Recitation Notes 7

Konrad Menzel

October 27, 2006

1 Educational Choice

1.1 The Discount Rate Bias Story

This part is based on Question I from the 2003 Labor Generals. We are given a human capital earnings function

$$y_i = \alpha_i + b_i S_i - \frac{1}{2} k_1 S_i^2$$

and individuals face a convex cost of schooling

$$c_i = \kappa_i + r_i S_i + \frac{1}{2} k_2 S_i^2$$

where r_i is often interpreted as the individual's discount rate, or the opportunity cost of capital (for empirical purposes, earnings y_i is actually understood to be in logs, but for the theoretical derivations, it should be just earnings). In order to determine the optimal level of schooling, the individual equalizes the marginal return to schooling to the marginal cost,

$$b_i - k_1 S_i = r_i + k_2 S_i \Longleftrightarrow S_i^* = \frac{b_i - r_i}{k_1 + k_2}$$

If $k_2 = 0$, the marginal return to schooling at the level of education chosen by the individual is

$$\frac{\partial y_i}{\partial S_i} = b_i - k_1 S_i^* = r_i$$

which is also the equilibrium condition for the classic model we saw in class. But if the marginal cost of education is strictly convex ($k_2 > 0$), we get instead

$$\frac{\partial y_i}{\partial S_i} = b_i - k_1 \frac{b_i - r_i}{k_1 + k_2} = \frac{k_2}{k_1 + k_2} b_i + \frac{k_1}{k_1 + k_2} r_i$$

which is a convex combination of the discount rate and b_i.

Now we want to estimate the return to schooling using a binary instrumental variable Z to address omitted variable bias and potential measurement error issues. More specifically, we assume that the instrumental variable shifts the capital cost (or discount rate for each individual, i.e. the individual faces $r_i = r_{0i}$ if the instrument $Z_i = 0$, and $r_i = r_{1i}$ if it takes the value 1. If the instrument Z is independent of (b_i, r_{0i}, r_{1i}), we can rewrite

$$\text{plim}_N \hat{\beta}_{IV} = \frac{\mathbb{E}[y_i|Z_i = 1] - \mathbb{E}[y_i|Z_i = 0]}{\mathbb{E}[S_i|Z_i = 1] - \mathbb{E}[S_i|Z_i = 0]}$$

$$\begin{aligned} &= \frac{\mathbb{E}[b_iS_i - k_1S_i^2|Z_i=1] - \mathbb{E}[b_iS_i - k_1S_i^2|Z_i=0]}{\mathbb{E}[S_i|Z_i=1] - \mathbb{E}[S_i|Z_i=0]} \\ &\overset{indep}{=} \frac{\mathbb{E}[\Delta S_i b_i] - \frac{1}{2}k_1\mathbb{E}[\Delta(S_i^2)]}{\mathbb{E}[\Delta S_i]} \end{aligned}$$

because of independence of (r_{0i}, r_{1i}, b_i) and Z_i. From the last part, we know that

$$\Delta S_i = \frac{b_i - r_{1i}}{k_1+k_2} - \frac{b_i - r_{0i}}{k_1+k_2} = \frac{r_{0i} - r_{1i}}{k_1+k_2}$$

and

$$\begin{aligned} \Delta(S_i^2) &= S_{1i}^2 - S_{0i}^2 \overset{(*)}{=} (S_{1i} + S_{0i})(S_{1i} - S_{0i}) \\ &= 2\bar{S}_i\Delta S_i = 2\bar{S}_i\frac{r_{0i} - r_{1i}}{k_1+k_2} \end{aligned}$$

where the "trick" in (*) is to use the third binomial formula $(a+b)(a-b) = a^2 - b^2$ (though brute force would have given the same result). Plugging this back into the plim for the Wald estimator, and noticing that the $(k_1 + k_2)$ terms cancel out,

$$\text{plim}_N\hat{\beta}_{IV} = \frac{\mathbb{E}[(r_{0i} - r_{1i})b_i] - k_1\mathbb{E}[(r_{0i} - r_{1i})\bar{S}_i]}{\mathbb{E}[r_{0i} - r_{1i}]}$$

From this we can see that the Wald estimator weights individual returns by $r_{0i} - r_{1i}$, so that individuals who experience a high shift in the discount rate as a result of the intervention are "over-represented" relative to the other persons.
Card's argument about "discount rate bias" (he doesn't call it like that) is that the instrumental variables which are typically used in the literature induce individuals to increase their schooling levels who would have received relatively low *levels* schooling otherwise, but therefore also have relatively high *marginal returns*. This means that if we are interested in a "population average" of the parameters in the HCEF, our IV estimates of the return to schooling are biased upwards. On the other hand one could argue that the effect on the subpopulation which is affected by the instrument is a more interesting policy parameters as long as that subpopulation is similar to the population which would be affected by a particular policy measure. E.g. as its title already states, the Angrist and Krueger paper which used quarter of birth dummies as instruments for high-school dropouts adequately estimates the effect of mandatory schooling on kids for which the legal constraints are actually "binding"/who are at the margin of dropping out in the absence of mandatory schooling.

1.2 A Simple Model for Signaling

This part is based on chapter 2 of Daron's lecture notes. We are now in a very simple world in which there are firms which compete on the labor market, and there are workers who learn about their own productivity, η_L or η_H ($\eta_H > \eta_L$) at the beginning of the game. At a first stage, workers can choose between two levels of education: no education $e = 0$, and some education $e = 1$ which doesn't affect productivity (say, memorizing Greek tragedies). The cost of education level $e = 1$ is different for the two types, and $c(1, \eta_L) > c(1, \eta_H) > 0$ (note that this is the two-educational-level version of the single-crossing property $c_{e\eta} < 0$ introduced in class).
Firms observe a particular worker's education level, but not his productivity, but it is common knowledge that a fraction $p(\eta_H) = \lambda$ of workers has productivity η_H. They can make a wage offer that depends on the level of education, and if the worker accepts, a high type will produce $y(\eta_H) = \eta_H$, and a low type will produce $y(\eta_L) = \eta_L$. Firms' profits are $\pi = y(\eta) - w(e)$, and in a competitive equilibrium, expected profits must be zero. Workers' final payoff is $u = w(e) - c(e, \eta)$.

