V. J. Hotz
May 28, 2007

# Notes on
# Willis & Rosen, "Education & Self-Selection"

***Model***

Two levels of Schooling: *A* (some College) and *B* (High School)

Observe earnings at 2 points in life cycle: soon after entrance into labor force & ~20 years later.

*Expected Earnings:*

If person *i* chooses *A* (some College), earnings are:

$$\begin{aligned} y_{ai}(t) &= 0, \quad 0 < t \leq S \\ y_{ai}(t) &= \overline{y}_{ai} \exp\left[g_{ai}(t-S)\right], \quad S \leq t < \infty \end{aligned} \tag{5}$$

where $S$ is incremental schooling period associated with *A* over *B* and $t - S$ is (potential) market work experience.

If person *i* chooses *B* (High School), earnings are:

$$y_{bi}(t) = \overline{y}_{bi} \exp\left[g_{bi}t\right], \quad 0 \leq t < \infty. \tag{6}$$

So, earnings prospects of individuals characterized by $(\overline{y}_a, g_a, \overline{y}_b, g_b)$, i.e., initial earnings and rates of growth in each of schooling alternatives.

*Present Value of Earnings under A* and *B*, respectively:

$$V_{ai} = \int_S^{\infty} y_{ai}(t)\exp(-r_i t)dt = \left[\frac{\overline{y}_{ai}}{(r_i - g_{ai})}\right]\exp(-r_i S) \tag{7}$$

$$V_{bi} = \int_0^{\infty} y_{bi}(t)\exp(-r_i t)dt = \frac{\overline{y}_{bi}}{(r_i - g_{bi})}, \tag{8}$$

where $r_i$ is person *i*'s discount rate, with $r_i > g_{ai}, g_{bi}$ and W&R ignore direct costs of school.

*Selection Rule:*

Choose $A$ if $V_{ai} > V_{bi}$ and choose $B$ if $V_{ai} \le V_{bi}$. Let $I_i = \ln(V_{ai}/V_{bi})$ or, substituting in for $V_{ai}$ and $V_{bi}$ from (5) – (8), we get

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

Use Taylor series approx. to nonlinear terms in above around population means, we get

$$I_i = \alpha_0 + \alpha_1(\ln \overline{y}_{ai} - \ln \overline{y}_{bi}) + \alpha_2 g_{ai} + \alpha_3 g_{bi} + \alpha_4 r_i \tag{9}$$

with

$$\begin{aligned} \alpha_1 &= 1 \\ \alpha_2 &= \partial I / \partial g_a = 1/\overline{r} - \overline{g}_a > 0 \\ \alpha_3 &= \partial I / \partial g_b = -1/\overline{r} - \overline{g}_b < 0 \\ \alpha_4 &= -\left[ S + \frac{(\overline{g}_a - \overline{g}_b)}{(\overline{r} - \overline{g}_a)(\overline{r} - \overline{g}_b)} \right] \end{aligned} \tag{10}$$

Then it follows that selection criteria are:

$$\begin{aligned} \Pr(\text{choose } A) &= \Pr(V_a > V_b) = \Pr(I > 0) \\ \Pr(\text{choose } B) &= \Pr(V_a \le V_b) = \Pr(I \le 0) \end{aligned} \tag{11}$$

*Earnings & Discount Functions:*

$$\begin{aligned} \ln \overline{y}_{ai} &= X_i \beta_a + u_{1i} \\ g_{ai} &= X_i \gamma_a + u_{2i} \end{aligned} \tag{12}$$

$$\begin{aligned} \ln \overline{y}_{bi} &= X_i \beta_b + u_{3i} \\ g_{bi} &= X_i \gamma_b + u_{4i} \end{aligned} \tag{13}$$

and

$$r_i = Z_i \delta + u_{5i} \tag{14}$$

Let vector $\boldsymbol{u}$ normally distributed with mean $\mathbf{0}$ and $\Sigma$ unrestricted.

*Reduced Form:*

$$
\begin{aligned}
I &= \alpha_0 + X\left[\alpha_1(\beta_a - \beta_b) + \alpha_2\gamma_a + \alpha_3\gamma_b\right] + \alpha_4 Z\delta + \alpha_1(u_1 - u_2) \\
&\quad + \alpha_2 u_2 + \alpha_3 u_3 + \alpha_5 u_5 \\
&\equiv W\pi - \varepsilon
\end{aligned}
\tag{15}
$$

where $W = [X,Z]$ and $-\varepsilon = \alpha_1(u_1 - u_2) + \alpha_2 u_2 + \alpha_3 u_3 + \alpha_5 u_5$. Then

$$
\Pr(A \text{ is observed}) = \Pr(W\pi > \varepsilon) = F\left(\frac{W\pi}{\sigma_\varepsilon}\right) \tag{16}
$$

