From Economic Model to Econometric Model (The Mincer Model)

Helle Bunzel
The Basic Model

★ Mincer (1974)
★ Theory of investment in human capital used to examine income distribution.
★ Relationship between schooling, earnings and post school investments in human capital.
★ Assumptions of the economic model:
  1. An individual with S years of schooling has earnings which do not depend on age.
  2. PV of lifetime incomes are the same across individual regardless of schooling if no post school investments are made.
  3. The number of years spent at work are independent of the number of years of schooling.
Definition: $E(S, t)$ is the earnings at time $t$ of a person with $S$ years of schooling.

PV of earnings of an individual who enters the labor market after $S$ years of schooling:

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

Under Assumption 1, $E(S, t)$ does not depend on $t$.

$$V(S) = \int_{S}^{R} E(S) e^{-rt} dt = E(S) \left[ e^{-rS} - e^{-rR} \right] / r$$
The Earning-Schooling Connection

★ Assumption 2 states that $V(S)$ should not depend on $S$, so $V(S) = V$.

★ Assumption 3 states that for some $T$, $R = S + T$, such that everyone works the same number of years. $T$

★ Using these two pieces of information:

$$V(S) = V = E(S) \left[ e^{-rS} - e^{-rS}e^{-rT} \right] / r \iff \quad rV = E(S) \left[ e^{-rS} - e^{-rS}e^{-rT} \right]$$

$$rV = E(0) \left[ 1 - e^{-rT} \right] = E(S) \left[ e^{-rS} - e^{-rS}e^{-rT} \right] \iff \quad E(0)e^{-r0} = E(S)e^{-rS} \iff \quad \ln E(S) = \ln E(0) + rS$$

★ The assumptions that we have made lead to a log-linear relationship between earnings and schooling.
Recall that we wanted to look at human capital investments and the effect on income distribution.
Log-linear Relationship

Earnings

Schooling

0  5  10  15  20  25  30  35  40
The Income Distribution

★ Even if the years of schooling is distributed symmetrically/evenly/uniformly the distribution of income is going to be very skewed.

★ To elaborate:

<table>
<thead>
<tr>
<th>Income</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 - $10,000</td>
<td>15</td>
</tr>
<tr>
<td>$10 - $20,000</td>
<td>8</td>
</tr>
<tr>
<td>$20 - $30,000</td>
<td>4</td>
</tr>
<tr>
<td>$30 - $40,000</td>
<td>3</td>
</tr>
<tr>
<td>$40 - $50,000</td>
<td>2</td>
</tr>
<tr>
<td>$50 - $60,000</td>
<td>2</td>
</tr>
</tbody>
</table>

★ VERY heavy at the bottom.
★ This was the connection between schooling and earnings. We were interested in the relationship between experience, schooling and earnings.
Post School Investment

★ We wish to add in post school investment. Make another assumption:

4. The return to postschool investment is a constant $p$.

★ We’ll assume that a worker devotes a fraction $k$ of his time to investment in human capital, and a fraction $(1 - k)$ to actual work.

★ This all implies that growth in earnings is determined by:

$$\frac{\partial E(S,t)}{\partial t} = pk(t)E(S,t)$$
Post School Investment
solving the differential equation

Solving this differential equation gives us:

\[
\frac{1}{E(S,t)} \frac{\partial E(S,t)}{\partial t} = pk(t) \iff \\
\frac{\partial \ln E(S,t)}{\partial t} = pk(t) \iff \\
\int_0^t \frac{\partial \ln E(S,t)}{\partial t} = \int_0^t pk(t) \iff \\
\ln E(S,t) = p \int_0^t k(t)
\]

For the specific solution, we know that

\[
\ln E(S,t) = C + p \int_0^t k(u)du
\]

We need one point to find the value of C.
Post School Investment

solving the differential equation

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

If we insert \( k(u) = 0 \) into the differential equation, we get:

\[ \ln E(S, t) = C, \]

implying that

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

This gives the solution:

\[ \ln E(S, t) = \ln E(0) + rS + p \int_0^t k(u) \, du \]

⋆ Why does time start at 0 in this earnings equation?
We need to assume something about the frequency of investment in human capital.

