Empirical Modeling: A Classy Example

- Mincer’s model of schooling, experience and earnings
- Develops empirical specification from theory of human capital accumulation
- Goal: Understanding the cross-section distribution of income
- Interaction of data and assumptions
Assumptions

• S years schooling => earnings E(S) (with no other investments)

• PDV of lifetime earnings is equated across identical individuals

• Years at work, T, is independent of S

• Comments?
Implications

\[ V(s) = \int_{s}^{R} E(S, t) e^{-rt} \, dt \]

- \( R = \) retirement time, \( E(S, t) \) is earnings at \( t \)
- \( E(S, t) = E(S, t') = E(S) \)
- \( V(S) = E(S) \left( e^{-rS} - e^{-rR} \right) / r \)
- \( V(S) = V(S') = V \)
- \( R = S + T \)
More implications

\[ rV = E(S)(e^{-rS} - e^{-rS} e^{-rT}) \]
\[ = E(0)(1 - e^{-rT}) \]
\[ \Rightarrow E(S) = E(0)e^{rS} \]

or

\[ \ln E(S) = \ln E(0) + rS \]

Nice clue to skewed income distribution: A symmetric dist. of S can imply skewed dist. of income (h.c. investment => skew)
Post-School Investments

• Impt. since earnings are not constant after schooling is over
• Distinguish actual earnings $Y$; potential $E$
• Assume: workers devote fraction $k$ of time to investment, $1-k$ to market work
• $Y = (1-k)E$
• Suppose the return on investment is $p$, so investment of $kE$ today yields $pkE$ in all future periods
Implications

- Potential Earnings growth

\[ \frac{\partial E(S,t)}{\partial t} = pk(t)E(S,t) \]

\[ \Rightarrow \ln E(S,t) = \ln E(0) + rS + p \int_{0}^{t} k(u)du \]

To proceed, we need the investment function \( k(t) \). Assume

\( k(t) = k(1-t/t^*) \) for \( t < t^* \), 0 for \( t > t^* \).
A little more work

\[ \ln E(S,t) = \ln E(0) + rS = pkt - (pk / 2t^*)t^2 \]

- The log of earnings is linear in schooling and quadratic in experience
- Glitch: This equation is for E, not Y; these differ if \( t < t^* \)
- \( \ln Y(S,t) = \ln E(S,t) + \ln(1-k(t)) \)
- \[ = \ln E(S,t) + \ln(1-k+kt/t^*) \]
- Approximate the second term by a quadratic (good?)
Finally, The Empirical Specification

\[
\ln Y(S, t) = \beta_1 + \beta_2 S + \beta_3 t + \beta_4 t^2
\]

\(\beta_2\) is the rate of return to schooling

• \(t\) is experience - of ten not measured. Assume continuous post-schooling employment. Then \(t = A - S\), where \(A\) is age.

• The interpretation of the coefficients depends on the model!!
Interpretation

• Mincer’s model:
• \( \ln Y(S,t) = \beta_1 + \beta_2 S + \beta_3 (A-S) + \beta_4 (A-S)^2 \)

• Suppose instead you fit the model
• \( \ln Y(S,t) = \beta_1' + \beta_2' S + \beta_3' A + \beta_4' A^2 \)
• \( \beta_2' = \beta_2 \) ?
• Difference is whether age or experience is held constant.
Fit to 1960 Census Data

- Annual earnings, white nonfarm nonstudent men, 31K obs
- $\ln Y = 6.2 + 0.1075S + 0.081t - 0.0012t^2$
- $R^2 = 0.285$, enormous t-statistics (basically a good fit)
- Rate of return approximately 11%
- This simple economic model explains 28.5% of the variance in cross-sectional earnings
Strategy

• Focussed on goal - relationship between schooling and earnings in the cross-section.

• Practical matters always at the forefront.

• Model is pushed as far as possible.

• Ignored: many “side issues,” unions, imperfect markets, regulations, other sources of individual heterogeneity, etc.

• Still explains 28.5% of variance in earnings!