Education and Labor Market Outcomes

Sergey V. Popov

Centre for Advanced Studies
Higher School of Economics

April 26, 2011
Table of Contents

1 Introduction
   - About Me
   - Plans for today

2 Papers in 90s
   - Signalling vs investment
   - Employment While Studying
   - Better Schools Vs Worse Schools

3 Econometrics
   - Omitted Variable Bias
   - Instrument Variable

4 Papers in 2000s
   - Grades
   - Better schools vs less good schools
   - Working while studying

5 Conclusion
About Me

Me

- My name is Sergey, I do economics.
My name is Sergey, I do economics.
I have a PhD from University of Illinois.
My name is Sergey, I do economics.

I have a PhD from University of Illinois.

This is the first economics of education lecture I teach.
My name is Sergey, I do economics.
I have a PhD from University of Illinois.
This is the first economics of education lecture I teach.
I have a website on Google Sites,
https://sites.google.com/site/sergeyvpopov/
My name is Sergey, I do economics.

I have a PhD from University of Illinois.

This is the first economics of education lecture I teach.

I have a website on Google Sites,

https://sites.google.com/site/sergeyvpopov/

I have a cat.
Today

- Empirically:
  - bias due to missing control variable
Today

- **Empirically:**
  - bias due to missing control variable
  - how to fix it with *instruments*
Today

- **Empirically:**
  - bias due to missing control variable
  - how to fix it with *instruments*

- **Economics-wise:**
  - Is education *signaling/screening* or *investment*?
  - Does it pay off to go to a better school?
  - Does it help to work while studying?
Table of Contents

1 Introduction
   - About Me
   - Plans for today

2 Papers in 90s
   - Signalling vs investment
   - Employment While Studying
   - Better Schools Vs Worse Schools

3 Econometrics
   - Omitted Variable Bias
   - Instrument Variable

4 Papers in 2000s
   - Grades
   - Better schools vs less good schools
   - Working while studying

5 Conclusion
Signalling vs investment

Groot and Oosterbeek (1994)

- Take Netherlands education system
  - there are required years
  - one can retake a year
  - one can skip a year
  - one can take an extra unnecessary year
  - one can drop out and have no diploma
Take Netherlands education system
- there are required years
- one can retake a year
- one can skip a year
- one can take an extra unnecessary year
- one can drop out and have no diploma

If signaling wage effects
Signalling vs investment

Groot and Oosterbeek (1994)

- Take Netherlands education system
  - there are required years > 0
  - one can retake a year < 0
  - one can skip a year > 0
  - one can take an extra unnecessary year ≈ 0
  - one can drop out and have no diploma ≈ 0

- If signaling wage effects
Take Netherlands education system
- there are required years > 0
- one can retake a year < 0
- one can skip a year > 0
- one can take an extra unnecessary year ≈ 0
- one can drop out and have no diploma ≈ 0

If signaling wage effects
If investment wage effects
Signalling vs investment

Groot and Oosterbeek (1994)

- Take Netherlands education system
  - there are required years $> 0 > 0$
  - one can retake a year $< 0 > 0$
  - one can skip a year $> 0 < 0$
  - one can take an extra unnecessary year $\approx 0 \geq 0$
  - one can drop out and have no diploma $\approx 0 \geq 0$

- If signaling wage effects
- If investment wage effects
Signalling vs investment

Groot and Oosterbeek (1994)

- Take Netherlands education system
  - there are required years $> 0 > 0$
  - one can retake a year $< 0 > 0$
  - one can skip a year $> 0 < 0$
  - one can take an extra unnecessary year $\approx 0 > 0$
  - one can drop out and have no diploma $\approx 0 \geq 0$

- If signaling wage effects
- If investment wage effects
- 928 males, 289 females
Signalling vs investment

Groot and Oosterbeek (1994)

- Take Netherlands education system
  - there are required years $> 0 > 0$
  - one can retake a year $< 0 > 0$
  - one can skip a year $> 0 < 0$
  - one can take an extra unnecessary year $\approx 0 > 0$
  - one can drop out and have no diploma $\approx 0 \geq 0$

- If signaling wage effects
- If investment wage effects
- 928 males, 289 females
- Wages at 43 years old
Signalling vs investment

Groot and Oosterbeek (1994)

Take Netherlands education system
- there are required years $> 0 > 0$
- one can retake a year $< 0 > 0$
- one can skip a year $> 0 < 0$
- one can take an extra unnecessary year $≈ 0 ≥ 0$
- one can drop out and have no diploma $≈ 0 ≥ 0$

If signaling wage effects
If investment wage effects
928 males, 289 females
Wages at 43 years old around 20 years of experience!
Signalling vs investment

