with University of Pennsylvania Career Services
  - Collect hundreds of real Penn resumes to cull components
  - Use real Penn seniors interested in being matched as candidate pool
- Career Services offers employers the opportunity to try a new pilot tool designed by Wharton professors
  - Framed and marketed as a way to help employers find candidates
  - Only participation incentive is to be matched with Penn seniors
- Employers rate 40 resumes (median employer takes 28 minutes)
  - Choose majors to view: Humanities/Social Sciences or Science/Math
  - Rate candidates on: "desirability" and "likelihood of acceptance"
- We use ML to match each employer to 10 real seniors based on their preferences (i.e., no deception) and email their resumes
- We repeat the experiment at University of Pittsburgh to show differences based on subject pool

Instructions | Sample Recruitment Email

Rating on two dimensions

Prospect Forum, Philadelphia, PA
Executive VP **2013-2015**

- Launched early stage entrepreneurial venture with peers to improve the career search process for college students
- Connected students with Philadelphia-based companies that match their interests and worked to expand to other Ivies

SKILLS

Public speaking, marketing, writing, fundraising, data analysis, PowerPoint, Excel

How interested would you be in hiring **Nathan Stewart**?

Not interested 1	2	3	4	5	6	7	8	9	Very interested 10
○	○	○	○	○	○	○	○	○	○

How likely do you think **Nathan Stewart** would be to accept a job with your organization?

Not likely 1	2	3	4	5	6	7	8	9	Very likely 10
○	○	○	○	○	○	○	○	○	○

Next

Resume creation and variables

Component	Randomization
GPA	Drawn from $U(2.90, 4.00)$
Major	Drawn from a list of Penn majors
First job	Pr(**Top Internship**) $= \frac{1}{2}$
Second job	Pr(**Second Internship**) $= \frac{13}{40}$,
	Pr(**Work for Money**) $= \frac{13}{40}$,
	Pr(Blank)$= \frac{14}{40}$
Leadership	Two items drawn independently
Skills	Pr(**Technical skills**)=0.25

Resume creation and variables

Component	Randomization
GPA	Drawn from $U(2.90, 4.00)$
Major	Drawn from a list of Penn majors
First job	Pr(**Top Internship**) $= \frac{1}{2}$
Second job	Pr(**Second Internship**) $= \frac{13}{40}$, Pr(**Work for Money**) $= \frac{13}{40}$, Pr(Blank)$= \frac{14}{40}$
Leadership	Two items drawn independently
Skills	Pr(**Technical skills**)=0.25

Component	Randomization
Name	Pr(**Not White Male**)=67.2%, Gender (50% Male, 50% Female), Race drawn from U.S. distribution (65.7% White, 16.8% Hispanic, 12.6% Black, 4.9% Asian)

Regression specification

- Recall expected employer productivity, $V_{ij} = \beta X_j + \xi_{ij}$
- Mean value in OLS (averaged over the space we created):

$$V_{ij} = \beta_0 + \beta_1 \, GPA + \beta_2 \, TopInt + \beta_3 \, SecondInt + \beta_4 \, WFM + \beta_5 \, TechSkills + \beta_6 \, NotWhiteMale + \alpha_i + \gamma_j + \xi_{ij}$$

where α_i are rater fixed effects and γ_j includes leadership and major fixed effects

Regression specification

- Recall expected employer productivity, $V_{ij} = \beta X_j + \xi_{ij}$
- Mean value in OLS (averaged over the space we created):

$$V_{ij} = \beta_0 + \beta_1 \, GPA + \beta_2 \, TopInt + \beta_3 \, SecondInt + \beta_4 \, WFM + \beta_5 \, TechSkills + \beta_6 \, NotWhiteMale + \alpha_i + \gamma_j + \xi_{ij}$$

  where α_i are rater fixed effects and γ_j includes leadership and major fixed effects
- Will also run quantile specifications to estimate marginal effects at 25th, 50th, 75th, 90th, and 95th percentiles
- Will first present results on the first rating: "How interested would you be in hiring [name]"?

