Economic Policy Analysis (EC406)

The Program Evaluation Problem

John Van Reenen

Centre for Economic Performance

London School of Economics

* These lecture notes are incomplete. Reading the lecture notes without attending lectures can be very misleading.
The Program

Twelve months ago a developing country Labas introduced an anti-poverty program PROSCOL
– Provides cash transfers to poor families with school-age children
– To be eligible for the transfer, households must have certain observable characteristics that suggest they are poor
– To continue receiving the transfer, they must keep their children in school until 18 years of age

Program’s objectives
– Reduce current poverty through cash transfers
– Reduce future poverty through increased schooling

How evaluate whether the program is effective in reducing poverty?
– How much is the program increasing school enrollment rates?
The Program Evaluation Problem

Let $d_i$ denote program participation of the $i^{\text{th}}$ individual
- $d_i = 1$ when individual $i$ participates
- $d_i = 0$ when she does not

Denote individual $i$’s schooling by
- $S_{1i}$ when $d_i = 1$
- $S_{0i}$ when $d_i = 0$

The gain in schooling due to the program for an individual $i$ who does in fact participate is

$$G_i = S_{1i} - S_{0i} \mid d_i = 1$$

Causal effect of program? Average effect of the treatment on the treated:

$$G = E(S_{1i} - S_{0i} \mid d_i = 1)$$
The Program Evaluation Problem

**Evaluation problem:** do not typically observe the *counterfactual*, \( S_{0i} \mid d_i = 1 \) (in this sense, program evaluation problem is a missing data problem)

We do observe the difference in mean outcomes between those who participate in the program and those who do not:

\[
D = E(S_{1i} \mid d_i = 1) - E(S_{0i} \mid d_i = 0)
\]

This is a biased estimate of average effect of the treatment on the treated

\[
D = G + B
\]

\[
B = E(S_{0i} \mid d_i = 1) - E(S_{0i} \mid d_i = 0)
\]

The bias is the expected difference in schooling in the absence of the program between those who participated and those who did not
Alternative Approaches to Program Evaluation

Treatment group
- Those who participated in the program

Control or comparison group
- Used to identify the counterfactual of what would have happened without the program
- Intended to be representative of the treatment group except for the fact that the control group did not participate in the program

(1) Experimental Methods (Randomization)
- Individuals from a well-defined group are randomly allocated to treatment and control groups
- No difference in expectation between the groups besides the fact that the treatment group received the program
  - Remain potential problems (eg sample attrition)
  - Example: reservation for women of head positions in Indian Village Councils (Chattopadhay and Duflo 2003)
Alternative Approaches to Program Evaluation

(2) Non-experimental or quasi-experimental methods
- When the allocation of individuals to treatment and control groups is non-random this gives rise to potential problems of
  - Selection on observables
  - Selection on unobservables (often referred as the econometric selection problem)
- Alternative non-experimental methods seek to address these problems in different ways

(2A) Selection Models
- Explicitly control for non-random selection of individuals into the treatment and control group (Heckman 1979)

(2B) Instrumental Variables Approaches
- Instrumental variable correlated with treatment but has no independent effect on the outcome of interest
- Example: compulsory schooling laws, Angrist and Krueger 1991
Alternative Approaches to Program Evaluation

(2C) Difference in Differences
  – Compare a treatment and control group (first difference) both before and after the program (second difference)

(2D) Matching
  – Construct a control group that resembles the treatment group as closely as possible in terms of observed characteristics
  – Example: National Supported Work (NSW) labour training program in the US, Dehejia and Wahba (1999)

(2E) Matching and Difference in Differences
  – Combine the above two methodologies
The Program Evaluation Problem

*Homogenous Treatment Effects and Linear Model*

Assume homogenous treatment effects
- The treatment has the same effect across individuals

Suppose that the program takes place at date $k$
- $d_i = 1$ if an individual participates in the program
- $d_i = 0$ otherwise

Consider the following linear regression model

$$y_{it} = \beta_0 + \alpha d_i + u_{it} \quad \text{if } t > k$$

$$y_{it} = \beta_0 + u_{it} \quad \text{if } t \leq k$$

Where, for simplicity, there are no control variables
The Program Evaluation Problem