# Results

<table>
<thead>
<tr>
<th>Variable</th>
<th>Equation (1)</th>
<th>Equation (2)</th>
<th>Equation (3a)</th>
<th>Equation (3b)</th>
</tr>
</thead>
<tbody>
<tr>
<td>constant</td>
<td>1.339 (4.0)^a</td>
<td>1.305 (3.8)^a</td>
<td>1.295 (3.8)^a</td>
<td>1.653 (5.2)^a</td>
</tr>
<tr>
<td>$s_1$</td>
<td>0.043 (8.7)^a</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$s_2$</td>
<td></td>
<td>0.050 (8.3)^a</td>
<td>0.065 (9.1)^a</td>
<td>0.071 (10.2)^a</td>
</tr>
<tr>
<td>$s_3$</td>
<td></td>
<td></td>
<td>-0.089 (2.9)^a</td>
<td>-0.093 (3.0)^a</td>
</tr>
<tr>
<td>$s_4$</td>
<td></td>
<td></td>
<td>0.005 (0.2)</td>
<td>0.009 (0.5)</td>
</tr>
<tr>
<td>$s_5$</td>
<td></td>
<td></td>
<td>-0.014 (0.9)</td>
<td>-0.012 (0.8)</td>
</tr>
<tr>
<td>$s_6$</td>
<td></td>
<td></td>
<td>0.050 (4.3)^a</td>
<td>0.056 (4.9)^a</td>
</tr>
<tr>
<td>$iq/10$</td>
<td>0.040 (3.4)^a</td>
<td>0.046 (3.9)^a</td>
<td>0.037 (3.2)^a</td>
<td></td>
</tr>
<tr>
<td>experience</td>
<td>0.061 (2.2)^a</td>
<td>0.063 (2.3)^a</td>
<td>0.057 (2.1)^a</td>
<td>0.056 (2.0)^a</td>
</tr>
<tr>
<td>experience^2 / 10</td>
<td>-0.012 (2.0)^a</td>
<td>-0.013 (2.2)^a</td>
<td>-0.010 (1.7)</td>
<td>-0.010 (1.6)</td>
</tr>
</tbody>
</table>

Adjusted $R^2$ | 0.157 | 0.151 | 0.171 | 0.163 |

No. of cases   | 928 | 928 | 928 | 928 |

Note: $t$-values in parentheses.

^a Significant at the 5% level.
Results

Table 2.—Earnings Equations: Females

<table>
<thead>
<tr>
<th>Variable</th>
<th>Equation (1)</th>
<th>Equation (2)</th>
<th>Equation (3a)</th>
<th>Equation (3b)</th>
</tr>
</thead>
<tbody>
<tr>
<td>constant</td>
<td>1.245 (2.0)</td>
<td>1.504 (2.4)</td>
<td>1.619 (2.6)</td>
<td>1.886 (3.7)</td>
</tr>
<tr>
<td>$s_{da}$</td>
<td>0.083 (4.9)</td>
<td>0.111 (5.6)</td>
<td>0.123 (4.7)</td>
<td>0.130 (5.4)</td>
</tr>
<tr>
<td>$s_{e}$</td>
<td></td>
<td></td>
<td>-0.038 (0.4)</td>
<td>-0.032 (0.3)</td>
</tr>
<tr>
<td>$s_{s}$</td>
<td></td>
<td></td>
<td>-0.005 (0.1)</td>
<td>-0.005 (0.1)</td>
</tr>
<tr>
<td>$s_{r}$</td>
<td></td>
<td></td>
<td>0.014 (0.3)</td>
<td>0.012 (0.2)</td>
</tr>
<tr>
<td>$s_{i}$</td>
<td></td>
<td></td>
<td>0.093 (1.8)</td>
<td>0.098 (1.9)</td>
</tr>
<tr>
<td>$s_{d}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>iq/10</td>
<td>0.053 (1.3)</td>
<td>0.036 (0.9)</td>
<td>0.030 (0.0)</td>
<td></td>
</tr>
<tr>
<td>experience</td>
<td>0.058 (1.0)</td>
<td>0.044 (0.8)</td>
<td>0.023 (0.4)</td>
<td>0.025 (0.4)</td>
</tr>
<tr>
<td>experience^2/10</td>
<td>-0.024 (1.7)</td>
<td>-0.020 (1.4)</td>
<td>-0.014 (1.0)</td>
<td>-0.014 (1.0)</td>
</tr>
<tr>
<td>Adjusted $R^2$</td>
<td>0.225</td>
<td>0.243</td>
<td>0.242</td>
<td>0.243</td>
</tr>
<tr>
<td>No. of cases</td>
<td>289</td>
<td>289</td>
<td>289</td>
<td>289</td>
</tr>
</tbody>
</table>

Note: t-values in parentheses.

a Significant at the 5% level.
Summary of Results

- Necessary years are useful.
- Skipped years hurt.
- Repeated years are harmless.
- Inefficient years are harmful but not significant.
- Dropout years are useful but not significant.
Summary of Results

- Necessary years are useful.
- Skipped years hurt.
- Repeated years are harmless.
- Inefficient years are harmful but not significant.
- Dropout years are useful but not significant.
- But do we know, who chooses to drop out?..
Ruhm (1997)

- High school students can choose to work.
Ruhm (1997)

- High school students can choose to work.
- Get experience, earn more.
Employment While Studying

Ruhm (1997)

- High school students can choose to work.
- Get experience, earn more.
- Underinvestment in human capital, earn less.
Ruhm (1997)

- High school students can choose to work.
- Get experience, earn more.
- Underinvestment in human capital, earn less.

Data

- 1149 observations
- Boys work more frequently, and more.
- Whites have a job more frequently.
- Older students work more.
- New covariates: if parents are foreign; whether went to library in childhood; marijuana smoking...
Ruhm (1997)

- High school students can choose to work.
- Get experience, earn more.
- Underinvestment in human capital, earn less.

Data

- 1149 observations
- Boys work more frequently, and more.
- Whites have a job more frequently.
- Older students work more.
- New covariates: if parents are foreign; whether went to library in childhood; marijuana smoking...

