LECTURE 1: HOW TO SPECIFY AN EMPIRICAL MODEL

1.1 CLASSY APPLICATION 1: CROSS SECTION DATA ON EARNINGS, SCHOOLING AND EXPERIENCE

Mincer (1974) provided an extremely ingenious development of an equation relating earnings to schooling and to experience. This work used the (then) recent theory of investment in human capital to address questions of income distribution. Mincer provided a set of assumptions leading to a particular functional form for the relationship. Most of the assumptions can be checked against data themselves. Some are plainly false, but may nevertheless be useful as empirical approximations. Mincer's approach is to push a fairly simple, and highly specific, model as far as possible, as long as it seems consistent with the data. The result was a simple organizing framework for understanding a large amount of data. The "human capital earnings function" proved to be surprisingly powerful as an empirical tool.

Mincer considers the relationship between earnings and schooling and "postschool investments." These postschool investments are mainly to be interpreted as on-the-job training, the accumulation of which is costly in that workers sacrifice income in order to obtain training. These transactions need not occur as formal arrangements: new workers are often paid less than experienced workers while they learn their way around the machinery, the site, and their coworkers. Mincer assumes (1) An individual with S years of schooling has earnings which do not vary with age (in the absence of postschool investments), (2) The present discounted value of lifetime earnings is equated across individuals, regardless of schooling, in the absence of postschool investments, and (3) The number of years spent at work, T, is independent of the
number of years spent in school.

These assumptions alone are sufficient to obtain one quite interesting empirical implication. The present value of earnings for an individual who will enter the labor market after $S$ years of schooling is

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

where a continuous time formulation has been adopted for convenience (it is a useful exercise to run through these calculations in discrete time if this argument seems puzzling).

Here, $E(t)$ is earnings, a flow in continuous time, at time $t$, $R$ is the retirement time and $r$ is the discount rate. Under assumption (1), $E(S,t) = E(S,t') = \text{def } E(S)$, which can be brought out of the integral, so (note $\int_{S}^{R} e^{-rt} dt = r^{-1} (e^{-rS} - e^{-rR})$)

$$V(S) = E(S) * (e^{-rS} - e^{-rR}) / r$$

With assumption (2), $V(S) = V(S') = V$, a constant. With (3), the time of retirement is $R = S + T$.

Therefore,

$$E(S)(e^{-rS} - e^{-rS-rT}) = rV = E(0) (1 - e^{-rT})$$

which implies on a cancellation

$$E(S) = E(0)e^{rS}$$

In natural logarithms

$$\ln E(S) = \ln E(0) + rS$$

The assumptions made lead to a loglinear relationship between earnings and schooling.

This result gives a very interesting clue about the determination of the distribution of income. The distribution of income is highly skewed, with a long right tail almost like a pareto distribution. If
we assume that productive ability is normally distributed, or more generally symmetrically distributed, then it is intuitive to suspect that the corresponding distribution of income would be symmetric. However, the loglinear relationship between earnings and schooling induces a skewed distribution of earnings if schooling is symmetrically distributed. Human capital investment can therefore lead to skewness in the distribution of income.

However, this relationship between earnings and schooling is not Mincer's main concern. In fact, earnings are not constant over time for individuals with a fixed amount of schooling. In part, this is due to post-school investments. At this point it is useful to distinguish "potential" earnings $E$ from actual earnings $Y$. Suppose a worker devotes a fraction $k$ of his time to investment in human capital and $(1-k)$ to market work. His observed earnings will be reduced by the investment, $Y = (1-k)E$. Assume (4) that the return on postschool investments is a constant, $p$. The worker who invests $kE$ enjoys a return in terms of potential earnings of $pkE$ in all future periods (the realized return may be smaller because of time spent investing in the future). Thus, earnings growth is governed by

$$\frac{\partial E(S,t)}{\partial t} = pk(t)E(S,t).$$

Solving this simple differential equations yields

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

an equation giving the log of earnings as a function of schooling $S$ and experience $t$. Mincer now assumes a simple function form for the investment function $k(t)$

