Causality

Expert Testimony

How to "control for other factors"? The pivot table as a way to compute conditional mean wages.

The multivariation regression captures the effect of a change in X, *holding other factors fixed*. How do we hold other factors fixed? We include them in the regression as additional explanatory variables.

Causal graphs for discrimination

```
flowchart LR
    gender -->
    performance -->
    salary
```

```
flowchart LR
    gender -->
    performance -->
    salary
    gender -- discrimination --> salary
```

```
flowchart LR
    gender -->
    performance -->
    metric -->
    salary
```

```
flowchart LR
    gender -->
    performance -->
```

Disparate treatment

```
graphLR
PERFORMANCE -> DECISION
DISCRIMINATION -> DECISION
```

How to detect

DECISION $=$ a $+$ b PERFORMANCE $+$ c MINORITY

c $=/=$ 0

Tainted variable

PERFORMANCE -> METRIC DISCRIMINATION -> METRIC METRIC -> DECISION

Defendant: DECISION = a + b METRIC + c MINORITY and c = 0

Plaintiff: but METRIC already includes discrimination

This can be subject to Title VII if Defendant cannot prove that METRIC is a good, non-discriminatory measure of performance.

Clearly tainted variable

PERFORMANCE -> METRIC1 DISCRIMINATION -> METRIC2 METRIC2 -> DECISION

Defendant: DECISION = a + b METRIC2 + c MINORITY and c = 0

Plaintiff: but METRIC2 already includes discrimination

In this case, METRIC2 is a bad proxy for PERFORMANCE, METRIC1 would be better. A clear case of disparate impact.

Societal discrimination

DISCRIMINATION -> PERFORMANCE PERFORMANCE -> DECISION

Hard to prove. Defendant may not be at fault. But still wrong.

A special case

DISCRIMINATION -> DECISION Expectation of DISCRIMINATION -> PERFORMANCE PERFORMANCE -> DECISION

For example, in an all-white college, a black student may not even apply. No application, no overt discrimination. Almost impossible to detect, need data outside the organization.