OLS results

	All
GPA	2.195***
	(0.129)
Top Internship	0.902***
	(0.0806)
Second Internship	0.463***
	(0.0947)
Work for Money	0.149
	(0.0913)
Technical Skills	-0.0680
	(0.0900)
Not White Male	-0.117
	(0.0842)
Observations	2880
F-test p-value for Majors	< 0.001
F-test p-value for Leadership	0.0649

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

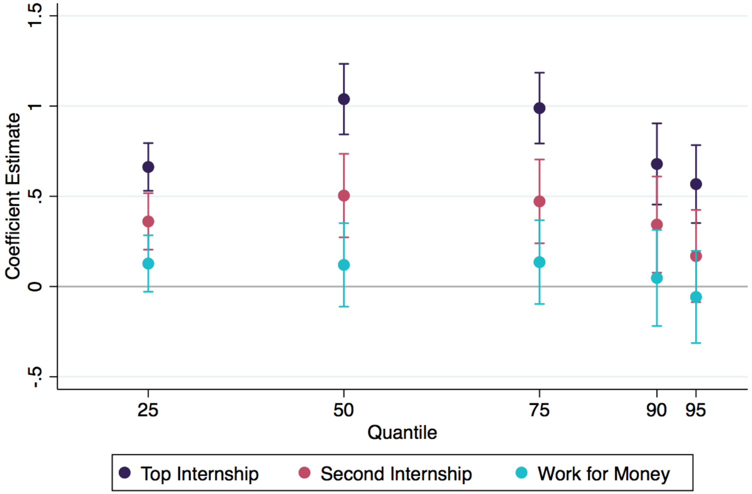
OLS results

	All	Humanities & Social Sciences	Science & Math
GPA	2.195***	2.300***	1.852***
	(0.129)	(0.153)	(0.243)
Top Internship	0.902***	1.039***	0.530***
	(0.0806)	(0.0944)	(0.173)
Second Internship	0.463***	0.514***	0.291
	(0.0947)	(0.114)	(0.187)
Work for Money	0.149	0.114	0.319*
	(0.0913)	(0.109)	(0.185)
Technical Skills	-0.0680	-0.0492	-0.171
	(0.0900)	(0.106)	(0.186)
Not White Male	-0.117	-0.0110	-0.399**
	(0.0842)	(0.0998)	(0.188)
Observations	2880	2040	840
F-test p-value for Majors	< 0.001	0.0036	< 0.001
F-test p-value for Leadership	0.0649	0.0246	< 0.001

Standard errors in parentheses

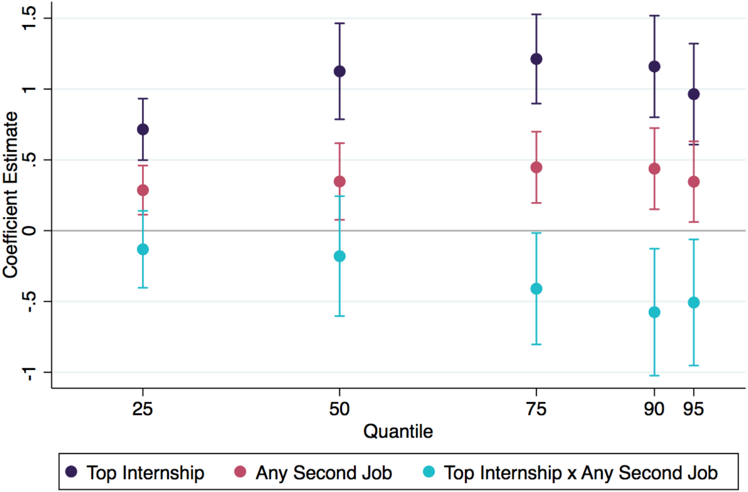
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Valuation of summer work experience



Bars indicate 95% Confidence Intervals.

Interactions between work experience

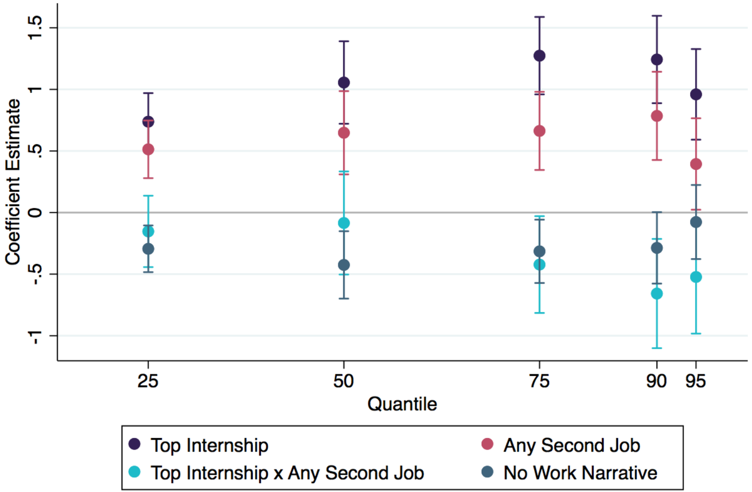


Bars indicate 95% Confidence Intervals.

Top Internship x Second Internship

Top Internship x Work for Money

Work experience narrative?