$$k(t) = k^*(1-t/t^*)$$

Here $t^*$ is the end of the investment period (i.e. after $t^*$, $k=0$). Investment, in terms of the fraction of potential earnings devoted to investment, begins at the fraction $k$ then declines linearly to zero at time $t^*$. This functional form is not derived formally from an analysis of the optimal
investment program for an individual (a better approach, followed by later authors) nevertheless it seems a reasonable starting point as an empirical approximation. With this assumption we have
\[ \ln E(S,t) = \ln E(0) + rS + pkt - \frac{(pk/2)}{t^*}t^2, \]
an equation giving the logarithm of earnings for an individual with S years of schooling and t years of experience as a linear function of schooling and a quadratic function of experience. A "glitch" is that the dependent variable here is the log of potential earnings, not actual earnings, so the equation is suitable for estimation and the coefficients for interpretation within the context of the model only for a sample of individuals who are no longer investing in human capital, i.e., who have experience t greater than t*.

Actual earnings \( Y(S,t) \) are related to potential earnings \( E(S,t) \) by the formula
\[ Y(S,t) = (1-k(t))E(S,t), \]
hence
\[ \ln Y(S,t) = \ln E(S,t) + \ln(1-k(t)) = \ln E(S,t) + \ln(1-k + (k/t^*)t). \]

Mincer considers approximating the second term here by a quadratic in t. The maximum error in this approximation will occur for individuals immediately out of school who are investing the maximum fraction of their potential earnings, k. Suppose k = .25, so new workers spend twenty-five percent of their potential earnings (time) in investment. Then the maximum relative error in the quadratic approximation is less than 2.3%. Assuming that this is satisfactory, we have an estimating equation for explaining the relationship between observed earnings and investments as a function of schooling and experience in the form
\[ \ln Y(S,t) = \beta_1 + \beta_2 S + \beta_3 t + \beta_4 t^2. \]

Mincer and many others have found that this functional form gives a surprisingly good explanation of cross-sectional variation in earnings. Note that \( \beta^2 \) can be interpreted as a rate of return to schooling.
Using data from the 1960 Census, Mincer fits this equation to a sample of 31,093 observations on annual earnings of white, nonfarm, nonstudent men. He finds

\[ \ln Y = 6.2 + 0.1075S + 0.081t - 0.0012t^2 \]

with enormous t-statistics and \( R^2 = 0.285 \). Thus the simple economic model explains 28.5% of the cross-section variation in earnings. The estimated rate of return to schooling is nearly 11%.

This application illustrates an extraordinarily productive research strategy. A simple economic model, designed to capture a key relationship, is pushed as far as possible. Practical issues were kept in mind throughout the development of the model - most importantly, the data at hand and the knowledge built up from previous work. The many "side issues" omitted -- unions, regulation, imperfect capital and labor markets, etc. notwithstanding, the simple model explains nearly 30% of the variance in earnings.

1.2 CLASSY APPLICATION 2: THE TERM STRUCTURE OF INTEREST RATES

The "term structure" of interest rates is the pattern of interest rates for bonds of different maturities. A six-month bill, for example, could be expected to carry a different rate of interest than a 20-year bond. The shorter term obligations will perhaps have much more volatile rates, while the long term obligations may not have rates which vary much with the point in the cycle at the issuing date. Nevertheless, short term rates and long term rates must be related, due to the effects of investors operating in both short and long term markets. In fact, it is useful to decompose long term rates into an average of current and expected future short term rates.

It is convenient in developing a theoretical model for the term structure to consider bonds which are traded in period \( t \) and which pay a lump sum in period \( t+n \). The return on such a bond is denoted \( R(t,n) \), so that one dollar invested at period \( t \) returns \((1+R(t,n))^n\) dollars in period \( t+n \). Ordinarily, bonds pay "coupons" in each period as well as a "face value" at maturity, a complication we will return to, but will ignore for the present. The pattern of the \( R(t,n) \) as \( n \), the
term of the bond, varies for \( t \), the period of issue, fixed, is known as the "yield curve" at period \( t \).

Now, we will define forward rates for one-period loans. These are rates agreed to in period \( t \) for one-period bonds to be purchased in period \( t+j-1 \) and redeemed in period \( t+j (j=1,...) \). Let \( F(t,j) \) be the one-period return on a one-period bond redeemed in period \( t+j \) when the transaction is made in period \( t \).

