Roadmap

1. Omitted Variable Bias
2. Coefficient Stability – Oster 2015 Paper
5. Recap
Omitted Variable Bias

• What is omitted variable bias (OVB)?

• OVB is something that we will always have to contend with because we do not have double-blind random experiments in empirical corporate finance research.
  • There are several reasons why a variable might be excluded from a regression. In ECF, it's typically because the variable can’t be measured, like CEO skill or investment opportunity set.
Omitted Variable Bias

A classical OVB problem:

\[ Salary_i = \beta_0 + \beta_1 \text{education}_i + \varepsilon_i \]
Omitted Variable Bias

Consider a general linear model of the form
\[ y_i = x_i \beta + z_i \delta + u_i \]

Where
- \( x_i \) is a 1 x \( p \) row vector of values of \( p \) independent variables
- \( \beta \) is a \( p \) x 1 column vector of unobservable parameters to be estimated
- \( z_i \) is a scalar and is the value of another independent variable
- \( \delta \) is a scalar and is an unobservable parameter to be estimated
Omitted Variable Bias

\[ y_i = x_i \beta + z_i \delta + u_i \]

To actually run the regression, we form the following matrices

\[ X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^{n \times p} \]

\[ Y = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}, \]

\[ Z = \begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix}, \]

\[ U = \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix} \in \mathbb{R}^{n \times 1} \]
Omitted Variable Bias

\[ y_i = x_i \beta + z_i \delta + u_i \]

Then, we use least squares calculation to get:

\[ \hat{\beta} = (X'X)^{-1}X'Y \]

Substituting for \( Y \) based on the linear model:

\[ \hat{\beta} = (X'X)^{-1}X'(X\beta + Z\delta + U) \]

\[ = (X'X)^{-1}X'X\beta + (X'X)^{-1}X'Z\delta + (X'X)^{-1}X'U \]

\[ = \beta + (X'X)^{-1}X'Z\delta + (X'X)^{-1}X'U \]
Omitted Variable Bias

\[ y_i = x_i \beta + z_i \delta + u_i \]

Taking the expectation:
\[ \mathbb{E}[\hat{\beta} \mid X] = \beta + (X'X)^{-1}X'Z\delta \]

\[ = \beta + \text{bias} \]

The second term is the OVB, which is non-zero if the omitted variable \( z \) is correlated with any of the included variables in the matrix \( X \). This is the problem that won’t go away in ECF that we have to deal with.
Omitted Variable Bias

The nature of the bias on the included independent variables depends on the nature of the correlation between (a) the dependent and the excluded variable and (b) the included and excluded independent variables. \( \beta \)

<table>
<thead>
<tr>
<th>Negative Correlation, Y and Z</th>
<th>Positive Correlation, X and Z</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \beta ) is \textit{OVEREstimated}</td>
<td>( \beta ) is \textit{UNDEREstimated}</td>
</tr>
<tr>
<td>( \beta ) is \textit{UNDEREstimated}</td>
<td>( \beta ) is \textit{OVEREstimated}</td>
</tr>
</tbody>
</table>
Omitted Variable Bias

\[
y_i = x_i \beta + z_i \delta + u_i
\]

The univariate omitted variable bias formula looks like:

\[
\hat{\beta} \approx \beta + \frac{\delta \text{Cov}(x_i, z_i)}{\text{Var}(x_i)}
\]

If \( \delta \text{Cov}(x_i, z_i) > 0 \), the omitted variable bias is positive.
If \( \delta \text{Cov}(x_i, z_i) < 0 \), the omitted variable bias is negative.
If \( \delta = 0 \) or \( \text{Cov}(x_i, z_i) = 0 \), there is no omitted variable bias.
Omitted Variable Bias

As preliminary check on whether your model has missing variables, Stata has two built in commands *linktest* and *ovtest* that should be run after your regression and can provide evidence on whether your model is misspecified. I tried them both and they are easy to use.

