Discrimination: Empirics

Heidi L. Williams

MIT 14.662

Spring 2015

Testing for evidence of discrimination

- Version 1.0: "unexplained" or residual male/female wage gap
  - Difficult to control for all relevant characteristics
  - Very indirect test of discrimination
  - Some covariates are potentially endogenous (education)
- Version 2.0: alternative methods
  - Audit studies
  - Quasi-experiments
  - Tests of equilibrium predictions

Roadmap for today

- Regression analysis
  - Goldberger (1984)
  - Neal and Johnson (1996)
- Audit studies
  - Overview by Riach and Rich (2002)
  - Bertrand and Mullainathan (2004)
- Quasi-experiments
  - Goldin and Rouse (2000)
  - Anwar, Bayer, and Hjalmarsson (2012)
- Testing models
  - Charles and Guryan (2008)
  - Chandra and Staiger (2010)

1. Regression analysis
   - Goldberger (1984)
   - Neal and Johnson (1996)
2. Audit studies
   - Bertrand and Mullainathan (2004)
3. Quasi-experiments
   - Goldin and Rouse (2000)
   - Anwar, Bayer, and Hjalmarsson (2012)
4. Testing models
   - Charles and Guryan (2008)
   - Chandra and Staiger (2010)
5. Looking ahead

Direct regression

Are men paid more than equally productive women?

Suppose that the conditional expectation of earnings given qualifications and gender is given by:

$$E(y|x,z) \quad = \quad b'\mathbf{x} + az$$

- y: earnings
- $\mathbf{x} = (x_1, x_2, ..., x_k)'$: vector of productivity qualifications
- z: gender indicator
  - $z = 1$ for men, $z = 0$ for women
- coefficient a: discriminatory premium paid to men

Direct regression

- Commonly estimated in both the academic literature (*e.g.* Oaxaca (1973)) and in discrimination-related law suits
- Usual finding: $a > 0$
  - Often interpreted as evidence of salary discrimination
  - Among men and women with equal $\mathbf{x}$, men are paid more
- Usual concern: omitted productivity-relevant characteristics
  - If $cov(z, \epsilon|x) > 0$ then expect upward bias

Reverse regression

Are men less qualified than equally paid women?

$$E(q|y,z) \quad = \quad c^* y + d^* z$$

- $q = b'\mathbf{x}$: scalar index of qualifications
- coefficient d^*: excess qualifications of men for same salary
  - $d^* < 0$: evidence of salary discrimination in favor of men
  - Among men and women with equal y, men less qualified

Do direct and reverse regressions provide similar estimates?

- If men are paid more than equally qualified women, they should be less qualified than equally paid women
  - $a > 0$ should imply that $d^* < 0$
- However: this reasoning relies on a deterministic relationship
  - $y = b'\mathbf{x} + az = q + az$ implies $q = y - az$
  - Likely not true empirically

Direct and reverse regression: Conflicting estimates

- In practice, often give conflicting results
- Example: 1976 U-Illinois study of male/female faculty salaries
  - Males paid \$2,000 more than females with same # publications
  - Females publish 2 fewer articles than males with same salary
  - Implies both a and d^* are positive
- In general, reverse regression suggests lower estimate of salary discrimination (in favor of men) than direct regression
  - Reverse regression often suggests reverse discrimination

Direct and reverse regression: Goldberger (1984)

- Goldberger paper very clearly written, but short on intuition
- Common notion at the time: direct biased, reverse unbiased
- Two alternative models for single qualification case:
  1. Model #1: errors in variables
     - ★ Direct regression estimate upward-biased
     - ★ Reverse regression estimate downward-biased
     - ★ Direct and reverse regression bound true parameter value
  2. Model #2: proxy variable
     - ★ Direct regression estimate unbiased
     - ★ Reverse regression estimate downward-biased (may be of the wrong sign)

Take-away from Goldberger (1984)

- **Take-away**: Without knowing the underlying data generating process there is no sense in which either the direct regression approach or the reverse regression approach is *a priori* more "correct"
- In general, both direct and reverse regression approaches are somewhat "out of style"
- One exception: Neal and Johnson (1996)

1. Regression analysis
   - Goldberger (1984)
   - Neal and Johnson (1996)
2. Audit studies
   - Bertrand and Mullainathan (2004)
3. Quasi-experiments
   - Goldin and Rouse (2000)
   - Anwar, Bayer, and Hjalmarsson (2012)
4. Testing models
   - Charles and Guryan (2008)
   - Chandra and Staiger (2010)
5. Looking ahead

Neal and Johnson (1996)

How much of the black-white earnings gap is explained by differences in skills acquired prior to labor market entry?