*Homogeneous Treatment Effects and Linear Model*

Under the assumption that $E(u|d) = 0$

- Ensured by randomization

\[ E(y | d = 1) - E(y | d = 0) = \alpha \]

OLS yields unbiased and consistent estimates of the treatment effect (average effect of the treatment on the treated)

\[ E(\hat{\alpha}^{OLS}) = \alpha \]

If $E(u|d) \neq 0$

- As will typically be the case in non-experimental contexts

\[ E(y | d = 1) - E(y | d = 0) = \alpha + E(u | d = 1) - E(u | d = 0) \]

OLS yields biased and inconsistent estimates of the treatment effect

\[ E(\hat{\alpha}^{OLS}) = \alpha + E(u | d = 1) - E(u | d = 0) \]
Controlling for Observables and Modelling Participation

Control for other determinants of outcomes besides the program

\[ y_{it} = X_{it} \beta + d_i \alpha + u_{it} \quad \text{if} \quad t > k \]

\[ y_{it} = X_{it} \beta + u_{it} \quad \text{if} \quad t \leq k \]

Consider an explicit index model of the individual’s (non-random) participation decision

\[ IN_i = Z_i \gamma + v_i \]

\[ d_i = 1 \quad \text{if} \quad IN_i > 0 \]

\[ d_i = 0 \quad \text{otherwise} \]
Controlling for Observables and Modelling Participation

Special case where estimating the outcome equation using OLS yields unbiased and consistent estimates of the treatment effect

- $X$ includes all the variables in $Z$ that also influence program participation and $E(\epsilon uv) = 0$ (selection on observables)
- The participation decision is linear in $Z$

(Suggests the importance of controlling for other determinants of outcomes in program evaluation)

More generally, even after controlling for other determinants in the outcome equation, there remains a sample selection problem

- When $E(\epsilon uv) \neq 0$, estimating the outcome equation using OLS yields biased and inconsistent estimates of the treatment effect
(2A) Heckman Selection Model (‘Heckit’)

Control directly for that part of the error term in the outcome equation \( (u) \) that is correlated with the participation dummy variable \( (d) \)

To do so, impose additional structure on the model:
- \( Z \) is exogenous in the outcome equation, \( E(u|X,Z) = 0 \)
- \( X \) is a strict subset of \( Z \)
  - Implies there is at least one additional regressor in the participation decision
- \( u_{it} \) and \( v_i \) are jointly normally distributed and, for simplicity, consider the standardization \( \sigma_v = I \)
(2A) Heckman Selection Model (‘Heckit’)

Under these assumptions

\[
E(y_{it} | d_i = 1) = X_{it} \beta + \alpha + \rho \frac{\phi(Z_i \gamma)}{\Phi(Z_i \gamma)}
\]

\[
E(y_{it} | d_i = 0) = X_{it} \beta - \rho \frac{\phi(Z_i \gamma)}{1 - \Phi(Z_i \gamma)}
\]

Final term on the right-hand side corresponds to the expected value of the error term \((u)\) conditional on the participation variable \((d)\)

\(\phi(Z_i \gamma) / \Phi(Z_i \gamma)\) is the inverse Mills ratio
- Ratio between the standard normal pdf and standard normal cdf evaluated at \(Z_i \gamma\)

Including it in the outcome equation controls for non-random selection, enabling us to identify the treatment effect
Derivation of the Role of the Inverse Mills Ratio

Taking expectations in the outcome equation

\[ E(y_{it} | d_i = 1) = X_{it} \beta + \alpha + E(u_{it} | d_i = 1) \]

\[ E(y_{it} | d_i = 1) = X_{it} \beta + \alpha + E(u_{it} | v_i) E(v_i | d_i = 1) \]

\[ E(y_{it} | d_i = 1) = X_{it} \beta + \alpha + \rho E(v_i | d_i = 1) \]

Using properties of the normal distribution

\[ E(v_i | d_i = 1) = E(v_i | v_i > -Z_i \gamma) \]

\[ E(v_i | v_i > -Z_i \gamma) = \frac{\phi(-Z_i \gamma)}{1 - \Phi(-Z_i \gamma)} = \frac{\phi(Z_i \gamma)}{\Phi(Z_i \gamma)} \]
Implementing the ‘Heckit’ Sample Selection Correction