You can learn more about these two commands at:

http://www.ats.ucla.edu/stat/stata/webbooks/reg/chapter2/statareg2.htm
<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Book-to-Bill Indicator</td>
<td>Book-to-Bill Indicator</td>
<td>Book-to-Bill Indicator</td>
<td>Book-to-Bill Indicator</td>
<td>Book-to-Bill Indicator</td>
<td>Book-to-Bill Indicator</td>
</tr>
<tr>
<td>Short Term Earnings</td>
<td>0.039***</td>
<td>0.040***</td>
<td>0.040***</td>
<td>0.040***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(3.46)</td>
<td>(3.55)</td>
<td>(3.54)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Long Term Earnings</td>
<td>0.033**</td>
<td>0.037**</td>
<td>0.037**</td>
<td>0.037**</td>
<td>0.037**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(2.05)</td>
<td>(2.31)</td>
<td>(2.29)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unexpected Revenue</td>
<td>-0.024</td>
<td>-0.026</td>
<td>-0.030*</td>
<td>-0.031*</td>
<td>-0.029</td>
<td>-0.031*</td>
</tr>
<tr>
<td></td>
<td>(-1.28)</td>
<td>(-1.37)</td>
<td>(-1.68)</td>
<td>(-1.76)</td>
<td>(-1.54)</td>
<td>(-1.66)</td>
</tr>
<tr>
<td>Unexpected Earnings</td>
<td>0.036**</td>
<td>0.027*</td>
<td>0.039***</td>
<td>0.029**</td>
<td>0.040**</td>
<td>0.029*</td>
</tr>
<tr>
<td></td>
<td>(2.35)</td>
<td>(1.76)</td>
<td>(2.61)</td>
<td>(1.97)</td>
<td>(2.55)</td>
<td>(1.88)</td>
</tr>
<tr>
<td>Pre-Ann. Return</td>
<td>-0.029</td>
<td>-0.029</td>
<td>-0.027</td>
<td>-0.029</td>
<td>-0.027</td>
<td>-0.028</td>
</tr>
<tr>
<td></td>
<td>(-1.34)</td>
<td>(-1.37)</td>
<td>(-1.28)</td>
<td>(-1.34)</td>
<td>(-1.27)</td>
<td>(-1.34)</td>
</tr>
<tr>
<td>Market Cap</td>
<td>0.379***</td>
<td>0.386***</td>
<td>0.021</td>
<td>0.029</td>
<td>0.026</td>
<td>0.030</td>
</tr>
<tr>
<td></td>
<td>(4.06)</td>
<td>(4.11)</td>
<td>(0.17)</td>
<td>(0.24)</td>
<td>(0.21)</td>
<td>(0.25)</td>
</tr>
<tr>
<td>ROA</td>
<td>-0.025</td>
<td>-0.021</td>
<td>-0.086**</td>
<td>-0.085**</td>
<td>-0.087**</td>
<td>-0.085**</td>
</tr>
<tr>
<td></td>
<td>(-0.75)</td>
<td>(-0.62)</td>
<td>(-2.00)</td>
<td>(-1.98)</td>
<td>(-2.02)</td>
<td>(-1.98)</td>
</tr>
<tr>
<td>Book-to-Market</td>
<td>0.002</td>
<td>0.004</td>
<td>-0.086**</td>
<td>-0.085**</td>
<td>-0.087**</td>
<td>-0.085**</td>
</tr>
<tr>
<td></td>
<td>(0.06)</td>
<td>(0.12)</td>
<td>(-2.00)</td>
<td>(-1.98)</td>
<td>(-2.02)</td>
<td>(-1.98)</td>
</tr>
<tr>
<td>Total Assets</td>
<td>0.536***</td>
<td>0.540***</td>
<td>0.533***</td>
<td>0.539***</td>
<td>0.539***</td>
<td>0.539***</td>
</tr>
<tr>
<td></td>
<td>(3.31)</td>
<td>(3.33)</td>
<td>(3.31)</td>
<td>(3.31)</td>
<td>(3.34)</td>
<td>(3.34)</td>
</tr>
<tr>
<td>Observations</td>
<td>12837</td>
<td>12837</td>
<td>12837</td>
<td>12837</td>
<td>12837</td>
<td>12837</td>
</tr>
<tr>
<td>Adjusted R-squared</td>
<td>0.397</td>
<td>0.396</td>
<td>0.399</td>
<td>0.399</td>
<td>0.399</td>
<td>0.399</td>
</tr>
</tbody>
</table>
Coefficient Stability