- Does wage later in life depend on high school working experience?
## Results

### Table 4
Regression Estimates of Log Earnings on High School Employment Hours

<table>
<thead>
<tr>
<th>Type of Employment</th>
<th>(a)</th>
<th>(b)</th>
<th>(c)</th>
<th>(d)</th>
<th>(e)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Employment hours in reference week:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sophomore hours:</td>
<td>.0171</td>
<td>.0055</td>
<td>.0013</td>
<td>.0016</td>
<td>.0005</td>
</tr>
<tr>
<td>Hours$^{2/10}$</td>
<td>(.0095)</td>
<td>(.0099)</td>
<td>(.0091)</td>
<td>(.0091)</td>
<td>(.0091)</td>
</tr>
<tr>
<td>Junior hours:</td>
<td>.0202</td>
<td>.0100</td>
<td>.0050</td>
<td>.0013</td>
<td>.0013</td>
</tr>
<tr>
<td>Hours$^{2/10}$</td>
<td>(.0090)</td>
<td>(.0096)</td>
<td>(.0088)</td>
<td>(.0088)</td>
<td>(.0088)</td>
</tr>
<tr>
<td>Senior hours:</td>
<td>.0238</td>
<td>.0210</td>
<td>.0185</td>
<td>.0178</td>
<td>.0168</td>
</tr>
<tr>
<td>Hours$^{2/10}$</td>
<td>(.0068)</td>
<td>(.0071)</td>
<td>(.0066)</td>
<td>(.0066)</td>
<td>(.0066)</td>
</tr>
<tr>
<td>$p$-value</td>
<td>[.000]</td>
<td>[.008]</td>
<td>[.007]</td>
<td>[.014]</td>
<td>[.019]</td>
</tr>
<tr>
<td>$R^2$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.022</td>
<td>.208</td>
<td>.252</td>
<td>.260</td>
<td></td>
</tr>
</tbody>
</table>
Results

- The employment at school increases the wage after school.
Results

- The employment at school increases the wage after school.
- Last year matters more than other years experience.
Results

- The employment at school increases the wage after school.
- Last year matters more than other years experience.
- But *who* works?..
Results

- The employment at school increases the wage after school.
- Last year matters more than other years experience.
- But *who* works?.. Ruhm (1997) deals with this question.
Table 7
Econometric Estimates of Future Labor Market Outcomes on Employment Hours in High School Reference Week and Other Covariates

<table>
<thead>
<tr>
<th>Reference Week Employment Hours</th>
<th>Annual Work Hours (a)</th>
<th>Annual Weeks Worked (b)</th>
<th>Hourly Wages (c)</th>
<th>Weekly Wages (d)</th>
<th>Hourly Compensation (e)</th>
<th>Duncan Occupation Index (f)</th>
<th>Employer Health Insurance (g)</th>
<th>Employer Pension Plan (h)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full sample:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Senior hours:</td>
<td>10.01</td>
<td>.4551</td>
<td>.0079</td>
<td>.0096</td>
<td>.0098</td>
<td>.3532</td>
<td>.0242</td>
<td>.0311</td>
</tr>
<tr>
<td>Hours²/10</td>
<td>(4.880)</td>
<td>(.1477)</td>
<td>(.0042)</td>
<td>(.0047)</td>
<td>(.0044)</td>
<td>(.1282)</td>
<td>(.0095)</td>
<td>(.0087)</td>
</tr>
<tr>
<td>Φ(·)</td>
<td>.979</td>
<td>.546</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Males:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Senior hours:</td>
<td>12.91</td>
<td>.4368</td>
<td>.0041</td>
<td>.0074</td>
<td>.0056</td>
<td>.2342</td>
<td>.0245</td>
<td>.0238</td>
</tr>
<tr>
<td>Hours²/10</td>
<td>(5.538)</td>
<td>(.1585)</td>
<td>(.0049)</td>
<td>(.0051)</td>
<td>(.0051)</td>
<td>(.1688)</td>
<td>(.0130)</td>
<td>(.0116)</td>
</tr>
<tr>
<td>Φ(·)</td>
<td>.979</td>
<td>.502</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Females:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Senior hours:</td>
<td>10.98</td>
<td>.5461</td>
<td>.0175</td>
<td>.0197</td>
<td>.0207</td>
<td>.5567</td>
<td>.0393</td>
<td>.0434</td>
</tr>
<tr>
<td>Hours²/10</td>
<td>(9.094)</td>
<td>(.2769)</td>
<td>(.0083)</td>
<td>(.0096)</td>
<td>(.0087)</td>
<td>(.2233)</td>
<td>(.0173)</td>
<td>(.0162)</td>
</tr>
<tr>
<td>Φ(·)</td>
<td>.972</td>
<td>.588</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

NOTE.—See tables 4 and 5, notes. Annual weeks and hours worked are estimated as tobit models, with Φ(·) indicating the predicted percentage of censored observations (estimated as the average value of Φ(χ/σ) in the single limit tobit case); 44 observations are left-censored at 0 hours and 394 are right-censored at 52 weeks. The wage, compensation, and Duncan score equations are estimated by ordinary least squares. Ordered probit models are used for employer health insurance and pension coverage. The dependent variables in these equations are equal to 0, 1, and 2 if the fringe benefit is provided at none, some, or all three of interview dates, respectively, with p-values obtained from likelihood-ratio tests. Sample sizes (in the top panel) are 1,048, 1,050, 979, 979, 977, 1,000, 961, and 957 for cols. a–h, respectively.
Better Schools Vs Worse Schools

Behrman et al. (1996)

■ Do better schools improve earnings?
Better Schools Vs Worse Schools

Behrman et al. (1996)

- Do better schools improve earnings?
- Just adding a dummy variable would be weird — we need to account for all the differences between people (Griliches, 1979; Ashenfelter and Krueger, 1994)
Better Schools Vs Worse Schools