(1) Using all observations (those who participated in the program and those who did not)
- Estimate a Probit model of participation $d_i$ on $Z_i$ and obtain the parameter estimates $\hat{\gamma}$
  \[ \hat{IN}_i = Z_i \hat{\gamma} + v_i \]
  \[ d_i = 1 \text{ if } \hat{IN}_i > 0, \ 0 \text{ otherwise} \]
- Compute the inverse Mills ratio $\hat{\lambda}(Z_i \hat{\gamma}) = \phi(z_i \hat{\gamma}) / \Phi(z_i \hat{\gamma})$

(2) Using all observations (those who participated in the program and those who did not)
- Estimate the outcome equation using OLS
  \[ y_{it} = X_{it} \beta + d_i \alpha + \rho \hat{\lambda}(Z_i \hat{\gamma}) + \omega_{it} \]
- The estimated coefficients are consistent and approximately normally distributed
- Test for selection bias by testing the null hypothesis $\rho = 0$
Empirical Application: Earnings & Demand for College Education

Education and Self-Selection (Willis and Rosen, JPE, 1979)

(1) What is the effect of college education on earnings?
   - College education is the treatment
   - Control for selection bias in examining the effect of college education on earnings
   - What is the relationship between individual characteristics and earnings after controlling for selection bias?

(2) Examine the extent to which alternative earnings prospects, as opposed to family background and financial constraints, influence the decision to attend college
General Model of Education Decision

\[ Y_{ij} = y_j(X_i, \tau_i) \quad j = 1,...,n \]

\[ V_{ij} = g(y_j, Z_i, \omega_i) \quad (\tau, \omega) \sim F(\tau, \omega) \]

\( i \) belongs to \( j \) if \( V_{ij} = \max (V_{i1},...,V_{in}) \)

- \( Y_{ij} \): potential lifetime earnings of individual \( i \) if schooling level \( j \) is chosen
- \( X_i \): vector of observed individual characteristics
- \( \tau_i \): unobserved individual characteristics (ability)
- \( V_{ij} \): value of choosing schooling level \( j \) for individual \( i \)
- \( Z_i \): vector of observed individual characteristics
- \( \omega_i \): unobserved individual characteristics (tastes)
Econometric Model of College Education Decision

Two levels of schooling
- More than high school (level A)
- High school (level B)

If individual \( i \) chooses schooling level \( A \), expected earnings stream is

\[
y_{ai}(t) = 0, \quad 0 < t \leq S
\]

\[
y_{ai}(t) = \bar{y}_{ai} e^{g_{ai}(t-S)}, \quad S \leq t < \infty
\]

If individual \( i \) chooses schooling level \( B \), expected earnings stream is

\[
y_{bi}(t) = \bar{y}_{bi} e^{g_{bi}t}, \quad 0 \leq t < \infty
\]
Econometric Model of College Education Decision

If individual $i$ chooses schooling level $A$, the expected present value of the earnings stream is

$$V_{ai} = \int_{s}^{\infty} y_{ai}(t)e^{-r_i t} dt = \left(\frac{\bar{y}_{ai}}{r_i - g_{ai}}\right)e^{-r_i S}$$

If individual $i$ chooses schooling level $B$, the expected present value of the earnings stream is

$$V_{bi} = \int_{s}^{\infty} y_{bi}(t)e^{-r_i t} dt = \left(\frac{\bar{y}_{bi}}{r_i - g_{bi}}\right)$$

Assume individual $i$ chooses $A$ if $V_{ai} > V_{bi}$ and chooses $B$ if $V_{ai} \leq V_{bi}$
Define $I_i = \ln \left( \frac{V_{ai}}{V_{bi}} \right)$

$$I_i = \ln \bar{y}_{ai} - \ln \bar{y}_{bi} - r_i S - \ln (r_i - g_{ai}) + \ln (r_i - g_{bi})$$

Linearization

$$I_i = \alpha_0 + \alpha_1 (\ln \bar{y}_{ai} - \ln \bar{y}_{bi}) + \alpha_2 g_{ai} + \alpha_3 g_{bi} + \alpha_4 r_i \quad (S)$$

