2012 AZilg Compliance Conference

Pre-Conference Session: Understanding Adverse Impact Analyses

November 15, 2012 • 9:00 – 10:00am

Daniel Kuang, Ph.D.
Patrick Nooren, Ph.D.
Biddle Consulting Group, Inc.

Contact Information

Biddle Consulting Group, Inc.
193 Blue Ravine, Suite 270
Folsom, CA 95630
916.294.4250
www.biddle.com | www.bcginstitute.org

Daniel Kuang, Ph.D. (dkuang@biddle.com)
Patrick Nooren Ph.D. (pnooren@biddle.com)

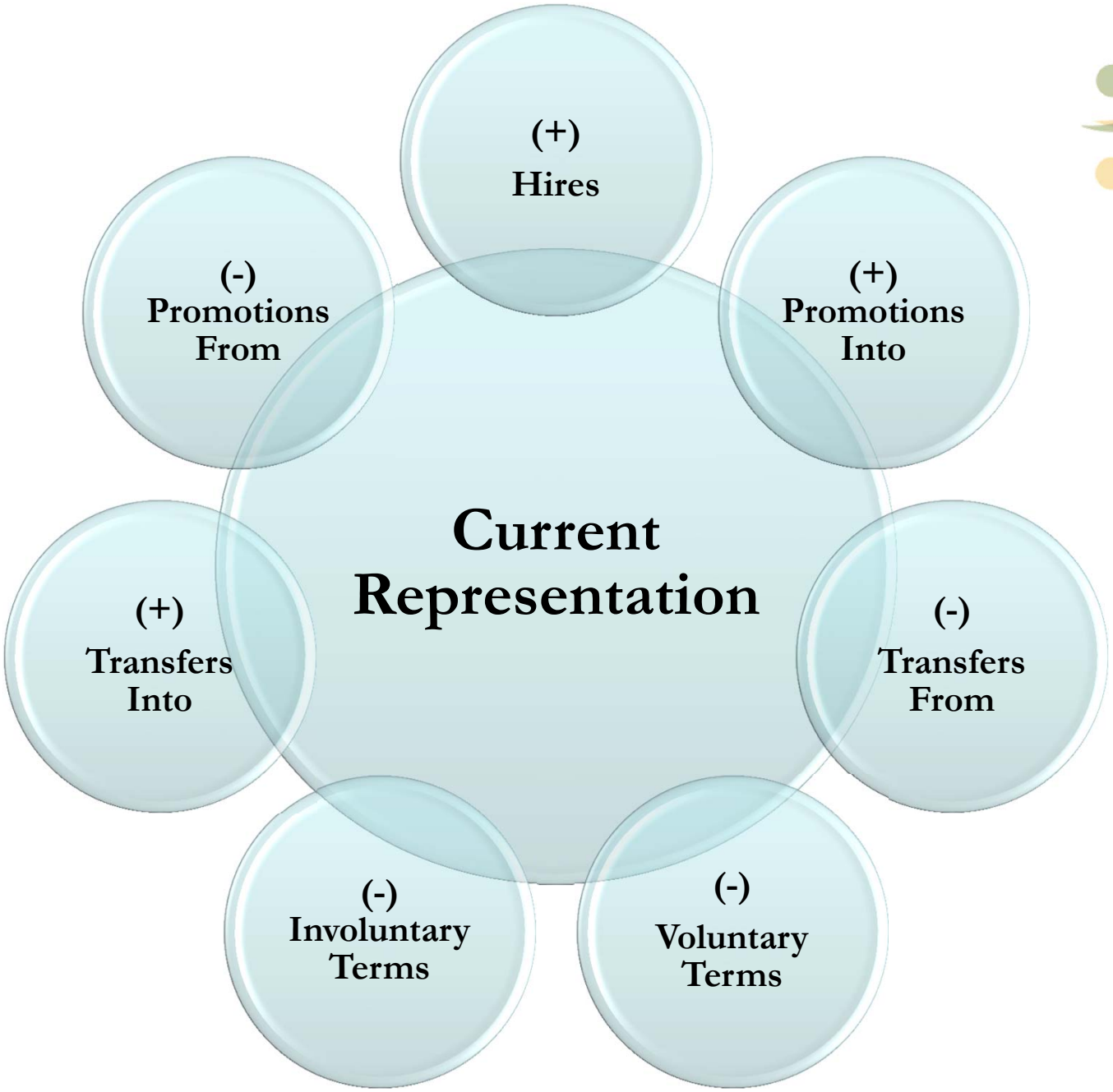
Agenda

Comparison of Selection Rates: How We Got to Where We Are Today

Adverse Impact: The Typical Approach

Adverse Impact: The Right Way

Comparison of Selection Rates: How We Got to Where We Are Today



Adverse/Disparate Impact: Legal Overview

DISPARATE IMPACT

An unlawful employment practice based on disparate impact is established only if:

1. A complaining party demonstrates that a respondent uses a particular employment practice that causes an adverse impact

and

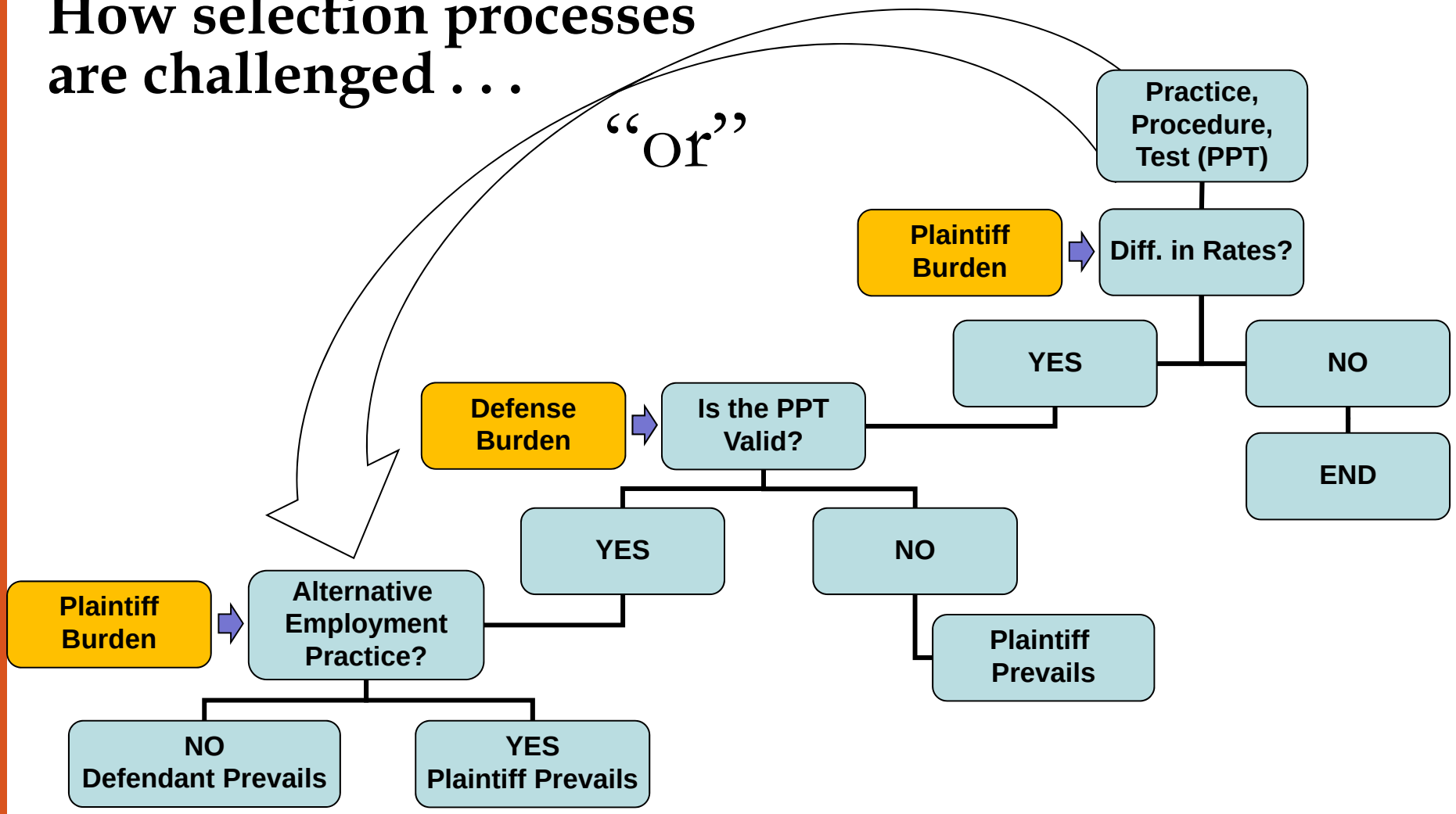
2. the respondent fails to demonstrate that the challenged practice is job-related for the position in question and consistent with business necessity

or

3. the complaining party makes the demonstration described above with respect to an alternate employment practice, and the respondent refuses to adopt such alternative employment practice.