- National Longitudinal Survey of Youth (NLSY) data
- Examine black-white wage gaps among workers in their late twenties as a function of AFQT score at age 18 or younger

Neal and Johnson (1996): Table 1

- Column (3): Adds linear and quadratic variables for AFQT
- Explains $\sim \frac{3}{4}$ of racial wage gap for young men

TABLE 1

LOG WAGE REGRESSIONS BY SEX

	MEN (N = 1,593)			WOMEN (N = 1,446)		
	(1)	(2)	(3)	(4)	(5)	(6)
Black	−.244	−.196	−.072	−.185	−.155	.035
	(.026)	(.025)	(.027)	(.029)	(.027)	(.031)
Hispanic	−.113	−.045	.005	−.028	.057	.145
	(.030)	(.029)	(.030)	(.033)	(.031)	(.032)
Age	.048	.046	.040	.010	.009	.023
	(.014)	(.013)	(.013)	(.015)	(.014)	(.015)
AFQT	...	...	.172	...	...	.228
			(.012)			(.015)
$AFQT^2$	...	...	−.013	...	...	.013
			(.011)			(.013)
High grade by 1991	...	.061	...	...	.088	...
		(.005)			(.005)	
R^2	.059	.155	.168	.029	.191	.165

NOTE.—The dependent variable is the log of hourly wages. The wage observations come from 1990 and 1991. All wages are measured in 1991 dollars. If a person works in both years, the wage is measured as the average of the two wage observations. Wage observations below $1.00 per hour or above $75 are eliminated from the data. The sample consists of the NLSY cross-section sample plus the supplemental samples of blacks and Hispanics. Respondents who did not take the ASVAB test are eliminated from the sample. Further, 163 respondents are eliminated because the records document a problem with their test. All respondents were born after 1961. Standard errors are in parentheses.

Is the AFQT racially biased?

1991 National Academy of Sciences (NAS) report

- Exhaustive study with the Department of Defense
- Focused on validity of the AFQT
- Special emphasis on racial fairness of the test
- No evidence AFQT under-predicts performance of blacks

Do blacks underinvest in skill because the return is lower?

Models of statistical discrimination (Lundberg-Startz 1983)

- Payoff to skill lower for blacks $\Rightarrow$ skill differences could reflect anticipation that returns from acquiring skills will be low
- Intuitive, but difficult to test
- Imperfect test: do returns to AFQT differ by race?

Neal and Johnson (1996): Table 2

- Can't reject that returns to skill are equal for blacks and whites
- But, problematic test: AFQT score an endogenous investment
- Ideally would have an instrument here

TABLE 2

TESTING FOR RACIAL DIFFERENCES IN THE RETURN TO AFQT: MEN

	All Races (N = 1,593) (1)	White (N = 825) (2)	Black (N = 466) (3)	Hispanic (N = 302) (4)
Black	−.107	...	...	...
	(.033)			
Hispanic	.003	...	...	...
	(.029)			
Age	.038	.052	.047	−.014
	(.013)	(.017)	(.025)	(.035)
AFQT	.172	.183	.208	.124
	(.015)	(.017)	(.031)	(.031)
AFQT^2	−.023	−.018	.031	−.066
	(.013)	(.015)	(.025)	(.031)
Black × AFQT	.037	...	...	...
	(.031)			
Black × AFQT^2	.056	...	...	...
	(.028)			
R^2	.170	.155	.129	.074

NOTE.—The "all races" sample includes all men from the sample described in table 1. All respondents were born after 1961. Standard errors are in parentheses.

What about labor market dropouts?