Behrman et al. (1996)

- Do better schools improve earnings?
- Just adding a dummy variable would be weird — we need to account for all the differences between people (Griliches, 1979; Ashenfelter and Krueger, 1994)
- Twin female study: 403 MZ and 305 DZ.
Better Schools Vs Worse Schools

Behrman et al. (1996)

- Do better schools improve earnings?
- Just adding a dummy variable would be weird — we need to account for all the differences between people (Griliches, 1979; Ashenfelter and Krueger, 1994)
- Twin female study: 403 MZ and 305 DZ.
- A lot of school properties (CASPAR, Barron’s)
Better Schools Vs Worse Schools

Behrman et al. (1996)

- Do better schools improve earnings?
- Just adding a dummy variable would be weird — we need to account for all the differences between people (Griliches, 1979; Ashenfelter and Krueger, 1994)
- Twin female study: 403 MZ and 305 DZ.
- A lot of school properties (CASPAR, Barron’s)
- Looking for within-variation explaining income difference.
Better Schools Vs Worse Schools

### Results

Table 3.—Estimates of Log of Full-Time Earnings, by Estimation Procedure and Sample

<table>
<thead>
<tr>
<th>Estimation Procedure</th>
<th>OLS Within-sibling</th>
<th>Within-twin</th>
<th>BRT MZ and DZ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample</td>
<td>DZ twins</td>
<td>DZ twins</td>
<td>MZ twins</td>
</tr>
<tr>
<td>Individual Characteristic</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>School years</td>
<td>0.0789 (5.29)</td>
<td>0.0843 (4.21)</td>
<td>0.0750 (2.94)</td>
</tr>
<tr>
<td>Work experience</td>
<td>0.0154 (7.42)</td>
<td>0.0114 (4.36)</td>
<td>0.0163 (5.44)</td>
</tr>
<tr>
<td>College Characteristic</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Private</td>
<td>0.212 (2.22)</td>
<td>0.174 (1.29)</td>
<td>0.193 (2.12)</td>
</tr>
<tr>
<td>PhD granting</td>
<td>−0.0673 (0.70)</td>
<td>0.0796 (0.54)</td>
<td>0.379 (3.18)</td>
</tr>
<tr>
<td>Total enrollment</td>
<td>0.154 (0.77)</td>
<td>−0.319 (1.07)</td>
<td>−0.494 (1.76)</td>
</tr>
<tr>
<td>(×10⁻⁵)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average Full Professor salary (×10⁻³)</td>
<td>0.00888 (1.45)</td>
<td>0.0221 (2.32)</td>
<td>0.0158 (1.68)</td>
</tr>
<tr>
<td>Expenditure per student (×10⁻³)</td>
<td>0.00514 (2.34)</td>
<td>0.0143 (3.68)</td>
<td>−0.104 (2.61)</td>
</tr>
<tr>
<td>Total students per faculty</td>
<td>0.00729 (2.55)</td>
<td>0.00664 (1.91)</td>
<td>0.00259 (0.75)</td>
</tr>
</tbody>
</table>

Note: Specifications include indicator variables for absent faculty salary, students per faculty and all college characteristics.

* Absolute value of t-ratio corrected for the clustering of twins in households in first column.
Better Schools Vs Worse Schools

Results

- In general
  - Extra year adds 6% of earnings (vs 1% for the experience)
  - Private schools provide extra 20% of income
Better Schools Vs Worse Schools

Results

- In general
  - Extra year adds 6% of earnings (vs 1% for the experience)
  - Private schools provide extra 20% of income

- But only when we consider twins, we can see
  - PhD-granting institutions are better
  - Bigger schools are worse
  - No point in spending money on students
  - Expensive professors are useful
Better Schools Vs Worse Schools

Results

- In general
  - Extra year adds 6% of earnings (vs 1% for the experience)
  - Private schools provide extra 20% of income
- But only when we consider twins, we can see
  - PhD-granting institutions are *better*
  - Bigger schools are *worse*
  - No point in spending money on students
  - Expensive professors are *useful*
- But how come two twins get into different schools?..
# Table of Contents

1. Introduction
   - About Me
   - Plans for today

2. Papers in 90s
   - Signalling vs investment
   - Employment While Studying
   - Better Schools Vs Worse Schools

3. Econometrics
   - Omitted Variable Bias
   - Instrument Variable

4. Papers in 2000s
   - Grades
   - Better schools vs less good schools
   - Working while studying

5. Conclusion
Linear Regression, One Variable

\[ y_i = \beta_0 + \beta_1 x_i + \epsilon_i \Rightarrow \]

\[ \hat{\beta}_1 = \frac{\text{cov}(x_i, y_i)}{\text{cov}(x_i, x_i)} \]

\[ = \frac{\text{cov}(x_i, \beta_0 + \beta_1 x_i + \epsilon_i)}{\text{cov}(x_i, x_i)} \rightarrow \frac{\text{cov}(x_i, \beta_1 x_i) + \text{cov}(x_i, \epsilon_i)}{\text{cov}(x_i, x_i)} = \beta_1 \]
Noisy Regressor Bias

What if we observe $\tilde{x}_i = x_i + \eta_i$ instead of $x$?