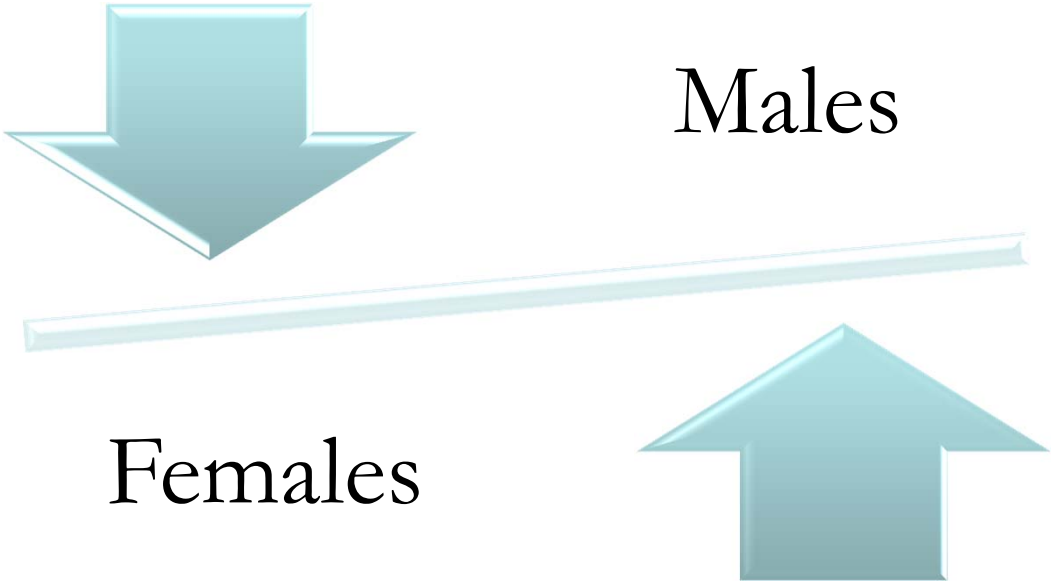
Title VII Disparate Impact Discrimination Flowchart

How selection processes are challenged . . .



Selection Rate Comparison

Does a practice, procedure or test (PPT) results in **disproportionate selection rates** by gender, race/ethnicity, or age group?



Requirements of an Affirmative Action Plan

Example of *Personnel Transaction Report*

Company Name - Plan Name

Personnel Transactions Summary

Job Group: 1B – Directors and Senior Managers

Transaction Dates: 01/01/2007 To 12/31/2007

	Applicants				*Hires*			*Terminations (I)*			*Terminations (V)*		
	MALE	FEMALE	UNKNOWN (GENDER)	TOTAL RACE	MALE	FEMALE	TOTAL RACE	MALE	FEMALE	TOTAL RACE	MALE	FEMALE	TOTAL RACE
White	64	0	0	64	16	0	16	1	0	1	5	0	5
African American	0	0	0	0	0	0	0	0	0	0	0	0	0
Hispanic	17	0	0	17	7	0	7	2	0	2	12	0	12
Asian	15	0	0	15	0	0	0	0	0	0	0	0	0
Native American	0	0	0	0	0	0	0	0	0	0	0	0	0
NHOPI	0	0	0	0	0	0	0	0	0	0	0	0	0
Multiple Race	0	0	0	0	1	0	1	0	0	0	0	0	0
Unknown (Race)	0	0	0	0									
Total	96	0	0	96	24	0	24	3	0	3	17	0	17
Total Minority	32	0	0	32	8	0	8	2	0	2	12	0	12

	Promotions From			*Promotions Into*			*Promotions Within*		
	MALE	FEMALE	TOTAL RACE	MALE	FEMALE	TOTAL RACE	MALE	FEMALE	TOTAL RACE
White	4	0	4	1	0	1	2	0	2
African American	0	0	0	0	0	0	0	0	0
Hispanic	0	0	0	0	0	0	1	0	1
Asian	0	0	0	0	0	0	0	0	0
Native American	0	0	0	0	0	0	0	0	0
NHOPI	0	0	0	0	0	0	0	0	0
Multiple Race	0	0	0	1	0	0	0	0	0
Total	4	0	4	2	0	1	3	0	3
Total Minority	0	0	0	1	0	0	1	0	1