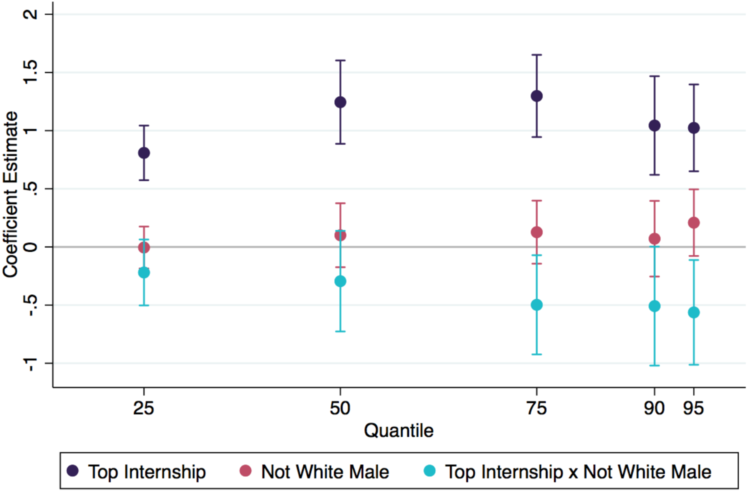
Bars indicate 95% Confidence Intervals.

Human capital matters, what about demographics?

- Have shown that firms value summer work experience
  - Both quality and quantity important—effects differ by quantile
  - Constraints students face in needing to earn money from summer work might be materially important
  - Interactions between different components, can be more closely examined with this design

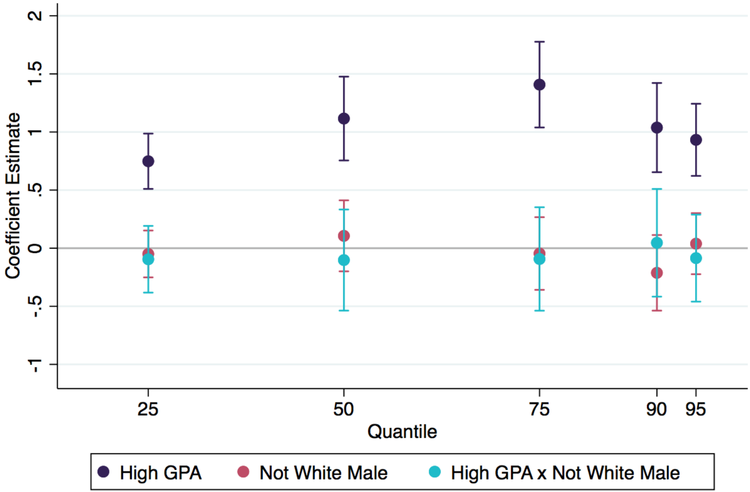
- Have also shown that firms recruiting in STEM are less interested in female/minority candidates
  - Will now examine impact of demographic characteristics more closely
  - In Bertrand and Mullainathan [2004], not only did resumes with black names receive fewer callbacks, there was also a lower return to quality improvements

Top Internship less valuable for women and minorities



Bars indicate 95% Confidence Intervals.

Effect absent for GPA



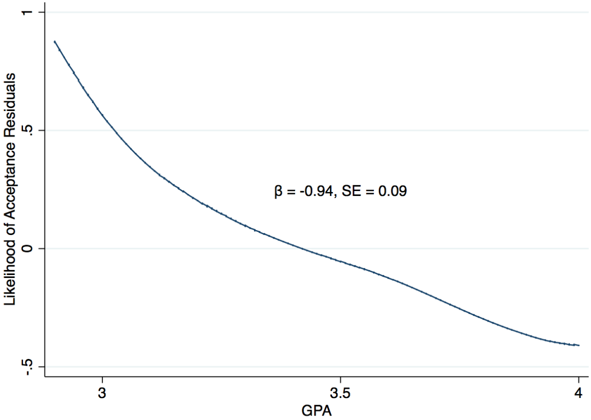
Bars indicate 95% Confidence Intervals.

Second measure: likelihood of acceptance

- Recall question: "How likely do you think [name] would be to accept a job with your organization?"
  - This is correlated positively with desirability rating
  - Holding desirability constant, negatively correlated with "objective" quality

Second measure: likelihood of acceptance

- Recall question: "How likely do you think [name] would be to accept a job with your organization?"
  - This is correlated positively with desirability rating
  - Holding desirability constant, negatively correlated with "objective" quality



Firms believe women and minorities are less likely to accept

	All
GPA	0.734***
	(0.120)
Top Internship	0.666***
	(0.0763)
Second Internship	0.393***
	(0.0910)
Work for Money	0.200**
	(0.0895)
Technical Skills	-0.105
	(0.0862)
Not White Male	-0.197**
	(0.0805)
Observations	2880