\[ \hat{\beta}_1 = \frac{\hat{\text{cov}}(\tilde{x}_i, y_i)}{\hat{\text{cov}}(\tilde{x}_i, \tilde{x}_i)} = \]

\[ = \frac{\text{cov}(x_i + \eta_i, \beta_0 + \beta_1 x_i + \epsilon_i)}{\text{cov}(x_i + \eta_i, x_i + \eta_i)} \rightarrow \frac{\beta_1 \text{cov}(x_i, x_i) + 0}{\text{cov}(x_i, x_i) + \text{var}(\eta_i)} \]

\[ = \beta_1 \frac{\text{var}(x_i)}{\text{var}(x_i) + \text{var}(\eta_i)} < \beta_1 \]
Omitted Variable Bias

**Linear Regression, Two Variables**

\[ y_i = \beta_0 + \beta_1 x_i + \beta_2 \tilde{x}_i + \epsilon_i \]

but estimate

\[ y_i = \beta_0 + \beta_1 x_i + \epsilon_i \Rightarrow \]

\[ \hat{\beta}_1 = \frac{\text{cov}(x_i, y_i)}{\text{cov}(x_i, x_i)} = \]

\[ = \frac{\text{cov}(x_i, \beta_0 + \beta_1 x_i + \beta_2 \tilde{x}_i + \epsilon_i)}{\text{cov}(x_i, x_i)} \rightarrow \frac{\beta_1 \text{cov}(x_i, x_i) + \beta_2 \text{cov}(x_i, \tilde{x}_i) + 0}{\text{cov}(x_i, x_i)} \]

\[ = \beta_1 + \beta_2 \times \frac{\text{cov}(x_i, \tilde{x}_i)}{\text{var}(x_i)}. \]

Not a problem if \( x_i \) and \( \tilde{x}_i \) are independent.
Can we *make* $x$ and $\tilde{x}$ non-correlated?

- We know that residuals and fits are uncorrelated...
- What if we have a $z$ such that $\text{cov}(x, z) \neq 0$, but $\text{cov}(\tilde{x}, z) = 0$?
2-stage Least Squares

Let

\[ x_i = \gamma_0 + \gamma_1 z_i + \gamma_2 \tilde{x}_i + \nu_i \]

- **Stage 1:** estimate a regression of \( x \) on \( z \), get a fit
  \[ \hat{x}_i = \hat{\gamma}_0 + \hat{\gamma}_1 z_i \]
  \[ \hat{\gamma}_1 \rightarrow \gamma_1 = \frac{\text{cov}(x_i, z_i)}{\text{cov}(z_i, z_i)} \]

- **Stage 2:** use \( \hat{x} \) instead of \( x \) in the main regression
  \[ \hat{x} \] is not correlated with \( \tilde{x} \) by construction

- But do we get a consistent estimate?
2SLS Estimate

Let true regression be

$$y_i = \beta_0 + \beta_1 x_i + \beta_2 \tilde{x}_i + \epsilon_i$$

but we instead estimate

$$y_i = \hat{\beta}_0 + \hat{\beta}_1 \hat{x}_i$$

$$\hat{\beta}_1 = \frac{\text{cov}(y_i, \hat{x}_i)}{\text{cov}(\hat{x}_i, \hat{x}_i)} \rightarrow \frac{\beta_1 \text{cov}(x_i, \hat{x}_i) + \beta_2 \text{cov}(\tilde{x}_i, \hat{x}_i)}{\text{var}(\hat{x}_i)} = 0$$

$$= \beta_1 \frac{\text{cov}(\gamma_0 + \gamma_1 z_i + \gamma_2 \tilde{x}_i, \hat{x}_i)}{\text{var}(\hat{x}_i)} = \beta_1$$

No problem from noisy $x$ variable!
Take-home message

- There is an omitted variable bias.
- If you have an instrument for the regressor of interest, you can use it to get rid of this bias.
- Estimates become consistent.
- Problems:
  - Getting a good instrument.
  - Estimates become *less efficient*, confidence intervals become *wider*.
Table of Contents

1 Introduction
   ■ About Me
   ■ Plans for today

2 Papers in 90s
   ■ Signalling vs investment
   ■ Employment While Studying
   ■ Better Schools Vs Worse Schools

3 Econometrics
   ■ Omitted Variable Bias
   ■ Instrument Variable

4 Papers in 2000s
   ■ Grades
   ■ Better schools vs less good schools
   ■ Working while studying

5 Conclusion
Gemus (2009)

Do grades improve your wage afterwards?
Gemus (2009)

- Do grades improve your wage afterwards?
- Cross-sectional data: grades at school and in the university, wages at graduation and 10 years after.
# Results: GPA affects wages

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>GPA</td>
<td>0.079</td>
<td>0.083</td>
<td>0.090</td>
<td>0.094</td>
</tr>
<tr>
<td></td>
<td>[0.028]***</td>
<td>[0.023]***</td>
<td>[0.023]***</td>
<td>[0.023]***</td>
</tr>
<tr>
<td>Second Tier</td>
<td>-0.058</td>
<td>0.005</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.067]</td>
<td>[0.068]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Third Tier</td>
<td>-0.129</td>
<td>-0.057</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.062]**</td>
<td>[0.064]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fourth Tier</td>
<td>-0.195</td>
<td>-0.115</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.066]***</td>
<td>[0.066]*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fifth Tier</td>
<td>-0.201</td>
<td>-0.100</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.071]***</td>
<td>[0.073]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sixth Tier</td>
<td>-0.300</td>
<td>-0.183</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.063]***</td>
<td>[0.065]***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Family Income</td>
<td>0.023</td>
<td>0.030</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.011]**</td>
<td>[0.011]***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>0.282</td>
<td>0.281</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.024]***</td>
<td>[0.025]***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>0.012</td>
<td>-0.024</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.035]</td>
<td>[0.037]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SAT/ACT Score</td>
<td>0.001</td>
<td>0.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.001]</td>
<td>[0.000]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experience</td>
<td>0.139</td>
<td>0.125</td>
<td>0.117</td>
<td>0.110</td>
</tr>
<tr>
<td></td>
<td>[0.049]***</td>
<td>[0.046]***</td>
<td>[0.044]***</td>
<td>[0.043]**</td>
</tr>
<tr>
<td>Experience(^2)</td>
<td>-0.003</td>
<td>-0.002</td>
<td>-0.003</td>
<td>-0.002</td>
</tr>
<tr>
<td></td>
<td>[0.003]</td>
<td>[0.003]</td>
<td>[0.003]</td>
<td>[0.003]</td>
</tr>
<tr>
<td></td>
<td>[0.212]***</td>
<td>[0.204]***</td>
<td>[0.354]***</td>
<td>[0.355]***</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.06</td>
<td>0.14</td>
<td>0.19</td>
<td>0.29</td>
</tr>
</tbody>
</table>
### Results: Richer get lower grades