Selection Rate Comparison

- 2 X 2 Table Comparison
- Impact Ratio Analysis (IRA)
- Fisher Exact / Chi-Square / 80% Test

Men Pass (50)	Men Fail (50)	⇨ Men Passing Rate (50%)
Women Pass (25)	Women Fail (75)	⇨ Women Passing Rate (25%)

Results in a value indicating if the observed difference in rates is due to chance (i.e., statistically significant).

Statistical Significance

- Statistical Significance (Thresholds):
  - 5%
  - 0.05
  - 1 chance in 20
  - 2.0 Standard Deviations *(actually 1.96)*
- Statistical Significance (Outputs)
  - Lower p-values = higher SD (or "Z") values
  - For example:
    - P-value: .05 = 1.96 SDs
    - P-value: .01 = 2.58 SDs

Types of Adverse Impact Analyses (Sample)

Male v. Female			
Analysis	**Starting**	**Completing**	**Result**
Overall (App v. Hired)	**Male - 100 Female - 100**	**Male - 50 Female - 30**	**2.81 SD**

Applied v. External Availability ("barrier" analysis)	**Took v. Pass/Fail Interview**
Applied v. Pass/Fail BQs	**Avail for Prom v. Prom (promo "from")**
Took v. Pass/Fail Test	**Applied v. Prom (competitive)**
Avail for Interview v. Interviewed	**Avail for Term v. Retained (Vol/Invol)**

Adverse Impact:
The Typical Approach

Statistical Significance and Power

- Statistical significance: The point at which differences become large enough that one can claim a trend exists.
- Statistical power: The ability to see those trends if, in fact, they do exist.
- Statistical power is directly related to effect size and sample size:
  - **Effect size**: The size of the difference in selection rates between two groups . . . the larger the difference the less number of transactions necessary to detect statistical significance
  - **Sample size**: With larger numbers of transactions it becomes much easier to detect statistical significance

Statistical Significance and Power

- Enforcement agencies have no control over effect size (i.e., the difference in selection rates), but they do have some control over sample size . . . which is why they often request two (2) years worth of data (or more: Frito-Lay) to analyze.
- However, simply aggregating all applicants and all hires across strata (as is typically done), can sometimes result in incorrect/misleading findings.

Statistical Significance and Power

	Men		Women	
	(#)	(%)	(#)	(%)
Pass	100	50.0	90	45.0
Fail	100	50.0	110	55.0
80% Test [1]	Company A		0.90	
Stat. Test-**EXACT** [2]			0.343	
Pract. Test (**to Exact Test**) [3]			N/A	
Stat. Test-**ESTIMATED** [2]			0.317	

	Men		Women	
	(#)	(%)	(#)	(%)
	200	50.0	180	45.0
	200	50.0	220	55.0
	Company B		0.90	
			0.168	
			N/A	
			0.157	

	Men		Women	
	(#)	(%)	(#)	(%)
	300	50.0	270	45.0
	300	50.0	330	55.0
	Company C		0.90	
			0.088	
			N/A	
			0.083	

	Men		Women	
	(#)	(%)	(#)	(%)
	400	50.0	360	45.0
	400	50.0	440	55.0
	Company D		0.90	
			0.048	
			N/A	
			0.045	

Adverse Impact: The Typical Approach

- Analyses by AAP job group regardless of different:
  - Job titles
  - Selection processes
  - Hiring managers
  - Basic qualifications
  - Locations (perhaps)
  - Applicant pools for separate requisitions (perhaps)
- Typically an aggregation of 12 months (sometimes 18/24 months) worth of transactions into a single 2x2 table
- Considers everyone who applied throughout the year as available for every hire throughout the year

Adverse Impact: The Typical Approach

ALL applicants and ALL hires for a 12-month period

Men Pass	Men Fail
Women Pass	Women Fail

There is nothing wrong with this approach . . . as an initial inquiry only. Sometimes this approach is used as the basis for a Notice of Violation (NOV) or plaintiff class action litigation; however, it is up to the employer to provide rebuttal analyses that may more accurately reflect reality.