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

More

Firms believe women and minorities are less likely to accept

	All	Desirability < 5	Desirability ≥ 5
GPA	0.734***	-0.341**	-0.133
	(0.120)	(0.140)	(0.144)
Top Internship	0.666***	0.435***	0.0632
	(0.0763)	(0.0910)	(0.0880)
Second Internship	0.393***	0.293***	0.194*
	(0.0910)	(0.105)	(0.104)
Work for Money	0.200**	0.0895	0.136
	(0.0895)	(0.0991)	(0.106)
Technical Skills	-0.105	0.00508	-0.119
	(0.0862)	(0.0982)	(0.0962)
Not White Male	-0.197**	-0.0664	-0.208**
	(0.0805)	(0.0913)	(0.0919)
Observations	2880	1367	1513

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

More

Why does this matter?

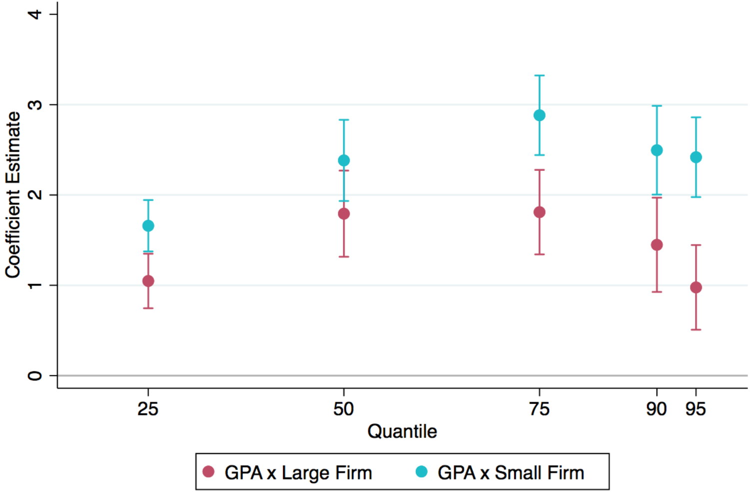
- Imagine the firm incurs costs to interview or recruit candidates (e.g., time/effort, limited slots)
- Could produce (or exacerbate) lower callback rates for under-represented groups
- Callback differences may reflect more than expected productivity
  - Essentially an omitted variable bias problem
  - But not solved with randomization, since appeal of trait and impact on likelihood of acceptance assigned simultaneously
  - Anything the firm finds appealing might also change their chance of "getting" candidate

Incentivized Resume Rating: future research opportunities

- IRR can be used to answer a wide array of human capital questions
- Can identify different dimensions of preferences
- Setup costs are substantial, but marginal costs of running are lower (we will gladly share our technology)
- Can be used outside of college setting
- Deployment with multiple groups possible for comparison

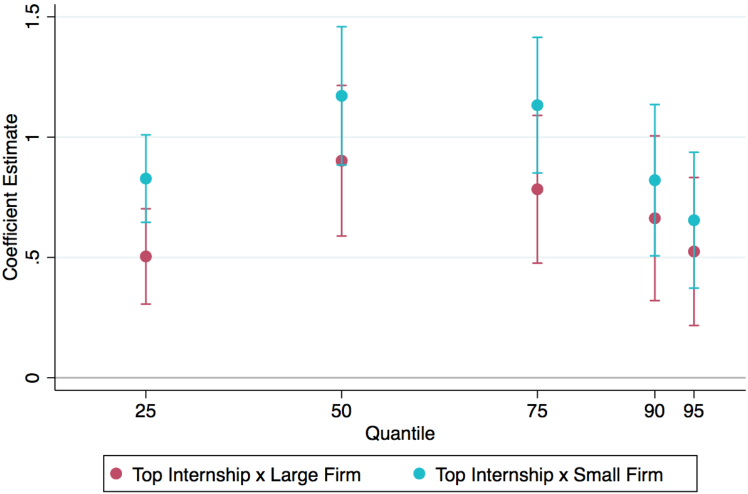
Conclusion

Firm Size & GPA



Bars indicate 95% Confidence Intervals.

Firm Size & Top Internship



Bars indicate 95% Confidence Intervals.