Selection rule becomes

$$\Pr(\text{choose A}) = \Pr(V_a > V_b) = \Pr(I > 0)$$

$$\Pr(\text{choose B}) = \Pr(V_a \leq V_b) = \Pr(I \leq 0)$$
Earnings, Growth and Discount Equations

\[
\ln \bar{y}_{ai} = X_i \beta_a + u_{1i} \quad (E1)
\]

\[
g_{ai} = X_i \gamma_a + u_{2i} \quad (G1)
\]

\[
\ln \bar{y}_{bi} = X_i \beta_b + u_{3i} \quad (E2)
\]

\[
g_{bi} = X_i \gamma_b + u_{4i} \quad (G2)
\]

\[
r_i = Z_i \delta + u_{5i} \quad (D)
\]

**Estimation**

– Substitute (D) in (S), variation in \(Z_i\) is used to separately identify the parameters of the earnings and selection equations
Econometric Estimation

Estimation Procedure

– Estimate the reduced form college selection equation
– Evaluate the inverse Mills ratio and include when estimating structural earnings and growth equations
– Generate predicted values from the structural earnings and growth equations
– Include these predicted values in the college selection equation to estimate the structural relationship
Data Description

NBER-Thorndike-Hagen survey, 1968-71
- 3,611 respondents
- Male WWII veterans who applied for the army air corps
  - Disadvantages
    - Non-random sample of the population
  - Advantages
    - Covers more than 20 years of labor-market experience
    - Extensive information on family background and talent
Data Description

Descriptive Statistics

- Mean and variance of earnings lower for high school graduates
- Fathers of high school graduates had less schooling and more likely to be manual workers
- High school graduates score lower in math and reading comprehension tests, about the same in manual dexterity, and somewhat better on mechanical ability (consistent with idea college education decision shaped by comparative advantage)
<table>
<thead>
<tr>
<th>VARIABLE</th>
<th>High School (Group B)</th>
<th>More than High School (Group A)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>Father's ED</td>
<td>8.671</td>
<td>2.966</td>
</tr>
<tr>
<td>Father's ED²</td>
<td>83.99</td>
<td>55.53</td>
</tr>
<tr>
<td>DK ED</td>
<td>.0999</td>
<td>...</td>
</tr>
<tr>
<td>Manager</td>
<td>.3628</td>
<td>...</td>
</tr>
<tr>
<td>Clerk</td>
<td>.1239</td>
<td>...</td>
</tr>
<tr>
<td>Foreman</td>
<td>.2238</td>
<td>...</td>
</tr>
<tr>
<td>Unskilled</td>
<td>.1492</td>
<td>...</td>
</tr>
<tr>
<td>Farmer</td>
<td>.1062</td>
<td>...</td>
</tr>
<tr>
<td>DK job</td>
<td>.0177</td>
<td>...</td>
</tr>
<tr>
<td>Catholic</td>
<td>.2993</td>
<td>...</td>
</tr>
<tr>
<td>Jew</td>
<td>.0405</td>
<td>...</td>
</tr>
<tr>
<td>Old sibs</td>
<td>1.143</td>
<td>1.634</td>
</tr>
<tr>
<td>Young sibs</td>
<td>.9381</td>
<td>1.486</td>
</tr>
<tr>
<td>Mother works:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Full 5</td>
<td>.0468</td>
<td>...</td>
</tr>
<tr>
<td>Part 5</td>
<td>.0392</td>
<td>...</td>
</tr>
<tr>
<td>None 5</td>
<td>.7168</td>
<td>...</td>
</tr>
<tr>
<td>Full 14</td>
<td>.0822</td>
<td>...</td>
</tr>
<tr>
<td>Part 14</td>
<td>.0708</td>
<td>...</td>
</tr>
<tr>
<td>None 14</td>
<td>.6384</td>
<td>...</td>
</tr>
<tr>
<td>H.S. shop</td>
<td>.2592</td>
<td>...</td>
</tr>
<tr>
<td>Read</td>
<td>20.57</td>
<td>10.17</td>
</tr>
<tr>
<td>NR read</td>
<td>.0291</td>
<td>...</td>
</tr>
<tr>
<td>Mech</td>
<td>59.24</td>
<td>18.27</td>
</tr>
<tr>
<td>NR mech</td>
<td>.0025</td>
<td>...</td>
</tr>
<tr>
<td>Math</td>
<td>18.13</td>
<td>11.82</td>
</tr>
<tr>
<td>NR math</td>
<td>.0683</td>
<td>...</td>
</tr>
<tr>
<td>Dext</td>
<td>50.04</td>
<td>9.359</td>
</tr>
<tr>
<td>NR dext</td>
<td>0</td>
<td>...</td>
</tr>
<tr>
<td>Exp</td>
<td>90.32</td>
<td>9.430</td>
</tr>
</tbody>
</table>
Reduced Form College Selection Equation