Simpson's Paradox

Job Title	Group	Applicants (#)	Selected (#)	Selection Rate (%)
Warehouse Person	Men	400	200	50.0%
	Women	100	50	50.0%
Laborer	Men	100	20	20.0%
	Women	100	20	20.0%
W/H Person + Laborer Combined	Men	500	220	44.0%
	Women	200	70	35.0%

- Fisher Exact Test: SD = 2.16 (Significant)
- Mantel-Haenszel: SD = .024 (NOT Significant)

Adverse Impact: The Right Way

Single Event v. Multiple Events

ALL applicants and ALL hires throughout the time period

Men Pass	Men Fail
Women Pass	Women Fail

= Chi-Square or Fisher's Exact

Req. 1 + **Req. 2** + **Req. 3**

Men Pass	Men Fail
Women Pass	Women Fail

Men Pass	Men Fail
Women Pass	Women Fail

Men Pass	Men Fail
Women Pass	Women Fail

Mantel-Haenszel (MH) Defined

- In the context of selection rate analyses, the MH:
  - is a statistical tool that allows researchers to appropriately combine separate and distinct selection processes (e.g., requisitions) into a single analysis
  - appropriately allows for the benefits of increased sample size while controlling for Simpson's Paradox
  - is a useful tool for evaluating whether the employer has a discriminatory "pattern or practice"
  - may (if appropriate) be used to rebut allegations based on a single (overly) aggregated analysis

Mantel-Haenszel (MH) Defined

- The Mantel-Haenszel is not for every situation. It requires separate and distinct pools of applicants/test takers, etc. For example, combining:
  - Requisitions
  - Locations
  - Different jobs in same job group
  - Different hiring seasons and/or groups
- Not applicable for "pooled requisitions" where the requisition stays "open" and applicants are just regularly selected from the pool.

Component "Step" Analyses

Component "Step" Analyses

Title VII of 1964/1991 Civil Rights Act

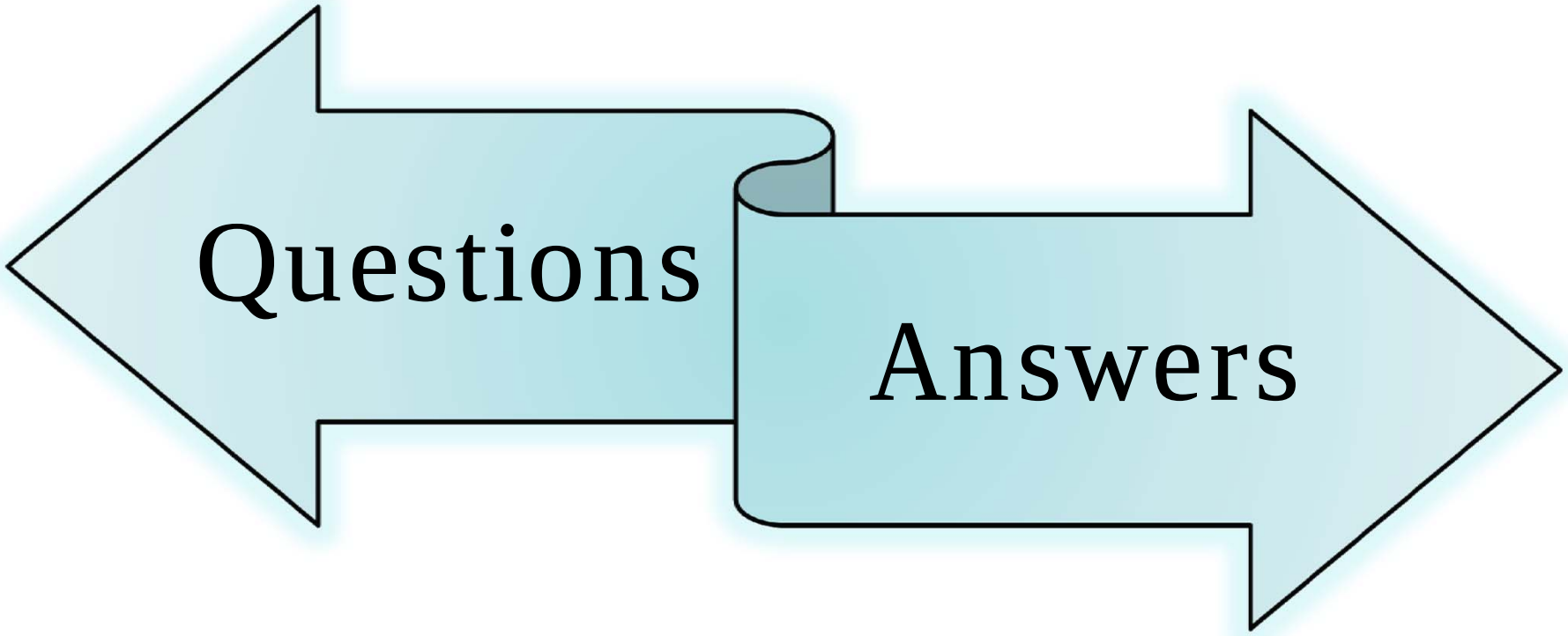
An unlawful employment practice based on disparate impact is established under this title only if a complaining party demonstrates that a respondent uses a particular employment practice that causes a disparate impact . . .