Results at Pitt directionally similar

	Penn	Pitt
GPA	2.195***	0.263**
	(0.129)	(0.113)
Top Internship	0.902***	0.222***
	(0.0806)	(0.0741)
Second Internship	0.463***	0.212**
	(0.0947)	(0.0844)
Work for Money	0.149	0.154*
	(0.0913)	(0.0807)
Technical Skills	-0.0680	0.107
	(0.0900)	(0.0768)
Not White Male	-0.117	0.00297
	(0.0842)	(0.0710)
Observations	2880	3440
F-test p-value for Majors	< 0.001	< 0.001
F-test p-value for Leadership	0.0649	0.937

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Summary

- In this paper, we introduce a new experimental paradigm, Incentivized Resume Rating, for measuring employers' preferences over candidate characteristics
- The key advantage is ability to elicit the full distribution of employer preferences
  - Estimate value of characteristics at different levels of selectivity
  - Independent randomization of many characteristics allows for analysis of conditional marginal effects
- Other benefits
  - Can access employers who don't respond to cold resumes
  - Can measure multiple dimensions driving employer callbacks
- We deploy IRR to investigate
  - Preferences of recruiters at elite colleges for student human capital investments
  - Impact of demographic characteristics, beyond current literature

Backup

The issue is that the variance of preferences could be correlated with underlying characteristic

- This is not a minor, statistical issue. In fact, it seems highly *likely* to occur in the resume setting
- To see this, let's write an expanded version of the error term

$$\xi_{ij} = \eta_i x_j + \Omega' \vec{Z}_j + \vec{\psi}_i \vec{Z}_j + \nu_{ij} \tag{1}$$

- η_i is firm-specific tastes for x_j, Ω' is common tastes for other characteristics, Z, and $\vec{\psi}_i$ is firm-specific tastes for other characteristics, with ν_{ij} representing match-specific quality

The issue is that the variance of preferences could be correlated with underlying characteristic

- This is not a minor, statistical issue. In fact, it seems highly *likely* to occur in the resume setting
- To see this, let's write an expanded version of the error term

$$\xi_{ij} = \eta_i x_j + \Omega' \vec{Z}_j + \vec{\psi}_i \vec{Z}_j + \nu_{ij} \tag{1}$$

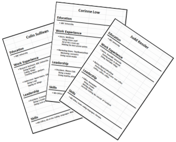
- η_i is firm-specific tastes for x_j, Ω' is common tastes for other characteristics, Z, and $\vec{\psi}_i$ is firm-specific tastes for other characteristics, with ν_{ij} representing match-specific quality
- Two natural reasons for heteroskedasticity:
  - Variance in firm specific tastes that's correlated with x_j
    - e.g., all firms feel equally about low GPAs, but some love high GPAs whereas some don't care. Or, for a binary, some firms love waitstaff jobs while some dislike them strongly
  - Variance in tastes for other outcome measures correlated with x_j
    - e.g., I don't even pay attention to work experience if GPA is too low, so variance is increasing in GPA

Back: Theory

Probability of Acceptance can be Correlated with X

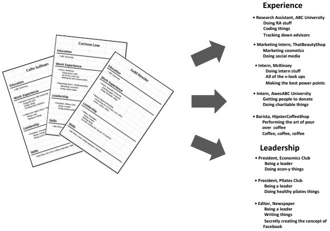
- Second issue. Imagine employers think there is some probability $\pi_{ij} \in [0,1]$ that a candidate will accept a position if offered
- But, costs incurred either way. Can write firm's utility as a function of: $\pi_{ij}(V_{ij} - V^* - c) - (1 - \pi_{ij})c$.
- If we imagine that firms hire whenever utility is positive, for example, require $\pi > \frac{c}{V_{ij} - V^*}$
- The problem comes in if the probability of acceptance, π_{ij}, is correlated with x_j. In this case, it can bias the sign of α
- Again, is this an esoteric statistical issue?
  - Anything the firm finds appealing might also decrease their chance of "getting" candidate
  - Interviews are obviously costly, or callback thresholds would not be so low
  - Picture top-coding in the economics job market
- This is essentially OVB, but that is not solved with random assignment, since the appeal of a trait and its impact on "gettability" are assigned simultaneously

Collect Sample Resumes



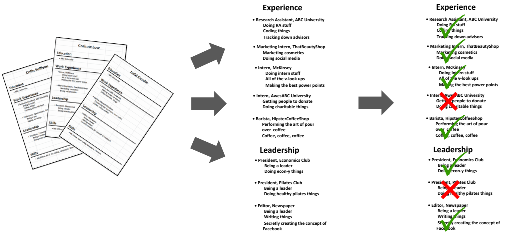
- Collected about 1200 resumes from graduating seniors
- Sought a variety of schools, majors, experiences, and skills

Extract Resume Elements



- Converted resumes into data using text parsing software
- Extracted major, skills, jobs and leadership experiences (organization, position, location, description)

Curate List of Elements



- Curated list of various experiences and skills
- Removed gender- or race-identifying information. E.g., "Women's Basketball," "South Asian Students Association"

Compile Elements into “Composite” Resumes

- Random name: conveys race and gender
- Random education elements: GPA, major
- Two randomized work experience blurbs
- Two randomized leadership blurbs and skills

Key Excerpt from Instructions

In this survey tool, you will be shown 40 resumes of hypothetical job candidates and asked to evaluate:
(a) how interested you would be in hiring the candidate; and
(b) how likely the candidate would be to accept a job at your organization

Importantly, when gauging your interest level in the candidate, imagine that the candidate is guaranteed to accept your job offer — think only about your perception of the candidate's quality.