*Observed Earnings & Selection Bias:*

$$
E\left(\ln \bar{y}_a \middle| I > 0\right) = X\beta_a + \frac{\sigma_{1\varepsilon}}{\sigma_\varepsilon}\lambda_a \tag{18}
$$

$$
E\left(g_a \middle| I > 0\right) = X\gamma_a + \frac{\sigma_{2\varepsilon}}{\sigma_\varepsilon}\lambda_a \tag{19}
$$

$$
E\left(\ln \bar{y}_b \middle| I \le 0\right) = X\beta_b + \frac{\sigma_{3\varepsilon}}{\sigma_\varepsilon}\lambda_b \tag{20}
$$

$$
E\left(g_b \middle| I \le 0\right) = X\gamma_b + \frac{\sigma_{4\varepsilon}}{\sigma_\varepsilon}\lambda_b \tag{21}
$$

where

$$
\lambda_a \equiv -f\left(W\pi/\sigma_\varepsilon\right) / F\left(W\pi/\sigma_\varepsilon\right) < 0 \tag{17}
$$

$$
\lambda_b \equiv f\left(W\pi/\sigma_\varepsilon\right) / \left[1 - F\left(W\pi/\sigma_\varepsilon\right)\right] > 0 \tag{22}
$$

$$
\sigma_{j\varepsilon} = \operatorname{cov}(u_j, \varepsilon), \quad j = 1, \ldots, 4.
$$

**Positive Selection Bias** if $\frac{\sigma_{j\varepsilon}}{\sigma_\varepsilon} < 0$, $j$ = 1, 2, since $\lambda_a < 0$ **and** $\frac{\sigma_{j\varepsilon}}{\sigma_\varepsilon} > 0$, $j$ = 3, 4, since $\lambda_b > 0$. Positive bias in both *A* and *B* implies **comparative advantage**.

*Estimation:*

Step 1: Estimate schooling choice (*A* or *B*) by probit to obtain $\widehat{\pi/\sigma_\varepsilon}$.

Step 2: Using $\widehat{\pi/\sigma_\varepsilon}$ to form $\hat{\lambda}_a$ and $\hat{\lambda}_b$ and then estimate

$$\begin{aligned} \ln \bar{y}_a &= X\beta_a + \beta_a^* \hat{\lambda}_a + \eta_1 \\ g_a &= X\gamma_a + \gamma_a^* \hat{\lambda}_a + \eta_2 \\ \ln \bar{y}_b &= X\beta_b + \beta_b^* \hat{\lambda}_b + \eta_3 \\ g_b &= X\gamma_b + \gamma_b^* \hat{\lambda}_b + \eta_4 \end{aligned} \tag{24}$$

with data on initial earnings and change in earnings to measure earnings growth rate.

Step 3: Can go back and form structural version of schooling choice probit, to see how well model based on maximizing earnings "fits" the observed schooling choices, i.e.,

$$\Pr(\text{choose } A) = \Pr\left\{ \frac{\alpha_0 + \alpha_1 \ln(\widehat{\bar{y}_a/\bar{y}_b} + \alpha_2 \hat{g}_a + \alpha_3 \hat{g}_b + \alpha_4 Z\hat{\delta})}{\sigma_\varepsilon} > \frac{\varepsilon}{\sigma_\varepsilon} \right\} \tag{26}$$

where estimated values formed from (24).