• As discussed, the vast majority of the work we do is non-experimental, thus we have to worry about omitted variable bias. This is particularly true when we are trying to find a causal relation.

• The most straightforward way to tackle OVB is to include a variable that captures the omitted variable. If the reason for the omission is because the variable unobservable, we include controls that are observable to help mitigate this.

• However, there is an implicit assumption when using observable control variables to mitigate omitted variable bias. Namely, that selection on observables is informative about selection on observables.
Coefficient Stability

1. Using observables to identify the bias from the unobservables requires making assumptions about the covariance properties. This can lead to problems. If the unobservables are completely unlike the observables, *nothing* about the remaining bias is learned from the inclusion of the observables.
   - E.g. maybe parental education tells us nothing about the innate, unobservable ability of person. Then including parental education in a regression of wages on education does nothing for this particular bias.

2. However, even in the opposite case, where the unobservables 100% share the same covariance properties as the observables, you cannot look at just coefficient movement.
Table 1 for Oster

This is a regression of wage returns to education with individual ability as the only confounding variable. Assume that ability has two orthogonal components, each of which contribute to determining wages. However, the researcher only gets to observe one of the two ability controls.

<table>
<thead>
<tr>
<th>Quality of Observed Control</th>
<th>Uncontrolled Coefficient [R^2]</th>
<th>Controlled Coefficient [R^2]</th>
<th>True Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>High Variance Control Observed</td>
<td>0.202 [.004]</td>
<td>0.002 [.990]</td>
<td>0</td>
</tr>
<tr>
<td>Low Variance Control Observed</td>
<td>0.202 [.004]</td>
<td>0.200 [.013]</td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Control Importance in Explaining Y</th>
<th>Uncontrolled Coefficient [R^2]</th>
<th>Controlled Coefficient [R^2]</th>
<th>True Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Very Important</td>
<td>0.20 [.004]</td>
<td>0.195 [.95]</td>
<td>0.194</td>
</tr>
<tr>
<td>Not Important</td>
<td>0.20 [.004]</td>
<td>0.195 [.01]</td>
<td>-0.63</td>
</tr>
</tbody>
</table>
Coefficient Stability

It many empirical economics papers, when discussing robustness of results, it is very uncommon to mention the movement of $R^2$ when different control variables are added. This needs to change.

Next, Oster will formalize a way to relate coefficient and $R^2$ movements to omitted variable bias into a test that can be run.
Coefficient Stability Theory

Goal: estimating an unbiased treatment effect from a model in which there are some observed confounders and some unobserved confounders.

*Important Caveat: This approach works only for unobservables which are related to the observables.*

Note: this is for linear models, if you need to perform a similar analysis for a non-linear model (like a bivariate probit model) consult Altonji, Elder and Taber (2005).
Coefficient Stability

Consider the following model:

\[ Y = \beta X + \gamma_1 w_1^0 + W_2 + \epsilon \]

Where \( X \) is the treatment and \( \beta \) is the coefficient of interest.
\( w_1^0 \) is an observed control variable.
\( W_2 \) is a vector which is a linear combination of unobserved control variables \( w_j^u \).
Coefficient Stability

\[ Y = \beta X + W_1 + W_2 + \varepsilon \]

Assume that \( Cov(W_1, W_2) = 0 \) and \( Var(X) = \sigma_{xx} \)

The covariance matrix associated with the vector \([X, W_1, W_2]'\) is PD.