When gauging how likely the candidate would be to accept a job, imagine that the candidate has been given a job — think only about whether you think the candidate would accept the job.

We will use both answers to recommend Penn students for you who we estimate to be strong candidates for your position(s).

Back: Incentivized Resume Rating

Sample Recruitment Email

From: upenn@csm.symplicity.com [mailto:upenn@csm.symplicity.com]
Sent: Tuesday, July 26, 2016 1:34 PM
To: [redacted]
Subject: Identify Top Penn Students for your Firm

Dear [redacted]

This year, Penn Career Services is participating in a pilot with two Wharton professors who are developing a new tool that can help you to identify potential job candidates from the University of Pennsylvania for post-graduate positions.

The tool is designed to identify top candidates for your open positions and provides you with those candidates' contact information and resumes so you can invite them to coffee chats, to info sessions, and to apply for a job at your organization. Since the tool uses data-driven methods to identify candidates, we see this as a useful complement to firms' existing methods for identifying promising candidates.

Completing the tool takes about 30 minutes and involves evaluating 40 hypothetical resumes. After evaluating these resumes, the tool uses a newly developed machine-learning algorithm to identify candidates who would be a particularly good fit for your job based on your evaluations. The Wharton professors will also use a completely anonymized version of your data to perform research on broader trends in what firms value in hiring, and they will be glad to share these insights with your company once the research is complete. To be provided with potential candidates for a position, at least one person from your firm must complete the tool. If possible, having multiple individuals participate will help increase the accuracy of the algorithm's recommendations. Additionally, if you are hiring for different positions within your organization, we recommend at least one person from your organization take the tool for each open position so you get a list of candidates tailored for each job opening. Rising Penn seniors will be invited to participate in the trial by submitting their resumes beginning on August 22nd, and we plan to have candidate recommendations to you by early September.

To take the tool, please click the link here:

https://wharton.qualtrics.com/SE/?SID=SV_3I3ohtNPn2R8c97

If you you would like to discuss more about how the tool could be useful for your firm, or have any questions, please contact the Wharton researchers: Judd B. Kessler (judd.kessler@wharton.upenn.edu) and Corinne Low (corlow@wharton.upenn.edu).

Sincerely,

Barbara Hewitt, Senior Associate Director, Career Services

Back: Incentivized Resume Rating

Breakdown by Race and Gender

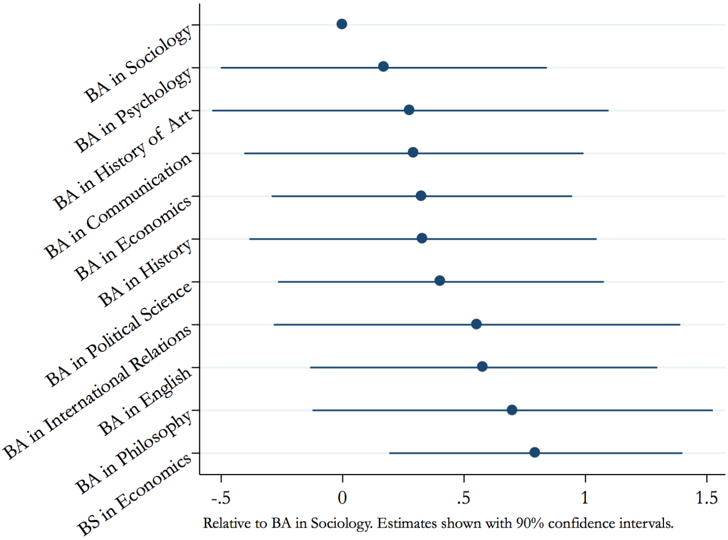
	All	Desirability < 5	Desirability ≥ 5
Female	-0.184**	-0.107	-0.248***
	(0.0747)	(0.0858)	(0.0843)
Not White	-0.0109	0.0172	-0.0293
	(0.0789)	(0.0891)	(0.0884)
Observations	2880	1367	1513