Reduced form Probit college selection equation
– Positive and significant effect of math score
– Negative and significant effect of mechanical score
– Positive and significant effect of father being a manager
– Negative and significant effect of father being a farmer
<table>
<thead>
<tr>
<th>VARIABLE</th>
<th>REDUCED FORM (16)</th>
<th></th>
<th>STRUCTURE (26)</th>
<th></th>
<th>STRUCTURE (29)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coefficient</td>
<td>$t$</td>
<td>Coefficient</td>
<td>$t$</td>
<td>Coefficient</td>
<td>$t$</td>
</tr>
<tr>
<td>Constant</td>
<td>.0485</td>
<td>.20</td>
<td>.1512</td>
<td>.22</td>
<td>.1030</td>
<td>.17</td>
</tr>
<tr>
<td>Background:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Father's ED</td>
<td>-.0145</td>
<td>-.41</td>
<td>-.0168</td>
<td>-.54</td>
<td>-.0152</td>
<td>-.49</td>
</tr>
<tr>
<td>Father's ED$^2$</td>
<td>.0037</td>
<td>2.05</td>
<td>.0038</td>
<td>2.26</td>
<td>.0037</td>
<td>2.26</td>
</tr>
<tr>
<td>DK ED</td>
<td>-.4059</td>
<td>-3.96</td>
<td>-.3924</td>
<td>-2.79</td>
<td>-.4001</td>
<td>-2.91</td>
</tr>
<tr>
<td>Manager</td>
<td>.1897</td>
<td>2.17</td>
<td>.1825</td>
<td>2.13</td>
<td>.1871</td>
<td>2.21</td>
</tr>
<tr>
<td>Clerk</td>
<td>.0556</td>
<td>.54</td>
<td>.0561</td>
<td>.59</td>
<td>.0554</td>
<td>.59</td>
</tr>
<tr>
<td>Foreman</td>
<td>.0182</td>
<td>.19</td>
<td>.0210</td>
<td>.23</td>
<td>.0200</td>
<td>.22</td>
</tr>
<tr>
<td>Unskilled</td>
<td>-.0910</td>
<td>-.85</td>
<td>-.0948</td>
<td>-.89</td>
<td>-.0928</td>
<td>-.87</td>
</tr>
<tr>
<td>Farmer</td>
<td>-.2039</td>
<td>-2.12</td>
<td>-.2256</td>
<td>-2.97</td>
<td>-.2094</td>
<td>-2.14</td>
</tr>
<tr>
<td>DK job</td>
<td>-.0413</td>
<td>-.19</td>
<td>-.0629</td>
<td>-.29</td>
<td>-.0609</td>
<td>-.28</td>
</tr>
<tr>
<td>Catholic</td>
<td>-.1144</td>
<td>-1.91</td>
<td>-.0982</td>
<td>-1.51</td>
<td>-.1083</td>
<td>-1.66</td>
</tr>
<tr>
<td>Jew</td>
<td>-.0293</td>
<td>-.23</td>
<td>.0143</td>
<td>.12</td>
<td>-.0158</td>
<td>-.14</td>
</tr>
<tr>
<td>Old sibs</td>
<td>-.0162</td>
<td>-.93</td>
<td>-.0162</td>
<td>-.93</td>
<td>-.0161</td>
<td>-.93</td>
</tr>
<tr>
<td>Young sibs</td>
<td>.0122</td>
<td>.63</td>
<td>.0096</td>
<td>.49</td>
<td>.0112</td>
<td>.57</td>
</tr>
<tr>
<td>Mother works:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Full 5</td>
<td>.1039</td>
<td>.66</td>
<td>.1168</td>
<td>.81</td>
<td>.1104</td>
<td>.76</td>
</tr>
<tr>
<td>Part 5</td>
<td>.2179</td>
<td>1.42</td>
<td>.2106</td>
<td>1.52</td>
<td>.2156</td>
<td>1.56</td>
</tr>
</tbody>
</table>
Table 2 (cont.) : College Selection Rule