Mincer assumed that:

\[ k(t) = k^* \left( 1 - \frac{t}{t^*} \right) = k^* - \frac{k^*}{t^*} t, \]

where \( t^* \) is the last period in which we invest in human capital.

Note that this definition is consistent with \( t \) running from the time we finish school.
Post School Investment
How frequently?

★ Then,

\[
\ln E(S, t) = \ln E(0) + rS + p \int_0^t k(u) \, du
\]

\[
= \ln E(0) + rS + pk^* \left[ u - \frac{u^2}{2t^*} \right]_0^t
\]

\[
= \ln E(0) + rS + pk^* t - pk^* \frac{t^2}{2t^*}
\]

★ This gives a relationship between potential earnings and schooling.

★ Now note that since we only work part of the time:

\[
Y(S, t) = (1 - k(t)) E(S, t) \Leftrightarrow \ln Y(S, t) = \ln E(S, t) + \ln (1 - k(t))
\]

★ The economic theory is done!
The Econometric Model

Now we need to make the model linear so that we can do econometrics.

\[ \ln Y(S,t) = \ln E(0) + rS + pk^*t - pk^* \frac{t^2}{2t^*} + \ln \left( 1 - k^* + k^* \frac{t}{t^*} \right) \]

What are the variables? Which terms are parameters?

We still have this investment term in the equation and it is not quite linear.

\[ \ln (1 - k(t)) = \ln \left( 1 - k^* + k^* \frac{t}{t^*} \right) \]
The Econometric Model

★ Taylor expansion around \( t = t^* \):

\[
\ln \left( 1 - k^* + k^* \frac{t}{t^*} \right) = \ln(1) + \frac{k^*/t^*}{1} (t - t^*) - \frac{1}{2} \left( \frac{k^*}{t^*} \right)^2 (t - t^*)^2
\]

\[
= -k^* - \frac{1}{2} (k^*)^2 + \left[ \frac{k^*}{t^*} + \frac{(k^*)^2}{t^*} \right] t - \frac{1}{2} \left( \frac{k^*}{t^*} \right)^2 t^2
\]

★ Note: The error of the expansion is highest when \( t = 0 \).
The Econometric Model

* The final equation is

\[
\ln Y(S, t) = \ln E(S, t) + \ln (1 - k(t)) \\
= \ln E(0) + rS + pk^* t - pk^* \frac{t^2}{2t^*} - k^* - \frac{1}{2} (k^*)^2 \\
+ \left[ \frac{k^*}{t^*} + \frac{(k^*)^2}{t^*} \right] t - \frac{1}{2} \left( \frac{k^*}{t^*} \right)^2 t^2 \\
= \left\{ \ln E(0) - k^* - \frac{1}{2} (k^*)^2 \right\} + rS + \left\{ \frac{k^*}{t^*} + \frac{(k^*)^2}{t^*} + pk^* \right\} t \\
+ \left\{ -\frac{1}{2} \left( \frac{k^*}{t^*} \right)^2 - \frac{pk^*}{2t^*} \right\} t^2
\]
This can be written as a regression model:

\[ \ln Y = \beta_0 + \beta_1 S + \beta_2 t + \beta_3 t^2 + \epsilon \]

What do we get from this regression?

Can we sign any of the coefficients?
Conclusion

★ Strategy:

◊ Use the economic model.
◊ Make assumptions that simplify and eventually lead us to something we can put to the data.

★ In spite of many omitted issues and simplifying assumption, this model does explain a lot of the variation in earnings.

★ What if we had just done the linear regression to start with.

★ What information does just running the regression give you?

◊ Simple correlation
◊ Were you to create a model, which features need to be included?