Futures markets for bonds are not common as futures in, for example, commodities, so it is worth stressing that the term structure given by the \( R(t,n) \) implies an implicit futures market in bonds with a corresponding set of futures rates \( F(t,j) \). Consider selling a two-period (no-coupon) bond and buying a three-period similar bond. This transaction is exactly the same as buying a forward commitment for a one-period bond purchased in period two at a rate determined now. Suppose both bonds have face value $1. In two periods you will pay out \((1+R(2,t))^2\) dollars and in the following period you will receive \((1+R(3,t))^3\) dollars. You had no money invested until period \( t+2 \) (recall that you both bought and sold $1 bonds in the current period) at which point you effectively invested in a one-period bond with rate of return

\[
F(3,t) = (1 + R(e,t))^3(1 + R(2,t))^2 - 1.
\]

The corresponding general formula is obtained by substituting \( j \) and \( j-1 \) for 3 and 2 respectively.

Implicit rates for forward contracts for multiperiod bonds can be constructed similarly. The first few forward rates \( F(t,j) \) and "long-term" rates \( R(t,n) \) are related by

\[
1 + R(1,t) = 1 + F(1,t)
\]
\[
(1 + R(2,t))^2 = (1 + R(1,t))(1 + F(2,t))
\]
\[
(1 + R(3,t))^3 = (1 + R(1,t))(1 + F(2,t)) (1 + F(3,t))
\]

These expressions can be simplified usefully by assuming simple rather than compound interest in each period is paid only on the principal and not previous interest. As an approximation to
compound interest, this assumption may be tolerable for small interest rates and few periods. With the assumption of simple interest we have
\[
\begin{align*}
R(1,t) &= F(1,t) \\
R(2,t) &= (R(1,t) + F(2,t))/2 \\
R(3,t) &= (R(1,t) + F(2,t) + F(3,t))/3 \\
\end{align*}
\]

and so on. In general, the long rate is the average of the current one-period rates. Alternatively, the forward rates can be expressed in terms of the long-term rates

\[
F(t,j) = jR(t+j,t) - (j-1)R(t+j-1,t)
\]

Meiselman (1962) argued that speculators indifferent to risk would enter securities markets, forcing forward rates to expected future spot rates. The expected future spot rates under the hypothesis of rational expectations are exactly the mathematical expectations of the spot rates, given current information. Thus

\[
F(j,t) = E(R(1,t+j-1)|R(1,t),R(1,t-1),...)
\]

where \( E \) is the mathematical expectation operator and current information consists only of current and lagged one-period rates. Under the assumption that conditional expectations are linear we can write

\[
E_{t+1}(R(1,t+j)) = E_t(R(1,t+j)) + \beta(R(1,t+1) - E_t(R(1,t+1)))
\]

where the subscript \( t \) on the expectation operator indicates a conditional expectation, with conditioning information dated \( t \) and earlier.

\[
E_t(R(1,t+1)) = E(R(1,t+1)|R(1,t),R(1,t-1),...)
\]

\[= F(1,t) \]

Upon substituting, we find

\[
F(j,t) - F(j,t-1) = \beta(R(1,t) - F(2,t-1))
\]

Meiselman (1962) estimated equations of the form

\[
F(j,t) - F(j+1,t-1) = \alpha_j + \beta_j (R(1,t) - F(2,t-1))
\]

for \( j = 1 \) to 8 years using data on annual yields on high grade corporate bonds over the period 1901-1954. There are difficulties in constructing suitable data, since bonds typically pay coupons.
This problem is avoided in practice by defining \( R(t,n) \) as the rate which discounts the payment stream from coupons and the final payment to the current (period \( t \)) price of the bond. This rate is the yield to maturity of the bond. For given \( t \), a plot was made of the yield to maturity versus the term of maturity. Observations are securities. A smooth curve fit through these data is a yield curve. The term structure data \( R(t,n) \) are then read off the curve. Thus, some estimation has already been done in setting up the data for the analysis to follow. The model suggests that the intercepts \( \alpha_j \) should be zero, since if there were a predictable trend in forward rates as a function of term, speculators could make expected profits. In fact, the constant terms are small and insignificantly different from zero. This equation is known as the "error learning model." Sargent (1979) gives a development of the model of information different from \( R(1,t-k)(k=0,1,...) \) but useful in predicting future spot rates. Sargent also works in the framework of projections rather than expectations, relaxing slightly the assumption of linear expectations.

Similar, but refined, models have since been used in a variety of studies of the efficient markets hypothesis using data from the stock market, the foreign exchange market, and other sources.