Neal and Johnson present estimates from two approaches:

- Median regressions
- Smith-Welch (1986) method

Doesn't hugely change conclusions

Determinants of AFQT scores

- Tables 5, 6: Large raw gap, significantly reduced by covariates
- Although sizable gaps remain, these results suggest "pre-market" factors may explain much of AFQT gap
- Results cast doubt on (*very* controversial) Herrnstein-Murray (1994) argument that AFQT measures inherent ability
  - Estimated racial gaps in scores are larger for older cohorts
  - Schooling increases AFQT scores (QOB instruments)

Take-aways from Neal and Johnson (1996)

- Very influential
- Focus solely on market discrimination is likely misplaced
- Suggests that some attention should be focused on understanding sources of large observed skill gaps between blacks and whites

1. Regression analysis
   - Goldberger (1984)
   - Neal and Johnson (1996)
2. Audit studies
   - Bertrand and Mullainathan (2004)
3. Quasi-experiments
   - Goldin and Rouse (2000)
   - Anwar, Bayer, and Hjalmarsson (2012)
4. Testing models
   - Charles and Guryan (2008)
   - Chandra and Staiger (2010)
5. Looking ahead

Audit studies: Overview

- *Long* literature ($>$ four decades old) has tested for evidence of discrimination in labor, housing, and product markets by conducting 'audit' field experiments
- Useful overview: Riach and Rich (2002)
- Two types of audit experiments:
  1. Audit tester studies
  2. Audit resume studies
- Conclusion of Riach and Rich: "...demonstrated pervasive and enduring discrimination against non-whites and women"

Audit studies: Criticisms

- Famously criticized by Heckman-Siegelman (1992)
  - Effectiveness of matched process
  - Unconscious bias
  - Small samples
- Despite these problems: results often quite compelling
- Audit resume studies can overcome many of these limitations

Bertrand and Mullainathan (2004)

- Well-known audit resume study
- Sent 5,000 resumes to help-wanted ads in Boston and Chicago
- Randomized otherwise equivalent resumes to have African-American or White sounding names: Emily Walsh or Greg Baker relative to Lakisha Washington or Jamal Jones
- Also experimentally vary credentials

Bertrand and Mullainathan (2004): Table 1

- Measured interview callbacks from each resume: 50% gap

TABLE 1—MEAN CALLBACK RATES BY RACIAL SOUNDINGNESS OF NAMES

	Percent callback for White names	Percent callback for African-American names	Ratio	Percent difference (p-value)
Sample:				
All sent resumes	9.65	6.45	1.50	3.20
	[2,435]	[2,435]		(0.0000)
Chicago	8.06	5.40	1.49	2.66
	[1,352]	[1,352]		(0.0057)
Boston	11.63	7.76	1.50	4.05
	[1,083]	[1,083]		(0.0023)
Females	9.89	6.63	1.49	3.26
	[1,860]	[1,886]		(0.0003)
Females in administrative jobs	10.46	6.55	1.60	3.91
	[1,358]	[1,359]		(0.0003)
Females in sales jobs	8.37	6.83	1.22	1.54
	[502]	[527]		(0.3523)
Males	8.87	5.83	1.52	3.04
	[575]	[549]		(0.0513)

Notes: The table reports, for the entire sample and different subsamples of sent resumes, the callback rates for applicants with a White-sounding name (column 1) an an African-American-sounding name (column 2), as well as the ratio (column 3) and difference (column 4) of these callback rates. In brackets in each cell is the number of resumes sent in that cell. Column 4 also reports the p-value for a test of proportion testing the null hypothesis that the callback rates are equal across racial groups.

Bertrand and Mullainathan (2004): Table 4

- Returns to higher-quality resume appear lower for African-Americans

TABLE 4—AVERAGE CALLBACK RATES BY RACIAL SOUNDINGNESS OF NAMES AND RESUME QUALITY

	Low	High	Ratio	Difference (p-value)
Panel A: Subjective Measure of Quality (Percent Callback)				
White names	8.50	10.79	1.27	2.29
	[1,212]	[1,223]		(0.0557)
African-American names	6.19	6.70	1.08	0.51
	[1,212]	[1,223]		(0.6084)
Panel B: Predicted Measure of Quality (Percent Callback)				
White names	7.18	13.60	1.89	6.42
	[822]	[816]		(0.0000)
African-American names	5.37	8.60	1.60	3.23
	[819]	[814]		(0.0104)

Notes: Panel A reports the mean callback percents for applicant with a White name (row 1) and African-American name (row 2) depending on whether the resume was subjectively qualified as a lower quality or higher quality. In brackets is the number of resumes sent for each race/quality group. The last column reports the p-value of a test of proportion testing the null hypothesis that the callback rates are equal across quality groups within each racial group. For Panel B, we use a third of the sample to estimate a probit regression of the callback dummy on the set of resume characteristics as displayed in Table 3. We further control for a sex dummy, a city dummy, six occupation dummies, and a vector of dummy variables for job requirements as listed in the employment ad (see Section III, subsection D, for details). We then use the estimated coefficients on the set of resume characteristics to estimate a predicted callback for the remaining resumes (two-thirds of the sample). We call "high-quality" resumes the resumes that rank above the median predicted callback and "low-quality" resumes the resumes that rank below the median predicted callback. In brackets is the number of resumes sent for each race/quality group. The last column reports the p-value of a test of proportion testing the null hypothesis that the callback percents are equal across quality groups within each racial group.