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

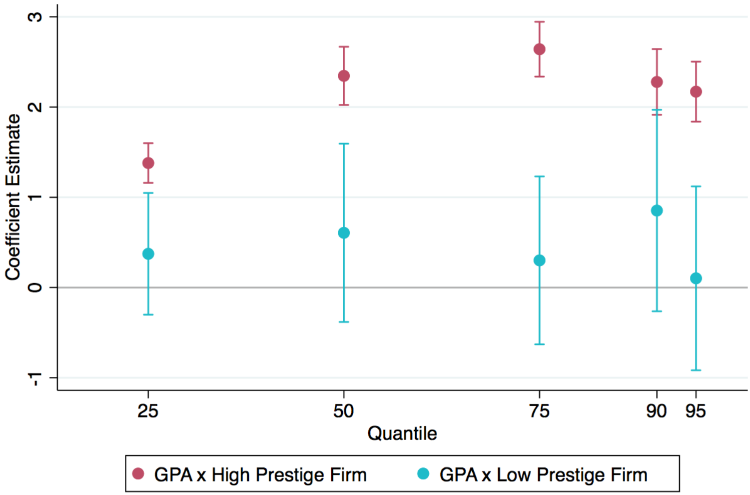
Back: Likelihood of Acceptance

Field of Study



Relative to BA in Sociology. Estimates shown with 90% confidence intervals.

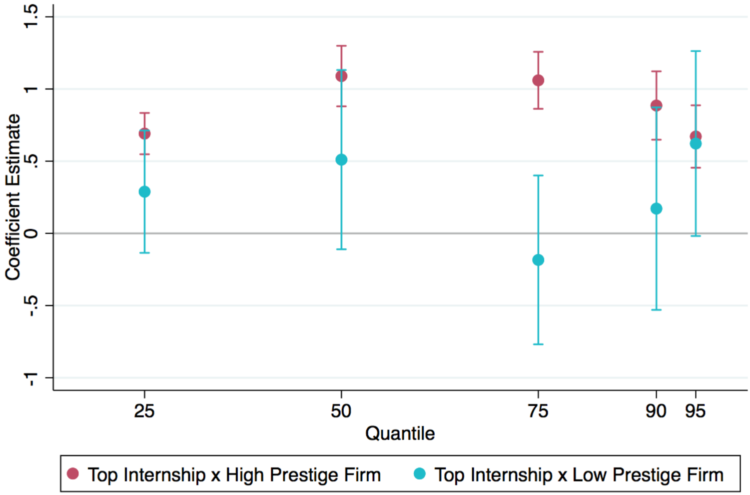
Firm Prestige & GPA



Bars indicate 95% Confidence Intervals.

Back: Firm Size & GPA

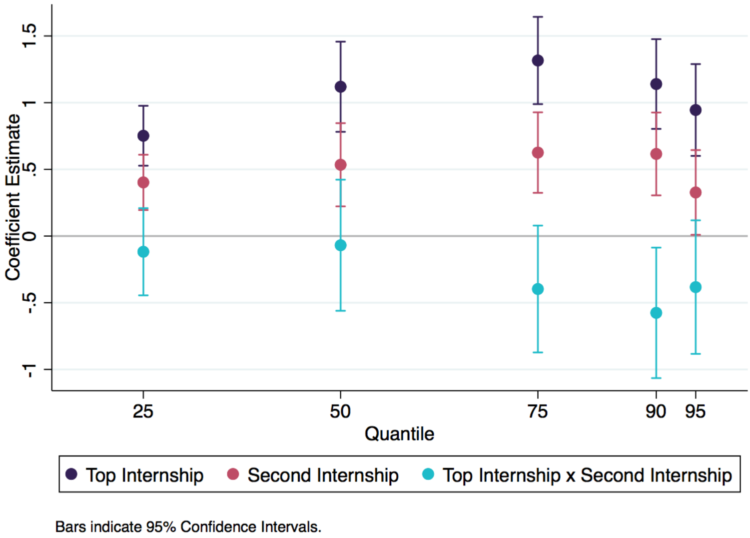
Firm Prestige & Top Internship



Bars indicate 95% Confidence Intervals.

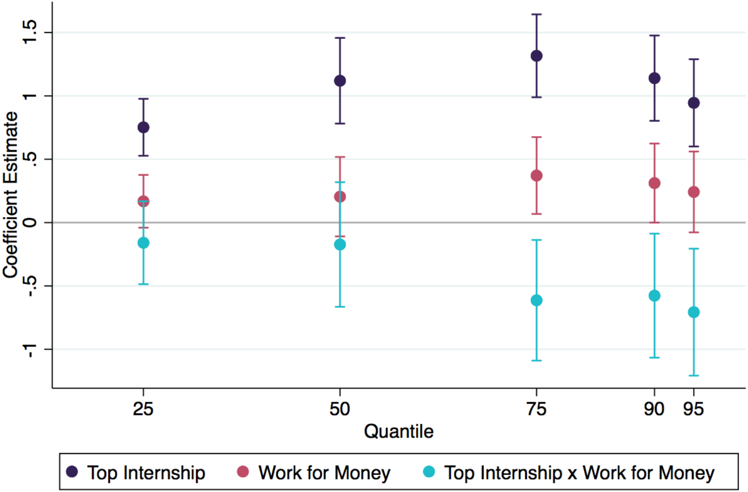
Back: Firm Size & GPA

Top Internship & Second Internship



Back: Top Internship x Any Second Job

Top Internship & Work for Money



Bars indicate 95% Confidence Intervals.