#### Table 3: Relationship Between GPA and Background Characteristics

<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>25th-75 Income Percentile</td>
<td>-0.051</td>
<td>-0.050</td>
<td>-0.048</td>
<td>-0.049</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.026]**</td>
<td>[0.026]*</td>
<td>[0.025]*</td>
<td>[0.025]*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Top Income Quartile</td>
<td>-0.061</td>
<td>-0.056</td>
<td>-0.051</td>
<td>-0.051</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.025]**</td>
<td>[0.026]**</td>
<td>[0.026]**</td>
<td>[0.026]**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Family Income ($50,000)</td>
<td>-0.162</td>
<td>-0.162</td>
<td>-0.136</td>
<td>-0.136</td>
<td>-0.17</td>
<td>-0.09</td>
</tr>
<tr>
<td></td>
<td>[0.019]***</td>
<td>[0.019]***</td>
<td>[0.019]***</td>
<td>[0.019]***</td>
<td>[0.018]***</td>
<td>[0.019]***</td>
</tr>
<tr>
<td>Male</td>
<td>0.006</td>
<td>0.006</td>
<td>0.007</td>
<td>0.007</td>
<td>0.007</td>
<td>0.008</td>
</tr>
<tr>
<td></td>
<td>[0.000]***</td>
<td>[0.000]***</td>
<td>[0.000]***</td>
<td>[0.000]***</td>
<td>[0.000]***</td>
<td>[0.000]***</td>
</tr>
<tr>
<td>SAT/ACT Score</td>
<td>-0.097</td>
<td>-0.098</td>
<td>-0.104</td>
<td>-0.104</td>
<td>-0.103</td>
<td>-0.114</td>
</tr>
<tr>
<td></td>
<td>[0.013]***</td>
<td>[0.013]***</td>
<td>[0.014]***</td>
<td>[0.013]***</td>
<td>[0.013]***</td>
<td>[0.011]***</td>
</tr>
<tr>
<td>Age at Graduation</td>
<td>0.181</td>
<td>0.181</td>
<td>0.163</td>
<td>0.162</td>
<td>0.161</td>
<td>0.137</td>
</tr>
<tr>
<td></td>
<td>[0.033]***</td>
<td>[0.033]***</td>
<td>[0.033]***</td>
<td>[0.033]***</td>
<td>[0.032]***</td>
<td>[0.032]***</td>
</tr>
</tbody>
</table>

Note: Standard errors clustered by FICE code are in brackets. * Significant at the 10% level, ** at the 5% level, *** at the 1% level. First Tier is the omitted category for school rank dummies. BA Major categories are Education, Engineering, Health Professions, Public Affairs, Biological Sciences, Math/Science, Social Science, History, Humanities, Psychology, and Other. Business and Management is the excluded category. All regressions are weighted by the 2003 sample weight.
## Results: Getting a degree matters

<table>
<thead>
<tr>
<th>Degree Type</th>
<th>(1) OLS</th>
<th>(2) OLS</th>
<th>(3) FE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Master's Degree</td>
<td>0.320</td>
<td>0.076</td>
<td>0.237</td>
</tr>
<tr>
<td></td>
<td>[0.028]**</td>
<td>[0.027]**</td>
<td>[0.037]**</td>
</tr>
<tr>
<td>Professional Degree</td>
<td>0.582</td>
<td>0.393</td>
<td>0.626</td>
</tr>
<tr>
<td></td>
<td>[0.058]**</td>
<td>[0.057]**</td>
<td>[0.114]**</td>
</tr>
<tr>
<td>Doctoral Degree</td>
<td>0.575</td>
<td>0.253</td>
<td>0.692</td>
</tr>
<tr>
<td></td>
<td>[0.061]**</td>
<td>[0.054]**</td>
<td>[0.103]**</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.04</td>
<td>0.37</td>
<td>0.51</td>
</tr>
<tr>
<td>School Rank Controls</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Major Controls</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Background Controls</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Person Fixed Effect</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
</tbody>
</table>

N=4397

Note: Standard errors, clustered by FICE in Column (1) and by person in Columns (2) - (4), are in brackets.