Bertrand and Mullainathan (2004): Discussion

- Manipulating perceptions of social class, not just race?
  - Birth certificate data on mother's education for first names
  - Little relationship between SES and name-specific callback rates
- Taste-based or statistical discrimination?
  - Argue neither model fits data especially well
- Randomization essentially assumes random search

Related paper: Fryer and Levitt (2004)

- Investigate relationship between Black names and life outcomes, controlling for background characteristics
- No compelling evidence of a relationship
- Reconciling this result with Bertrand-Mullainathan:
  1. Black names used as signals of race by discriminatory employers at resume stage, but unimportant later
  2. Black names provide useful signal to employers about labor market productivity conditional on resume information
  3. Black names have causal impact on job callbacks that Fryer and Levitt are unable to detect

1. Regression analysis
   - Goldberger (1984)
   - Neal and Johnson (1996)
2. Audit studies
   - Bertrand and Mullainathan (2004)
3. Quasi-experiments
   - Goldin and Rouse (2000)
   - Anwar, Bayer, and Hjalmarsson (2012)
4. Testing models
   - Charles and Guryan (2008)
   - Chandra and Staiger (2010)
5. Looking ahead

Goldin and Rouse (2000)

- US symphony orchestras long conducted non-blind auditions
- Over time, some began using screens to hide performers
- Over time, notable increase in share female
- Historically, many viewed women as unsuitable for orchestras
  - *"I just don't think women should be in an orchestra"*
  - *"women are more temperamental and more likely to demand special attention or treatment"*
  - *"the more women [in an orchestra], the poorer the sound"*
  - Some European orchestras continue (as of 2000) to have stated policies not to hire women
- Can blind auditions eliminate discrimination?

Data and empirical framework

- Collect audition records from major symphony orchestras
- Examine blind auditions in differences-in-differences framework
- Compare individuals in blind and non-blind auditions (FE)

Goldin and Rouse (2000): Table 7

Table 7: estimates for 3 orchestras that changed policies

- Less precise than other estimates, but same conclusions
- Without individual FE: compositional change
- With individual FE: blind auditions help women

TABLE 7—LINEAR PROBABILITY ESTIMATES OF THE LIKELIHOOD OF BEING ADVANCED: WITH INDIVIDUAL AND ORCHESTRA FIXED EFFECTS

	Include individual fixed effects		Exclude individual fixed effects
	(1)	(2)	(3)
Blind	0.404	0.399	0.103
	(0.027)	(0.027)	(0.018)
Female × Blind	0.044	0.041	−0.069
	(0.039)	(0.039)	(0.022)
Female			−0.005
			(0.019)
p-value of H_0: Blind + (Female × Blind) = 0	0.000	0.000	0.090
Individual fixed effects?	Yes	Yes	No
Orchestra fixed effects?	No	Yes	Yes
Year fixed effects?	Yes	Yes	Yes
Other covariates?	Yes	Yes	Yes
R^2	0.615	0.615	0.048
Number of observations	8,159	8,159	8,159

Notes: The unit of observation is a person-round. The dependent variable is 1 if the person is advanced to the next round and 0 if not. Standard errors are in parentheses. All specifications include an interaction for the sex being missing and a blind audition; "Other covariates" include automatic placement, years since last audition, number of auditions attended, size of the audition round, proportion female in audition round, whether a principal or substitute position, and a dummy indicating whether years since last audition is missing. These regressions include only the orchestras that changed their audition policy during our sample years and for which we observe individuals auditioning for the audition round both before and after the policy change. These regressions include 4,836 separate persons and are identified off of 1,776 person-rounds comprised of individuals who auditioned both before and after the policy change for a particular orchestra.
Source: Eight-orchestra audition sample (three orchestras of which are used; see Notes). See text.