Back: Top Internship x Any Second Job

Reconciling our discrimination results with the literature

- Much of the prior literature, including *un*-incentivized resume rating studies and resume audit studies, shows evidence of discrimination

Reconciling our discrimination results with the literature

- Much of the prior literature, including *un*-incentivized resume rating studies and resume audit studies, shows evidence of discrimination
- So, why do we see a null effect on hiring interest? Possible explanations:
  - VERY different pool of employers than past literature
    - These firms may have *positive* tastes for diversity
    - Recruiters might also have implicit biases, which could cancel one another out

Reconciling our discrimination results with the literature

- Much of the prior literature, including *un*-incentivized resume rating studies and resume audit studies, shows evidence of discrimination
- So, why do we see a null effect on hiring interest? Possible explanations:
  - VERY different pool of employers than past literature
    - These firms may have *positive* tastes for diversity
    - Recruiters might also have implicit biases, which could cancel one another out
  - Social desirability bias?
    - Cannot self-censor if majority of bias is implicit
    - And there are indeed real stakes here, since they "waste" matches on candidates they're not interested in if they distort their preferences
  - Other features of our design?
    - e.g., respondents ignore the name data, because they think we won't use it for matches
    - Again, requires conscious awareness of bias
- In the next section, we will also show that the lower expected probability of acceptance we find can produce a lower callback rate for minority groups

Joseph G Altonji and Rebecca M Blank. Race and gender in the labor market. *Handbook of labor economics*, 3:3143–3259, 1999.

David H Autor and Susan N Houseman. Do temporary-help jobs improve labor market outcomes for low-skilled workers? evidence from" work first". *American Economic Journal: Applied Economics*, pages 96–128, 2010.

Marianne Bertrand and Sendhil Mullainathan. Are Emily and Greg more employable than Lakisha and Jamal? *The American Economic Review*, 94(4):991–1013, 2004.

David J. Deming, Noam Yuchtman, Amira Abulafi, Claudia Goldin, and Lawrence F. Katz. The value of postsecondary credentials in the labor market: An experimental study. *American Economic Review*, 106(3): 778–806, March 2016. doi: 10.1257/aer.20141757. URL http://www.aeaweb.org/articles?id=10.1257/aer.20141757.

Stefan Eriksson and Dan-Olof Rooth. Do employers use unemployment as a sorting criterion when hiring? evidence from a field experiment. *American Economic Review*, 104(3):1014–39, March 2014. doi: 10.1257/aer.104.3.1014. URL http://www.aeaweb.org/articles?id=10.1257/aer.104.3.1014.

James J. Heckman. Detecting discrimination. *Journal of Economic Perspectives*, 12(2):101–116, 1998.

James J. Heckman and Peter Siegelman. The Urban Institute audit studies: their methods and findings. In *Clear and convincing evidence: measurement of discrimination in America*, pages 187–258. Lanhan, MD: Urban Institute Press, 1992.

Kory Kroft, Fabian Lange, and Matthew J. Notowidigdo. Duration dependence and labor market conditions: Evidence from a field experiment*. *The Quarterly Journal of Economics*, 128(3):1123–1167, 2013. doi: 10.1093/qje/qjt015. URL http://dx.doi.org/10.1093/qje/qjt015.

Corinne Low. Pricing the biological clock: Reproductive capital on the us marriage market. *Manuscript, Wharton School of Business*, 2017.

John M. Nunley, Adam Pugh, Nicholas Romero, and R. Alan Seals. The effects of unemployment and underemployment on employment opportunities: Results from a correspondence audit of the labor market for college graduates. *ILR Review*, 70(3):642–669, 2017. doi: 10.1177/0019793916654686. URL https://doi.org/10.1177/0019793916654686.

Amanda Pallais. Inefficient hiring in entry-level labor markets. *American Economic Review*, 104(11):3565–99, 2014.

Margery Turner, Michael Fix, and Raymond J. Struyk. Opportunities denied, opportunities, diminished: racial discrimination in hiring. *Washington, DC: Urban Institute Press*, 1991.