Important Note: Enforcement agencies have every right to investigate the practices, procedures, and tests contractors use to screen applicants. However, in the past, due to resource constraints they wouldn't typically do so unless there was adverse impact in the overall hiring process.

Times have changed!

Component "Step" Analyses

Male v. Female			
Steps	**Starting**	**Completing**	**Result**
Overall (App vs. Hired)	**Male - 100** **Female - 100**	**Male - 50** **Female - 30**	**2.81 SD**
1. Basic Qualifications	Male - 100 Female -100	Male – 79 Female - 77	0.25 SD
2. Test	**Male - 79** **Female - 77**	**Male – 65** **Female - 35**	**4.80 SD**
3. Interview	Male - 65 Female - 35	Male – 60 Female - 32	0.18 SD
4. Final Selection	Male - 60 Female - 32	Male – 50 Female - 30	0.00 SD



Logistic Regression and Hiring: I Understand There is a Significant Difference in Hiring Rates . . . But It's Based on Job-Related Practices, Procedures, or Tests

Logistic Regression

- The goal is not to instruct how to perform proper logistic regression analyses.
- The goal is to inform you that there is a way to analyze whether applicant differences in job-related criteria are the "real" reason why there is a disparity in hiring rates.
- Particularly useful when/if the OFCCP claims discrimination but you ***know*** that there is a legitimate reason for the disparity.

Logistic Regression

- Classic adverse impact analyses can only determine if the numerical difference in passing rates between two groups is significantly different.

- Logistic Regression (LR) can identify if that numerical difference in passing rates is due to applicant differences in job-related criteria (e.g., experience or education).

Logistic Regression

Logistic Regression

- LR needs to be applied to job-related factors that were ***actually used or considered*** in the selection process
- LR is useful for weighing the practical importance of job-related factors in the hiring or promotion process
- LR can potentially "pin" the impact on specific job-related criteria
- LR can also be useful for determining "shortfall" calculations
  - For example, how many women would have been hired "but for" the possible discrimination?
  - What is the total shortfall for women, given what the model can explain?

Logistic Regression

Gender Alone						
	B	S.E.	Wald	df	Sig.	Exp(B)
Gender	1.39	.41	11.61	1	0.0006	4.03
Constant	-3.23	0.38	70.26	1	0.00	0.039

Gender is significant in the absence of any job-related explanatory variables.

Gender and Qualification Factors						
	B	S.E.	Wald	df	Sig.	Exp(B)
Gender	0.36	0.45	0.69	1	0.414	1.24
Prior Experience	0.915	0.68	1.78	1	0.18	2.49
Education	1.87	0.93	4.04	1	0.04	6.51
Constant	-5.07	0.76	43.71	1	0.00	0.006

After controlling for gender differences in the job-related variables, gender (itself) is no longer significant.

Summary and Conclusion

I will leave you with the following:

- Adverse impact in hiring drives audits (at least the really painful ones)
- Applicant data is paramount to analyzing adverse impact in hiring (therefore, the OFCCP spends a lot of time focusing on it)
- It is up to you (the employer) to ensure the OFCCP is properly analyzing the data (and that "analyses reflect reality")
- Adverse impact is simply a numerical disparity in rates between two groups . . . there might be a defensible, job-related reason why that is occurring . . . consider logistic regression if the stakes are high and the audit is getting "ugly"
- Remember: Adverse impact (alone) is NOT discrimination (only in the absence of validity evidence does it become discrimination)