Define the equal selection relation as \( \frac{\sigma_{1X}}{\sigma_{11}} = \frac{\sigma_{2X}}{\sigma_{22}} \), where \( \sigma_{iX} = Cov(W_i, X) \), \( \sigma_{ii} = Var(W_i) \)

This equal selection relation is important. It implies that the ratio of the movement in coefficients is equal to the ratio of the movement in \( R^2 \).
Coefficient Stability

\( \hat{R} \) is the \( R^2 \) from the regression \( Y = \hat{\beta}X + \varepsilon \)

\( \tilde{R} \) is the \( R^2 \) from the regression \( Y = \tilde{\beta}X + \gamma_1 w_1^0 + \varepsilon \)

\( R_{\text{max}} \) is the \( R^2 \) from a hypothetical regression of \[ Y = \beta X + \gamma_1 w_1^0 + W_2 + \varepsilon \]

The OVB on \( \hat{\beta} \) and \( \tilde{\beta} \) are controlled by the auxiliary regressions
Coefficient Stability

You can express the probability limits of the short and intermediate regression coefficients as:

\[
\hat{\beta} \overset{p}{\rightarrow} \beta + \gamma_1 \lambda_{W_1^0|X} + \lambda_{W_2|X} \\
\tilde{\beta} \overset{p}{\rightarrow} \beta + \lambda_{W_2|X,W_1^0}
\]

The asymptotic bias on \( \tilde{\beta} \) (the coefficient on \( X \) with controls included) is \( \lambda_{W_2|X,W_1^0} \) which is equal to \( \frac{\sigma_{22} \sigma_{1X}}{\sigma_{11}(\sigma_{XX} - \frac{\sigma_{1X}^2}{\sigma_{11}})} = \Pi \)
Coefficient Stability

Define the following:

\[ \beta^* = \tilde{\beta} - [\dot{\beta} - \tilde{\beta}] \frac{R_{max} - \tilde{R}}{\tilde{R} - \dot{R}} \]

Which finally gives us Oster’s proposition 1:

\[ \beta^* \overset{p}{\rightarrow} \beta \]

Check the proof in the paper if you want.
Coefficient Stability

Oster then provides a general estimator for the case where selection is proportional and there are possibly multiple variables in the set of observable controls.

The main difference is that, instead of the equal selection relation that we had in the single variable case, we now have a proportional selection relationship, defined as:

$$\delta \frac{\sigma_{1X}}{\sigma_{11}} = \frac{\sigma_{2X}}{\sigma_{22}}$$

Where $\delta$ is the coefficient of proportionality. Most of the derivation is then very similar to the simple estimator.
Coefficient Stability

Implementation:

1. Assume a value for $R_{\text{max}}$ and calculate $\delta$ for which $\beta = 0$.
   (For example, a value of $\delta = 2$ would suggest that the unobservables would need to be twice as important as the observables to produce a treatment effect of zero. A value of $\delta = 1$ is common heuristic.)

2. Use bounding assumptions on both $R_{\text{max}}$ and $\delta$ to develop a set of bounds for $\beta$.

Either of these methods can be used in STATA with the `psacalc` command that Oster has provided.
Coefficient Stability

Main takeaways:
1. Heuristic coefficient stability is, at best, uninformative and, at worst, misleading.
2. You must consider $R^2$ movements to determine if the treatment effect truly is robust to control inclusion.
3. However, when using multiple controls, you must use the formal test in the Oster paper to get a clear idea of the robustness of the result.
Poorly Measured Confounders

• A common presumption in empirical work is that if an added control variable does not affect the coefficient of interest then the treatment effect is properly measured or the study is reliable.

• However, it may be the case that the added control variable is poorly measured/captured and does not actually capture the underlying confounder. If this is the case, the fact that the coefficient of interest does not vary might be meaningless.

• There are several tests that can help you to validate the use of a particular control variable and these tests should be performed as a part of your empirical work.
Poorly Measured Confounders

Let’s say you have a candidate confounder variable. There are two tests to see if including this variable will help reduce selection bias.