TABLE 1

DESCRIPTIVE STATISTICS

| Variable | High School (Group B) Mean | High School (Group B) SD | More than High School (Group A) Mean | More than High School (Group A) SD |
|---|---|---|---|---|
| Father's ED | 8.671 | 2.966 | 10.26 | 3.623 |
| Father's ED$^2$ | 83.99 | 55.53 | 118.4 | 78.09 |
| DK ED | .0999 | ... | .0464 | ... |
| Manager | .3628 | ... | .4954 | ... |
| Clerk | .1239 | ... | .1450 | ... |
| Foreman | .2238 | ... | .1695 | ... |
| Unskilled | .1492 | ... | .0819 | ... |
| Farmer | .1062 | ... | .0720 | ... |
| DK job | .0177 | ... | .0124 | ... |
| Catholic | .2933 | ... | .2138 | ... |
| Jew | .0405 | ... | .0617 | ... |
| Old sibs | 1.143 | 1.634 | .9035 | 1.383 |
| Young sibs | .9381 | 1.486 | .8138 | 1.266 |
| Mother works: | | | | |
| Full 5 | .0468 | ... | .0486 | ... |
| Part 5 | .0392 | ... | .0504 | ... |
| None 5 | .7168 | ... | .7507 | ... |
| Full 14 | .0822 | ... | .0936 | ... |
| Part 14 | .0708 | ... | .0851 | ... |
| None 14 | .6384 | ... | .6713 | ... |
| H.S. shop | .2592 | ... | .0908 | ... |
| Read | 20.57 | 10.17 | 24.06 | 11.63 |
| NR read | .0291 | ... | .0128 | ... |
| Mech | 59.24 | 18.27 | 58.88 | 18.96 |
| NR mech | .0025 | ... | 0 | ... |
| Math | 18.13 | 11.82 | 28.94 | 17.17 |
| NR math | .0683 | ... | .0188 | ... |
| Dext | 50.04 | 9.359 | 50.68 | 9.811 |
| NR dext | 0 | ... | .0071 | ... |
| Exp | 29.33 | 2.439 | 24.54 | 2.907 |
| Exp$^2$ | 866.1 | 147.1 | 610.4 | 147.4 |
| S13–15 | ... | ... | .3106 | ... |
| S16 | ... | ... | .3993 | ... |
| S20 | ... | ... | .0823 | ... |
| Year 48 | 46.62 | 1.584 | 48.05 | 1.869 |
| Year 69 | 69.11 | .3691 | 69.08 | .3437 |
| ln $\bar{y}$ | 8.635 | .4107 | 8.526 | .3871 |
| ln $y(\bar{t})$ | 9.326 | .4573 | 9.639 | .4904 |
| $g$ | .0309 | .0251 | .0535 | .0283 |
| $\lambda_a$ | −1.2870 | .2873 | −.3193 | .2256 |
| $\lambda_b$ | .4666 | .3763 | 1.605 | .5212 |
| No. observations | 791 | | 2820 | |

NOTE.—Variables are defined in Appendix A.

TABLE 2

College Selection Rules: Probit Analysis

| Variable | Reduced Form (16) Coefficient | Reduced Form (16) $t$ | Structure (26) Coefficient | Structure (26) $t$ | Structure (29) Coefficient | Structure (29) $t$ |
|---|---|---|---|---|---|---|
| Constant | .0485 | .20 | .1512 | .22 | .1030 | .17 |
| Background: | | | | | | |
| Father's ED | −.0145 | −.41 | −.0168 | −.54 | −.0152 | −.49 |
| Father's ED$^2$ | .0037 | 2.05 | .0038 | 2.26 | .0037 | 2.26 |
| DK ED | −.4059 | −3.96 | −.3924 | −2.79 | −.4001 | −2.91 |
| Manager | .1897 | 2.17 | .1825 | 2.13 | .1871 | 2.21 |
| Clerk | .0556 | .54 | .0561 | .59 | .0554 | .59 |
| Foreman | .0182 | .19 | .0210 | .23 | .0200 | .22 |
| Unskilled | −.0910 | −.85 | −.0948 | −.89 | −.0928 | −.87 |
| Farmer | −.2039 | −2.12 | −.2256 | −2.27 | −.2094 | −2.14 |
| DK job | −.0413 | −.19 | −.0629 | −.29 | −.0609 | −.28 |
| Catholic | −.1144 | −1.91 | −.0982 | −1.51 | −.1083 | −1.66 |
| Jew | −.0293 | −.23 | .0143 | .12 | −.0158 | −.14 |
| Old sibs | −.0162 | −.93 | −.0162 | −.93 | −.0161 | −.93 |
| Young sibs | .0122 | .63 | .0096 | .49 | .0112 | .57 |
| Mother works: | | | | | | |
| Full 5 | .1039 | .66 | .1168 | .81 | .1104 | .76 |
| Part 5 | .2179 | 1.42 | .2106 | 1.52 | .2156 | 1.56 |
| None 5 | .0655 | .63 | .0677 | .65 | .0661 | .64 |
| Full 14 | .2898 | 2.29 | .2884 | 2.30 | .2888 | 2.33 |
| Part 14 | .2709 | 2.20 | .2768 | 2.02 | .2693 | 2.03 |
| None 14 | .1980 | 1.91 | .1990 | 1.92 | .1966 | 1.92 |
| H.S. shop | −.4411 | −6.14 | −.4397 | −3.74 | −.4379 | −3.90 |
| Ability: | | | | | | |
| Read | .0047 | 1.67 | ... | ... | ... | ... |
| NR read | −.2575 | −1.41 | ... | ... | ... | ... |
| Mech | −.0070 | −4.29 | ... | ... | ... | ... |
| NR mech | −3.0236 | −1.04 | ... | ... | ... | ... |
| Math | .0244 | 12.34 | ... | ... | ... | ... |
| NR math | −.7539 | −5.75 | ... | ... | ... | ... |
| Dext | .0019 | .72 | ... | ... | ... | ... |
| NR dext | 2.2797 | .47 | ... | ... | ... | ... |
| Earnings: | | | | | | |
| ln $(\bar{y}_a/\bar{y}_b)$ | ... | ... | 5.1486 | 2.25 | ... | ... |
| $g_a$ | ... | ... | 138.3850 | 1.83 | 7.6632 | .11 |
| $g_b$ | ... | ... | −44.2697 | −1.28 | 71.8981 | 2.34 |
| ln $y_a(t)/y_b(t)$ | ... | ... | ... | ... | 5.1501 | 2.57 |
| Observations | | 3611 | | 3611 | | 3611 |
| Limit observations | | 791 | | 791 | | 791 |
| Nonlimit observations | | 2820 | | 2820 | | 2820 |
| −2 ln (likelihood ratio) | | 579.5 | | 568.8 | | 576.6 |
| $\chi^2$ degree freedom | | 28 | | 23 | | 23 |