* Significant at the 10% level, ** at the 5% level, *** at the 1% level. First Tier is the omitted category for school rank dummies. BA Major categories are Education, Engineering, Health Professions, Public Affairs, Biological Sciences, Math/Science, Social Science, History, Humanities, Psychology, and Other. Business and Management is the excluded category. Background variables include mother's and father's college attainment, family income, race, gender, SAT/ACT scores, and age. Column (1) results are weighted by the 2003 sample weight and Columns (2) - (4) are weighted by the panel weights.
Results: Getting a degree matters

Table 8: Alternative Specifications and Samples

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>GPA</td>
<td>0.044</td>
<td>0.001</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>[0.051]</td>
<td>[0.023]</td>
<td>[0.022]</td>
</tr>
<tr>
<td>GPA*Science Major (Undergrad)</td>
<td>0.114</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.131]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GPA*No Graduate Degree</td>
<td>0.052</td>
<td>0.054</td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.009]***</td>
<td>[0.008]***</td>
<td></td>
</tr>
<tr>
<td>GPA*Education Graduate Degree</td>
<td></td>
<td>-0.031</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>[0.018]*</td>
<td></td>
</tr>
<tr>
<td>R-squared</td>
<td>0.13</td>
<td>0.15</td>
<td>0.15</td>
</tr>
<tr>
<td>Number of Observations</td>
<td>1290</td>
<td>3858</td>
<td>4577</td>
</tr>
</tbody>
</table>
| Sample         | Graduate Degree Holders | Students from 3rd through 6th tier universities | Whole Sample
People who are very different should have different wages.

Hoekstra (2009)
Better schools vs less good schools

Hoekstra (2009)

- People who are very different should have different wages.
- People who are very similar should not.
Better schools vs less good schools

**Hoekstra (2009)**

- People who are very different should have different wages.
- People who are very similar should not.
- Flagship state university admits based on SAT (USE).
Better schools vs less good schools

**Hoekstra (2009)**

- People who are very different should have different wages.
- People who are very similar should not.
- Flagship state university admits based on SAT (USE).
- Take people who are a bit above and people who are a bit below, compare their earnings.
Better schools vs less good schools

Hoekstra (2009)

- People who are very different should have different wages.
- People who are very similar should not.
- Flagship state university admits based on SAT (USE).
- Take people who are a bit above and people who are a bit below, compare their earnings.
- Get 20% wage difference.
Better schools vs less good schools

Dale and Krueger (2011)

- US students apply to a lot of schools
- After they get accepted to some, they choose which to attend
Better schools vs less good schools

Dale and Krueger (2011)

- US students apply to a lot of schools
- After they get accepted to some, they choose which to attend
- Dale and Krueger have data on where else students got accepted.

Data
- 1976 and 1989 university attendants, earning money in 2000s
- 12075/6479 observations
- a lot of controls (SAT, high school grade)
<table>
<thead>
<tr>
<th>Variable</th>
<th>1976 Cohort</th>
<th>1989 Cohort</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Basic</td>
<td>Self-Revelation</td>
</tr>
<tr>
<td>School SAT Score/100</td>
<td>0.061</td>
<td>-0.023</td>
</tr>
<tr>
<td></td>
<td>(.013)</td>
<td>(.014)</td>
</tr>
<tr>
<td>Student SAT Score/100</td>
<td>0.022</td>
<td>0.014</td>
</tr>
<tr>
<td></td>
<td>(.005)</td>
<td>(.005)</td>
</tr>
<tr>
<td>Student SAT missing</td>
<td>-0.141</td>
<td>-0.122</td>
</tr>
<tr>
<td></td>
<td>(.030)</td>
<td>(.030)</td>
</tr>
<tr>
<td>Female</td>
<td>-0.479</td>
<td>-0.469</td>
</tr>
<tr>
<td></td>
<td>(.013)</td>
<td>(.013)</td>
</tr>
<tr>
<td>Black</td>
<td>-0.028</td>
<td>-0.037</td>
</tr>
<tr>
<td></td>
<td>(.029)</td>
<td>(.029)</td>
</tr>
<tr>
<td>Hispanic</td>
<td>-0.063</td>
<td>-0.077</td>
</tr>
<tr>
<td></td>
<td>(.069)</td>
<td>(.069)</td>
</tr>
<tr>
<td>Asian</td>
<td>0.171</td>
<td>0.151</td>
</tr>
<tr>
<td></td>
<td>(.046)</td>
<td>(.046)</td>
</tr>
<tr>
<td>Average SAT Score/100 of Schools</td>
<td>0.100</td>
<td></td>
</tr>
<tr>
<td>Applied to</td>
<td>(.012)</td>
<td></td>
</tr>
<tr>
<td>One Additional Application</td>
<td>0.062</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.017)</td>
<td></td>
</tr>
<tr>
<td>Two Additional Applications</td>
<td>0.057</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.018)</td>
<td></td>
</tr>
<tr>
<td>Three Additional Applications</td>
<td>0.073</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.020)</td>
<td></td>
</tr>
<tr>
<td>Four Additional Applications</td>
<td>0.085</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.034)</td>
<td></td>
</tr>
<tr>
<td>R-Squared</td>
<td>0.147</td>
<td>0.153</td>
</tr>
<tr>
<td>Sample Size (Unweighted)</td>
<td>12,075</td>
<td></td>
</tr>
</tbody>
</table>
Results

- Without controlling for other schools, +6%.
Better schools vs less good schools

Results

- Without controlling for other schools, +6%.
- With controlling, zero or possibly negative.
Better schools vs less good schools

Results

- Without controlling for other schools, $+6\%$.
- With controlling, zero or possibly negative.
- Omitted variable bias at its best.
Better schools vs less good schools

Results

- Without controlling for other schools, $+6\%$.
- With controlling, zero or possibly negative.
- Omitted variable bias at its best.
- Same story for Behrman et al. (1996); Hoekstra (2009).
Results