1. Coefficient Comparison Test - Confounder added to the RHS and if the estimated coefficient of interest is unaffected, you are good.

\[ Y = \alpha + \beta X + \delta Z + \epsilon \]

2. Balancing Test – Confounder is placed on LHS instead of outcome variable and you are looking for a zero coefficient on the parameter of interest.

\[ Z = \alpha + \beta X + \epsilon \]
Poorly Measured Confounders

Start with the following regression:
\[ y_i = \alpha^s + \beta^s s_i + \varepsilon_i^s \]

Now consider a possible confounder
\[ y_i = \alpha + \beta s_i + \gamma x_i + \varepsilon_i \]

The Balancing test:
\[ x_i = \delta_0 + \delta s_i + u_i \]

The change in the coefficient $\beta$ from adding $x_i$ to the first regression is given by the OVB formula
\[ \beta^s - \beta = \gamma \delta \]

Where $\delta$ is the coefficient on $s_i$ in the balancing regression.
Poorly Measured Confounders

What is the relation between the Coefficient Comparison Test and the Balancing Test? The Balancing Test looks at the null hypothesis:

$$H_0: \delta = 0$$

The Coefficient Comparison Test looks at the null hypothesis:

$$H_0: \beta^s - \beta = 0$$

Remember that $$\beta^s - \beta = \gamma \delta$$

So then the actual test can be written as:

$$H_0: \beta^s - \beta = 0 \iff \gamma = 0 \text{ or } \delta = 0$$
Poorly Measured Confounders

Why is the Balancing Test more powerful when $x_i$ is noisy?

Formally: $x_i^m = x_i + m_i$

You then compare the two regressions

$$y_i = \alpha^s + \beta^s s_i + e_i^s$$
$$y_i = \alpha^m + \beta^m s_i + \gamma^m x_i^m + \epsilon_i^m$$

In the long regression $\beta^m$ and $\gamma^m$ are biased.

$$\beta^m = \beta + \gamma \delta \frac{1 - \lambda}{1 - R^2} = \beta + \gamma \delta \theta$$
$$\gamma^m = \gamma \frac{\lambda - R^2}{1 - R^2} = \gamma (1 - \theta)$$
Poorly Measured Confounders

In the paper they perform a few more manipulations but then we are left with the following:

\[ \gamma = \gamma^m \frac{1 - R^2}{\lambda - R^2} \]

\[ \beta = \beta^m - \delta \gamma^m \frac{1 - \lambda}{\lambda - R^2} \]

Since \( R^2 \) is observed from the data this only involves the unknown parameter \( \lambda \). Even if \( \lambda \) is not know precisely, we can use these equations to bound \( \beta \) for a range of plausible reliabilities.
Poorly Measured Confounders

Comparing the power function of the two tests:

Balancing Test

$$Power_{t_{\delta m}}(d) = 1 - \Phi \left( 1.96 - d \frac{\sqrt{n} \sigma_s \sqrt{1 - \theta}}{\sigma_u} \right) + \Phi \left( -1.96 - d \frac{\sqrt{n} \sigma_s \sqrt{1 - \theta}}{\sigma_u} \right)$$

Coefficient Comparison Test

$$Power_{t_{(\beta s - \beta m)}}(d; \gamma) = 1 - \Phi \left( 1.96 - d \frac{\sqrt{n} \gamma (1 - \theta)}{\sqrt{V_\beta(d; \gamma)}} \right) + \Phi \left( -1.96 - d \frac{\sqrt{n} \gamma (1 - \theta)}{\sqrt{V_\beta(d; \gamma)}} \right)$$
Poorly Measured Confounders

After some algebraic manipulation you are left with:

$$\text{Power}_{t_\delta} (d) > \text{Power}_{t(\beta^s - \beta^m)} (d; \gamma)$$