Note.—$t$ is asymptotic $t$-statistic: DK: Don't know, dummy variable; NR: No response, dummy variable; other variables are defined in Appendix A.

TABLE 3

STRUCTURAL EARNINGS ESTIMATES: EQUATIONS (24) AND (28), OLS

| | DEPENDENT VARIABLE | | | | | |
|---|---|---|---|---|---|---|
| REGRESSOR | $\ln \bar{y}_a$ (1) | $\ln \bar{y}_b$ (2) | $g_a$ (3) | $g_b$ (4) | $\ln y_a(\bar{t})$ (5) | $\ln y_b(\bar{t})$ (6) |
| Constant | 8.7124 | 2.8901 | .1261 | .2517 | 10.3370 | 7.5328 |
| | (16.51) | (1.37) | (3.90) | (2.11) | (5.52) | (2.08) |
| Read | .0009 | −.0019 | .0001 | .0003 | .0027 | .0057 |
| | (1.21) | (−1.17) | (1.11) | (3.20) | (2.80) | (3.28) |
| NR read | .0791 | .0506 | −.0034 | −.0046 | .0033 | −.0402 |
| | (1.24) | (.58) | (−.76) | (−.89) | (.04) | (−.42) |
| Mech | −.0002 | −.0005 | −.0001 | −.0001 | −.0021 | −.0017 |
| | (−.48) | (−.54) | (−2.16) | (−1.13) | (−3.59) | (−1.73) |
| NR mech | ... | .1969 | ... | .0002 | ... | .2196 |
| | | (.69) | | (.01) | | (.68) |
| Math | .0015 | −.0013 | .0001 | −.0000 | .0030 | −.0019 |
| | (2.02) | (.74) | (1.18) | (−.20) | (3.31) | (−1.00) |
| NR math | −.1087 | .0562 | .0015 | .0006 | −.0877 | .0712 |
| | (−1.94) | (.83) | (.38) | (.15) | (−1.24) | (.96) |
| Dext | .0008 | −.0019 | −.0000 | .0003 | .0002 | .0036 |
| | (1.03) | (−1.21) | (−.78) | (2.77) | (.16) | (2.19) |
| NR dext | .0751 | ... | −.0004 | ... | .1466 | ... |
| | (.28) | | (−.02) | | (.43) | |
| Exp | −.0523 | .4260 | −.0028 | −.0154 | −.0129 | .0776 |
| | (−1.49) | (3.10) | (−1.11) | (−1.93) | (−.29) | (.53) |
| $Exp^2$ | .0015 | −.0067 | .0000 | .0002 | −.0000 | −.0012 |
| | (2.22) | (−2.95) | (.21) | (1.82) | (−.01) | (−.49) |
| Year 48 | −.0020 | −.0156 | ... | ... | ... | ... |
| | (−.48) | (−1.72) | | | | |
| Year 69 | ... | ... | ... | ... | −.0067 | .0039 |
| | | | | | (−.26) | (.09) |
| S13–15 | .1288 | ... | −.0062 | ... | .0168 | ... |
| | (5.15) | | (−3.49) | | (.52) | |
| S16 | .0760 | ... | .0026 | ... | .1095 | ... |
| | (3.82) | | (1.79) | | (4.26) | |
| S20 | .1318 | ... | .0049 | ... | .2560 | ... |
| | (4.10) | | (2.13) | | (6.15) | |
| $\lambda_a$ | −.1069 | ... | .0058 | ... | .0206 | ... |
| | (−3.21) | | (2.45) | | (.49) | |
| $\lambda_b$ | ... | −.0558 | ... | .0118 | ... | .2267 |
| | | (−.66) | | (2.39) | | (2.48) |
| $R^2$ | .0750 | .0439 | .1578 | .0513 | .0740 | .0358 |

NOTE.—NR: No response, dummy variable; other variables are defined in Appendix A; $t$-values are shown in parentheses.