Goldin and Rouse (2000): Discussion

- Headline estimate: blind additions increase relative probability that women advance from preliminary round by 50%
  - In general, results are quite noisy
  - One puzzling result for semi-final rounds
- Suggests blind auditions reduced discrimination against women and can explain a large share of the time-series increase in the share female of orchestras since 1970
- Can't distinguish between taste-based, statistical
- Can't examine whether performance affected by screen
- Writing this now, would include event study graphs

1. Regression analysis
   - Goldberger (1984)
   - Neal and Johnson (1996)
2. Audit studies
   - Bertrand and Mullainathan (2004)
3. Quasi-experiments
   - Goldin and Rouse (2000)
   - Anwar, Bayer, and Hjalmarsson (2012)
4. Testing models
   - Charles and Guryan (2008)
   - Chandra and Staiger (2010)
5. Looking ahead

Anwar, Bayer, and Hjalmarsson (2012)

- Examine the impact of jury racial composition on trial outcomes using data on felony trials in FL from 2000-2010
- Exploit day-to-day variation in the composition of the jury pool to isolate quasi-random variation in the composition of the seated jury

Anwar, Bayer, and Hjalmarsson (2012): Table 2

- Composition of jury pool appears uncorrelated with characteristics of the defendant and case

TABLE II

THE RELATIONSHIP BETWEEN THE RACIAL COMPOSITION OF THE JURY POOL AND DEFENDANT/CASE CHARACTERISTICS

	(1) Indicator for any blacks in pool	(2) Proportion of blacks in pool	(3) Proportion of whites in pool	(4) Proportion of other races in pool
Defendant characteristics				
Black	−0.008	0.003	−0.004	0.001
	[0.039]	[0.003]	[0.005]	[0.003]
Hispanic	0.005	0.004	−0.003	−0.001
	[0.088]	[0.008]	[0.011]	[0.006]
Male	0.043	0.006	−0.009	0.002
	[0.067]	[0.005]	[0.007]	[0.004]
Case characteristics				
Any drug charge	−0.029	−0.0003	0.004	−0.003
	[0.051]	[0.004]	[0.006]	[0.004]
Any murder charge	0.093	−0.002	−0.006	0.006
	[0.076]	[0.006]	[0.008]	[0.005]
Any other charge	0.007	0.002	−0.004	−0.0005
	[0.040]	[0.004]	[0.005]	[0.003]

Anwar, Bayer, and Hjalmarsson (2012): Table 4

- Large racial gap (16pp) in conviction rates when no blacks in jury pool
- ≥ 1 black member in jury pool eliminates this gap
- White conviction rates sharply higher with ≥ 1 black member

TABLE IV

REDUCED-FORM BENCHMARK REGRESSIONS

	(1)	(2)	(3)	(4)
Dependent variable	Any guilty conviction		Proportion guilty convictions	
Black defendant	0.150***	0.164***	0.156***	0.160***
	[0.056]	[0.058]	[0.055]	[0.057]
Any black in pool	0.069	0.105**	0.063	0.090*
	[0.048]	[0.051]	[0.047]	[0.050]
Black defendant * any black in pool	−0.168**	−0.166**	−0.174**	−0.155**
	[0.070]	[0.074]	[0.069]	[0.072]
Constant	0.656***	0.627***	0.600***	0.576***
	[0.039]	[0.041]	[0.038]	[0.040]
Includes controls for:				
Gender/age of pool	No	Yes	No	Yes
County dummy	No	Yes	No	Yes
Year of filing dummies	No	Yes	No	Yes
Observations	712	712	712	712
R-squared	0.01	0.07	0.01	0.08

Anwar, Bayer, and Hjalmarsson (2012): Discussion

- Headline estimate: racial gap in conviction rates is entirely eliminated when the jury pool includes at least one black member
- Don't estimate IV (argue exclusion restriction isn't plausible)
  - First stage is 0.40
- Note: broader law/economics literature + data

1. Regression analysis
   - Goldberger (1984)
   - Neal and Johnson (1996)
2. Audit studies
   - Bertrand and Mullainathan (2004)
3. Quasi-experiments
   - Goldin and Rouse (2000)
   - Anwar, Bayer, and Hjalmarsson (2012)
4. Testing models
   - Charles and Guryan (2008)
   - Chandra and Staiger (2010)
5. Looking ahead

Testing models

- Most papers documenting evidence of discrimination can't distinguish between taste-based and statistical models of discrimination
- Recent papers speaking more closely to theory:
  - Testing taste-based: Charles-Guryan (2008)
  - Testing statistical: Altonji-Pierret (2001)
  - Testing between models: List (2004), Chandra-Staiger (2010)