<table>
<thead>
<tr>
<th>Ability</th>
<th>None 5</th>
<th>Full 14</th>
<th>Part 14</th>
<th>None 14</th>
<th>H.S. shop</th>
</tr>
</thead>
<tbody>
<tr>
<td>Read</td>
<td>.0655</td>
<td>.2898</td>
<td>.2709</td>
<td>.1980</td>
<td>-.4411</td>
</tr>
<tr>
<td>NR read</td>
<td>-.2575</td>
<td>-1.41</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>Mech</td>
<td>-.0070</td>
<td>-4.29</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>NR mech</td>
<td>-3.0236</td>
<td>-1.04</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>Math</td>
<td>.0244</td>
<td>12.34</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>NR math</td>
<td>-.7539</td>
<td>-5.75</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>Dext</td>
<td>.0019</td>
<td>.72</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>NR dext</td>
<td>2.2797</td>
<td>.47</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>Earnings:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>in ((\hat{v}_a\hat{y}_a))</td>
<td>...</td>
<td>...</td>
<td>5.1486</td>
<td>2.25</td>
<td>...</td>
</tr>
<tr>
<td>(\hat{g}_a)</td>
<td>...</td>
<td>...</td>
<td>138.3850</td>
<td>1.83</td>
<td>7.6632</td>
</tr>
<tr>
<td>(\hat{g}_e)</td>
<td>...</td>
<td>...</td>
<td>-44.2697</td>
<td>-1.28</td>
<td>71.8981</td>
</tr>
<tr>
<td>ln (y_a(t)/y_b(t))</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>5.1501</td>
</tr>
<tr>
<td>Observations</td>
<td>3611</td>
<td>3611</td>
<td>3611</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Limit observations</td>
<td>791</td>
<td>791</td>
<td>791</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nonlimit observations</td>
<td>2820</td>
<td>2820</td>
<td>2820</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-2 ln (likelihood ratio)</td>
<td>579.5</td>
<td>568.8</td>
<td>576.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\chi^2) degree freedom</td>
<td>98</td>
<td>93</td>
<td>23</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note.—t is asymptotic t-statistic; DK: Don’t know, dummy variable; NR: No response, dummy variable; other variables are defined in Appendix A.
Structural Earnings and Growth Equations

Determinants of initial earnings
- Experience
- Positive and significant effect of math score for college attendees

Determinants of earnings growth
- Positive and significant effect of reading and dexterity score for high school graduates
- Negative and significant effect of mechanical score for college attendees
Selection effects

- College attendees
  - Positive selection bias for initial earnings
  - Negative selection bias for earnings growth
- High school graduates
  - No selectivity bias for initial earnings
  - Positive selection bias for earnings growth

Positive selection bias for high school graduates earnings growth

- Observed earnings growth for high school graduates more rapid than would have been observed for the average member of the sample with the same $X$ characteristics had they chosen to remain a high school graduate
Selection and Comparative Advantage

Positive selection effects for both college attendees and high school graduates provide evidence of sorting according to comparative advantage.

Consider individuals with the same $X$ characteristics.

Empirical results on selection bias imply:

- Those who remained high school graduates had better prospects as high school graduates than the average person with the same $X$ characteristics.
- Those who attended college had better prospects as college attendees than the average person with the same $X$ characteristics.

Important implications for the rate of return to college education.
## Table 3
**Structural Earnings Estimates: Equations (24) and (28), OLS**