- Without controlling for other schools, +6%.
- With controlling, zero or possibly negative.
- Omitted variable bias at its best.
- Same story for Behrman et al. (1996); Hoekstra (2009).
  ... adjust their 20% bonus is too high.
Working while studying

Stinebrickner and Stinebrickner (2003)

- Data on 2372 students of Berea College, where everyone works.
- They have grades, wages and hours.
Working while studying

Stinebrickner and Stinebrickner (2003)

- Data on 2372 students of Berea College, where everyone works.
- They have grades, wages and hours.
- They also have *job assignments*, which can be an instrument for work hours.
Working while studying

Results before instruments

Table 1
Descriptive statistics and OLS regression for 1st and 2nd semester GPA
first semester and second semester separately

<table>
<thead>
<tr>
<th></th>
<th>data</th>
<th>1st semester GPA</th>
<th>2nd semester GPA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>mean std dev</td>
<td>estimate (std error)</td>
<td>estimate (std error)</td>
</tr>
<tr>
<td>weekly work hours</td>
<td></td>
<td>.043 (.010)**</td>
<td>.021 (.006)**</td>
</tr>
<tr>
<td>black</td>
<td>.10</td>
<td>-.183 (.053)**</td>
<td>-.185 (.051)**</td>
</tr>
<tr>
<td>male</td>
<td>.46</td>
<td>-.154 (.031)**</td>
<td>-.117 (.030)**</td>
</tr>
<tr>
<td>math ACT</td>
<td>22.35 4.36</td>
<td>.038 (.004)**</td>
<td>.036 (.004)**</td>
</tr>
<tr>
<td>verbal ACT</td>
<td>20.51 3.92</td>
<td>.036 (.004)**</td>
<td>.039 (.004)**</td>
</tr>
<tr>
<td>constant</td>
<td></td>
<td>.614 (.153)**</td>
<td>.890 (.111)**</td>
</tr>
<tr>
<td>Estimated Variance Error</td>
<td></td>
<td>.547</td>
<td>.451</td>
</tr>
<tr>
<td>Estimated Variance Error</td>
<td></td>
<td>R^2 = .145</td>
<td>R^2 = .167</td>
</tr>
</tbody>
</table>
### Table 4
Instrumental Variables Estimates for GPA

<table>
<thead>
<tr>
<th></th>
<th>first semester GPA</th>
</tr>
</thead>
<tbody>
<tr>
<td>weekly hours</td>
<td>-.162 (.050)**</td>
</tr>
<tr>
<td>black</td>
<td>-.202 (.057)**</td>
</tr>
<tr>
<td>male</td>
<td>-.172 (.034)**</td>
</tr>
<tr>
<td>math ACT</td>
<td>.039 (.004)**</td>
</tr>
<tr>
<td>verbal ACT</td>
<td>.038 (.004)**</td>
</tr>
<tr>
<td>constant</td>
<td>2.82 (.558)**</td>
</tr>
</tbody>
</table>

* n=2372
* t statistic greater than 1.65
** t statistic greater than 2.0
Results

- Some students work more, same students exhibit better grades.
- One needs a way of measuring of the working capacity of the person to get the estimate of the marginal effect of an extra hour.
- Extra hour worked *hurts* your investment in your human capital.
Results

- Some students work more, same students exhibit better grades.
- One needs a way of measuring of the working capacity of the person to get the estimate of the marginal effect of an extra hour.
- Extra hour worked *hurts* your investment in your human capital.

... that is, if you studied in Berea College...
Results

- Some students work more, same students exhibit better grades.
- One needs a way of measuring of the working capacity of the person to get the estimate of the marginal effect of an extra hour.
- Extra hour worked *hurts* your investment in your human capital.
  
  … that is, if you studied in Berea College...
- Ruhm (1997) has no measure of working capacity...
Table of Contents

1 Introduction
   ■ About Me
   ■ Plans for today

2 Papers in 90s
   ■ Signalling vs investment
   ■ Employment While Studying
   ■ Better Schools Vs Worse Schools

3 Econometrics
   ■ Omitted Variable Bias
   ■ Instrument Variable

4 Papers in 2000s
   ■ Grades
   ■ Better schools vs less good schools
   ■ Working while studying

5 Conclusion
Conclusion

- Even the most trivial questions are hard to answer.
Conclusion

- Even the most trivial questions are hard to answer.
- Does higher GPA improve your wage? Yes — if you don’t get to be a doctor.
Conclusion

- Even the most trivial questions are hard to answer.
- Does higher GPA improve your wage? Yes — if you don’t get to be a doctor.
- Does working while studying improve your wage?
Conclusion

- Even the most trivial questions are hard to answer.
- Does higher GPA improve your wage? Yes — if you don’t get to be a doctor.
- Does working while studying improve your wage? *Maybe*, but it makes your GPA lower.
Even the most trivial questions are hard to answer.

Does higher GPA improve your wage? Yes — if you don’t get to be a doctor.

Does working while studying improve your wage? *Maybe*, but it makes your GPA lower.

Does a better school make you a better worker?
Conclusion

- Even the most trivial questions are hard to answer.
- Does higher GPA improve your wage? Yes — if you don’t get to be a doctor.
- Does working while studying improve your wage? *Maybe*, but it makes your GPA lower.
- Does a better school make you a better worker? *Maybe*. 
Conclusion

- Even the most trivial questions are hard to answer.
- Does higher GPA improve your wage? Yes — if you don’t get to be a doctor.
- Does working while studying improve your wage? *Maybe*, but it makes your GPA lower.
- Does a better school make you a better worker? *Maybe*.
- So the optimal