This sometimes matters. Let’s look at the examples.
Poorly Measured Confounders

Figure 1: Theoretical Rejection Rates

- balancing test
- cc test
- balancing test rel=0.5
- cc test rel=0.5
- balancing test rel=0.25
- cc test rel=0.25
Poorly Measured Confounders

1. Both the Coefficient Comparison Test and the Balancing Test are useful and should be used to validate research designs.
2. The Balancing Test is generally more powerful than the Coefficient Comparison Test, especially when the candidate confounder is noisy.
3. Successfully passing the balancing test should be a necessary condition for a successful research design but it is not sufficient.
## Adding Covariates Sequentially

**Table 1: Results from Table 4 of Altonji & Blank (1999)**

<table>
<thead>
<tr>
<th>Included covariates</th>
<th>Model: (4)</th>
<th>Model: (5)</th>
<th>Model: (6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coefficient on black indicator</td>
<td>-0.207</td>
<td>-0.117</td>
<td>-0.089</td>
</tr>
<tr>
<td>Hispanic and female indicators</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Education, experience, &amp; region</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Occupation, industry, job characteristics</td>
<td>N</td>
<td>N</td>
<td>Y</td>
</tr>
</tbody>
</table>
Adding Covariates Sequentially

### Table 2: Results from Tables 1 and 2 of Levitt & Syverson (2008)

<table>
<thead>
<tr>
<th>Included covariates</th>
<th>Table 1 Column (4)</th>
<th>Table 2 (1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coefficient on agent dummy</td>
<td>0.128</td>
<td>0.048</td>
<td>0.042</td>
<td>0.038</td>
<td>0.037</td>
</tr>
<tr>
<td>City-by-year interactions</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Basic house characteristics</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Indicators of house quality</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Keywords in description</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Block fixed effects</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>Y</td>
</tr>
</tbody>
</table>
Adding Covariates Sequentially

Table 3: NLSY Wage Results for Young Men, Similar to Table 1 of Neal & Johnson (1996)

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coefficient on black indicator (log points)</td>
<td>-22.1</td>
<td>-19.6</td>
<td>-7.2</td>
<td>-11.7</td>
</tr>
<tr>
<td></td>
<td>(3.6)</td>
<td>(3.5)</td>
<td>(3.8)</td>
<td>(3.9)</td>
</tr>
<tr>
<td>Included covariates:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hispanic indicator, and age</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Education</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
</tr>
<tr>
<td>AFQT</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>N</td>
<td>1,749</td>
<td>1,749</td>
<td>1,749</td>
<td>1,749</td>
</tr>
</tbody>
</table>

Base Specification

Full Specification
Adding Covariates Sequentially

**Panel B: Implied decomposition components**

<table>
<thead>
<tr>
<th>Wage difference attributed to:</th>
<th>Sequence in which covariates are added</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>First add: Education</td>
</tr>
<tr>
<td>Education completed as of 1990</td>
<td>AFQT -2.5</td>
</tr>
<tr>
<td>AFQT</td>
<td>Education -7.9</td>
</tr>
<tr>
<td>Total</td>
<td>AFQT -10.4</td>
</tr>
<tr>
<td></td>
<td>Education -14.9</td>
</tr>
<tr>
<td></td>
<td>Total -10.4</td>
</tr>
</tbody>
</table>
Adding Covariates Sequentially

1. Do not make an interpretation of the amount of variation in the treatment effect based on the sequence of the controls being added in. The sequence is going to be arbitrary.

2. Gelbach provides a method for empirically decomposing controls if you actually want to provide this kind of analysis.
Recap

1. Omitted Variable Bias is something that you must contend with as an empirical researcher and is a major threat to identification.

2. Simple heuristics, or eyeballing the coefficients, when adding controls variables is not very informative.

3. If you have a poorly measured confounder, it is better to run a balancing test with the confounder replacing the outcome variable rather than a coefficient comparison test.

4. If you want to interpret the change in coefficient values as you add more covariates, you cannot arbitrarily do so.