<table>
<thead>
<tr>
<th>Dependent Variable</th>
<th>( \ln \tilde{y}_a )</th>
<th>( \ln \tilde{y}_b )</th>
<th>( g_a )</th>
<th>( g_b )</th>
<th>( \ln y_a(\tilde{y}) )</th>
<th>( \ln y_b(\tilde{y}) )</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Regressor</strong></td>
<td>( (1) )</td>
<td>( (2) )</td>
<td>( (3) )</td>
<td>( (4) )</td>
<td>( (5) )</td>
<td>( (6) )</td>
</tr>
<tr>
<td>Constant</td>
<td>8.7124</td>
<td>2.8901</td>
<td>.1261</td>
<td>.2517</td>
<td>10.8370</td>
<td>7.5828</td>
</tr>
<tr>
<td>( \hat{R}^2 )</td>
<td>( (16.51) )</td>
<td>( (1.37) )</td>
<td>( (3.90) )</td>
<td>( (2.11) )</td>
<td>( (5.52) )</td>
<td>( (2.08) )</td>
</tr>
<tr>
<td>Read</td>
<td>.0009</td>
<td>-.0019</td>
<td>.0001</td>
<td>.0003</td>
<td>.0027</td>
<td>.0057</td>
</tr>
<tr>
<td>( \hat{R} )</td>
<td>( (1.21) )</td>
<td>( (-1.17) )</td>
<td>( (1.11) )</td>
<td>( (3.20) )</td>
<td>( (2.80) )</td>
<td>( (3.28) )</td>
</tr>
<tr>
<td>NR read</td>
<td>.0791</td>
<td>.0566</td>
<td>-.0034</td>
<td>-.0046</td>
<td>.0033</td>
<td>-.0402</td>
</tr>
<tr>
<td>( \hat{R} )</td>
<td>( (1.24) )</td>
<td>( (.58) )</td>
<td>( (-.76) )</td>
<td>( -.89) )</td>
<td>( (.04) )</td>
<td>( (.42) )</td>
</tr>
<tr>
<td>Mech</td>
<td>-.0002</td>
<td>-.0005</td>
<td>-.0001</td>
<td>-.0001</td>
<td>-.0021</td>
<td>-.0017</td>
</tr>
<tr>
<td>( \hat{R} )</td>
<td>( (-.48) )</td>
<td>( (-.54) )</td>
<td>( (-2.16) )</td>
<td>( -1.13) )</td>
<td>( -1.35) )</td>
<td>( -1.73) )</td>
</tr>
<tr>
<td>NR mech</td>
<td>(...)</td>
<td>(.1969)</td>
<td>(...)</td>
<td>(.0002)</td>
<td>(...)</td>
<td>(.2196)</td>
</tr>
<tr>
<td>Math</td>
<td>.0015</td>
<td>-.0013</td>
<td>.0001</td>
<td>-.0000</td>
<td>.0030</td>
<td>-.0019</td>
</tr>
<tr>
<td>( \hat{R} )</td>
<td>( (2.02) )</td>
<td>( (.74) )</td>
<td>( (1.18) )</td>
<td>( -2.20 )</td>
<td>( (3.31) )</td>
<td>( -1.00) )</td>
</tr>
<tr>
<td>NR math</td>
<td>-.1087</td>
<td>.0562</td>
<td>.0015</td>
<td>.0006</td>
<td>-.0877</td>
<td>.0712</td>
</tr>
<tr>
<td>( \hat{R} )</td>
<td>( (-1.94) )</td>
<td>( (.83) )</td>
<td>( (.38) )</td>
<td>( .15) )</td>
<td>( -1.24) )</td>
<td>( (.96) )</td>
</tr>
<tr>
<td>Dext</td>
<td>.0008</td>
<td>-.0019</td>
<td>-.0000</td>
<td>.0003</td>
<td>.0002</td>
<td>.0036</td>
</tr>
<tr>
<td>( \hat{R} )</td>
<td>( (1.05) )</td>
<td>( (-.78) )</td>
<td>( (2.77) )</td>
<td>( .16) )</td>
<td>( (2.19) )</td>
<td>( (2.19) )</td>
</tr>
<tr>
<td>NR dext</td>
<td>(.0751)</td>
<td>(...)</td>
<td>(.0004)</td>
<td>(...)</td>
<td>(.1466)</td>
<td>(...)</td>
</tr>
<tr>
<td>Exp</td>
<td>(-.0523)</td>
<td>(.4260)</td>
<td>-.0028</td>
<td>-.0154</td>
<td>-.0129</td>
<td>.0776</td>
</tr>
<tr>
<td>( \hat{R} )</td>
<td>( (-1.49) )</td>
<td>( (3.10) )</td>
<td>( -1.11) )</td>
<td>( -1.93) )</td>
<td>( -2.29) )</td>
<td>( (.53) )</td>
</tr>
<tr>
<td>Exp*</td>
<td>(.0015)</td>
<td>-.0007</td>
<td>.0000</td>
<td>.0002</td>
<td>-.0000</td>
<td>-.0012</td>
</tr>
<tr>
<td>( \hat{R} )</td>
<td>( (2.22) )</td>
<td>( (-2.95) )</td>