1. Regression analysis
   - Goldberger (1984)
   - Neal and Johnson (1996)
2. Audit studies
   - Bertrand and Mullainathan (2004)
3. Quasi-experiments
   - Goldin and Rouse (2000)
   - Anwar, Bayer, and Hjalmarsson (2012)
4. Testing models
   - Charles and Guryan (2008)
   - Chandra and Staiger (2010)
5. Looking ahead

Charles and Guryan (2008)

Tests key predictions of Becker taste-based discrimination model

- Combine 'standard' measures of CPS residual wage gap with 'direct' measures of prejudice from General Social Survey
- Although not definitive, results are supportive of Becker model

Charles and Guryan (2008): Table 3

- Prejudice of the 'marginal' white more strongly predictive of racial wage gaps than is the average prejudice

TABLE 3
ESTIMATED RELATIONSHIP BETWEEN RACIAL PREJUDICE OF WHITES IN A LABOR MARKET AND BLACK-WHITE RELATIVE WAGES
Dependent Variable: Residual Black-White Wage Gap in Market

Measure of Prejudice among All Whites	(1)	(2)	(3)	(4)	(5)	(6)
Average	−.036		.097	.050		
	(.030)		(.029)	(.033)		
Marginal		−.213	−.328	−.202		
		(.040)	(.050)	(.068)		
10th percentile					−.212	−.292
					(.180)	(.125)
Median					−.006	.007
					(.062)	(.043)
90th percentile					.016	.016
					(.029)	(.020)
Fraction black				−.157		−.304
				(.062)		(.045)
State	45	45	45	45	45	45
R^2	.03	.40	.52	.59	.05	.56

NOTE.—The table reports coefficients (standard errors) from OLS regressions of residual state-level black-white wage gaps on various measures of prejudice among all whites (the mean of the black-white wage gap across states is −0.123, and the standard deviation is .044). Residual black-white wage gaps are estimated using 1977–2002 May/MORG CPS data and control for education, a quadratic in experience, race-specific year effects, and state effects. Data from 1973–76 are dropped because the CPS reports states in groups in those years. States are dropped if they are not sampled in the GSS in the years necessary to measure the marginal index of prejudice. The "marginal" is the pth percentile of the prejudice distribution of the relevant population of whites, where p is the fraction of the population that is black. See the text for details.

1. Regression analysis
   - Goldberger (1984)
   - Neal and Johnson (1996)
2. Audit studies
   - Bertrand and Mullainathan (2004)
3. Quasi-experiments
   - Goldin and Rouse (2000)
   - Anwar, Bayer, and Hjalmarsson (2012)
4. Testing models
   - Charles and Guryan (2008)
   - Chandra and Staiger (2010)
5. Looking ahead

Health care: Chandra and Staiger (2010)

- Gigantic literature documenting evidence of disparities in health care treatment and health outcomes
- Taste-based: providers use higher benefit threshold for providing care to minority patients
  - Implies that returns to the marginal minority patient receiving treatment will be higher than the returns to the marginal non-minority patient receiving treatment
- Statistical: minorities may have lower benefit from treatment
- In both models, minorities receive less treatment, but statistical implies "under-treatment" may be optimal

Chandra and Staiger (2010): Key test

- With prejudice, treatment-on-the-treated effect should be larger for minorities (conditional on propensity to be treated)
- Similar in spirit to Knowles, Persico, and Todd (2001), who analyze racial bias in motor vehicle searches

Chandra and Staiger (2010): Discussion

- Do not find evidence of taste-based discrimination
  - If anything, women and minorities appear to have slightly *smaller* benefits from treatment relative to men and whites
- Section VI discusses several potential explanations
- Argue results most consistent with statistical discrimination
- Unclear why minorities, women are less appropriate for treatment: key to interpreting findings, public policy relevance

1. Regression analysis
   - Goldberger (1984)
   - Neal and Johnson (1996)
2. Audit studies
   - Bertrand and Mullainathan (2004)
3. Quasi-experiments
   - Goldin and Rouse (2000)
   - Anwar, Bayer, and Hjalmarsson (2012)
4. Testing models
   - Charles and Guryan (2008)
   - Chandra and Staiger (2010)
5. Looking ahead

Looking ahead

Discrimination and learning

14.662 Labor Economics II

Spring 2015