<td>( .21) )</td>
<td>( 1.82) )</td>
<td>( -.01) )</td>
<td>( -1.49) )</td>
</tr>
<tr>
<td>Year 48</td>
<td>-.0020</td>
<td>-.0156</td>
<td>(...)</td>
<td>(...)</td>
<td>(...)</td>
<td>(...)</td>
</tr>
<tr>
<td>Year 69</td>
<td>(...)</td>
<td>(...)</td>
<td>(...)</td>
<td>(...)</td>
<td>(...)</td>
<td>(...)</td>
</tr>
<tr>
<td>S15–15</td>
<td>.1288</td>
<td>(...)</td>
<td>-.0062</td>
<td>(...)</td>
<td>.0168</td>
<td>(...)</td>
</tr>
<tr>
<td>( \hat{R} )</td>
<td>( (5.15) )</td>
<td>( (-3.49) )</td>
<td>(...)</td>
<td>(...)</td>
<td>(...)</td>
<td>(...)</td>
</tr>
<tr>
<td>S16</td>
<td>.0760</td>
<td>(...)</td>
<td>.0026</td>
<td>(...)</td>
<td>.1095</td>
<td>(...)</td>
</tr>
<tr>
<td>( \hat{R} )</td>
<td>( (3.82) )</td>
<td>( (1.79) )</td>
<td>(...)</td>
<td>(...)</td>
<td>(...)</td>
<td>(...)</td>
</tr>
<tr>
<td>S20</td>
<td>.1318</td>
<td>.0049</td>
<td>(...)</td>
<td>(...)</td>
<td>.2560</td>
<td>(...)</td>
</tr>
<tr>
<td>( \hat{R} )</td>
<td>( (4.10) )</td>
<td>( (2.13) )</td>
<td>(...)</td>
<td>(...)</td>
<td>(...)</td>
<td>(...)</td>
</tr>
<tr>
<td>( \lambda_a )</td>
<td>-.1069</td>
<td>(...)</td>
<td>(...)</td>
<td>(...)</td>
<td>(...)</td>
<td>(...)</td>
</tr>
<tr>
<td>( \hat{R} )</td>
<td>( (-3.21) )</td>
<td>( (2.45) )</td>
<td>(...)</td>
<td>(...)</td>
<td>(...)</td>
<td>(...)</td>
</tr>
<tr>
<td>( \lambda_b )</td>
<td>(...)</td>
<td>(...)</td>
<td>(...)</td>
<td>(...)</td>
<td>(...)</td>
<td>(...)</td>
</tr>
<tr>
<td>( \hat{R} )</td>
<td>( (-5.66) )</td>
<td>( (2.39) )</td>
<td>(...)</td>
<td>(...)</td>
<td>(...)</td>
<td>(...)</td>
</tr>
<tr>
<td>( R^2 )</td>
<td>.0750</td>
<td>.0439</td>
<td>.1578</td>
<td>.0513</td>
<td>.0740</td>
<td>.0358</td>
</tr>
</tbody>
</table>

*Note:* NR: No response; dummy variable; other variables are defined in Appendix A; t-values are shown in parentheses.
Structural Probit college selection equation

- Anticipated earnings gains are an important determinant of the decision to attend college

Evaluated at the observed proportion of college attendees and high school graduates

- Elasticity of the number enrolled in college with respect to initial earnings is 1.94
- Elasticity of the number enrolled in college with respect to an increment in father's education of 1.59 years (the difference in means between the two groups) is 0.0337
Conclusions

Introduced the program evaluation problem
  – Interested in the average effect of the treatment on the treated
  – But do not typically observe the counterfactual

Alternative methods of program evaluation address this problem in different ways
  – Randomization (experimental methods)
    – No difference in expectation between the treatment and control groups other than the treatment group received the program
  – Non-experimental methods
    – Non-random selection into treatment and control groups
      – Heckman selection model (Heckit)

Empirical application of the Heckman selection model
  – Estimating the return to education (Willis and Rosen 1979)