1.2.1 First-Best Solution

If firms observed each worker's productivity, the first welfare theorem tells us that a market equilibrium would be Pareto-efficient. In order to get there, use zero-profit condition to get $w(\eta) = \eta$, regardless of the level of education (since education doesn't enhance productivity), so that workers of both types would choose not to get any education, $e = 0$. This is clearly efficient, since all workers will be employed, and no resources are wasted on education.
But now let's go back to our original assumption that firms don't know individual types. Firms could ask workers about their types and continue to offer the first-best wage scheme $w(\eta) = \eta^*$ based on workers' reports η^*. However this clearly doesn't work: while high types have every incentive to tell the truth since $w(\eta_H) = \eta_H > \eta_L = w(\eta_L)$, low types are tempted to disguise themselves as η_H-types. Therefore, firms would make losses under the first-best wage scheme relying on voluntary reports by workers, since $\mathbb{E}\pi = \lambda\eta_H + (1-\lambda)\eta_L - \eta_H = (1-\lambda)(\eta_L - \eta_H) < 0$.

1.3 Bayesian Perfect Equilibrium in Signaling Games

In order to analyze formally what happens in this setting, we need an equilibrium concept that takes into account imperfect information and the dynamic structure of the problem while ruling out "unreasonable" outcomes which aren't compatible with rational expectations about other players' types or future behavior. This leads us to the notion of Bayesian Perfect Equilibrium, which in the context of this type of signaling games is defined as meeting the following requirements

1. (Expected Utility Maximization) Each type of workers choose their signal (i.e. education) as to maximize their utility given the wage schedule they will be offered by firms, i.e.

$$e^* = \arg\max_e w(e) - c(e, \eta)$$

2. (Beliefs) After observing a level of education, the firm must have *beliefs* $\mu(\eta|e)$ which correspond to the probability that a worker with education e actually is of type η, such that

$$\sum_{\eta \in \{\eta_L, \eta_H\}} \mu(\eta|e) = 1$$

3. (Bayes' Rule) As long as there is any type η who would have chosen educational level e with some probability in equilibrium, beliefs conditional on observing e must satisfy Bayes' rule, i.e.

$$\mu(\eta|e) = \frac{p(\eta)\mathbb{P}(e^* = e|\eta)}{\sum_{\tilde{\eta}} p(\eta)\mathbb{P}(e^* = e|\eta)}$$

where $\mathbb{P}(e^* = e|\eta)$ is the probability that type η chooses e in this particular equilibrium, and in our example $p(\eta_H) = \lambda$ and $p(\eta_L) = 1 - \lambda$.

4. (Zero Profits) Firms pay workers their expected marginal product,

$$\mathbb{E}[\pi|e] = \mathbb{E}[y(\eta)|e] - w(e) = 0$$

Conditions 1 and 4 are standard conditions on any kind of equilibrium (where zero profits is equivalent to profit maximization under perfect competition of firms on the labor market) - these 2 conditions look similar to the usual Nash equilibrium in a sequential game in which workers play e, and $w(e)$ is the firm's best response to any possible education level. Condition 2 explicitly includes beliefs as part of a complete description of the equilibrium - each equilibrium will turn out to be supported only by a particular constellation of beliefs, and in a sense we could think of a players' beliefs as part of his/her own strategy.

Condition 3 requires that players' beliefs are consistent with rational expectations about past play by the sender of the signal. It is explicitly agnostic about firms' beliefs at educational choices which were not supposed to happen in a given equilibrium, so that this definition allows for *any* beliefs off the equilibrium path. The reason for this omission is mainly technical (probabilities conditional on impossible events aren't defined), but also turns out to be the weak point of this equilibrium concept, since in some games, some equilibria are supported only by "unreasonable" beliefs at nodes which aren't reached in equilibrium.

1.4 Separating Equilibrium

A separating equilibrium, is a Perfect Bayesian Equilibrium (PBE) in which different types will chose different education levels, e_H and e_L. Imposing the conditions for a PBE one after another, note that the firm will perfectly back out workers' abilities (by condition 3 for a PBE, $\mu(\eta_H|e_H) = 1$ and $\mu(\eta_L|e_L) = 1$) and therefore pay each type their marginal product, i.e. $w(e_H) = \eta_H$ and $w(e_L) = \eta_L$ (by condition 4).

When we looked at the first-best wage rule under voluntary reporting of types, we noticed that only the low types had an incentive not to tell the truth, so the separating equilibrium has to impose an educational level which will make "lying" through the choice of a different type's signal unattractive for both types of workers (note that this is where condition 1 for PBE comes in), and this translates to the *incentive compatibility constraints*

$$
\begin{aligned}
\mathrm{IC}_H: \quad & w(e_H) - c(e_H, \eta_H) \quad \geq \quad w(e_L) - c(e_L, \eta_H) \\
\mathrm{IC}_L: \quad & w(e_L) - c(e_L, \eta_L) \quad \geq \quad w(e_H) - c(e_H, \eta_L)
\end{aligned}
$$

In our case there is only one way to satisfy IC_L if $w(e_H) = \eta_H > w(e_L) = \eta_L$, which is by setting $e_L = 0$ and $e_H = 1$, so that IC_L is satisfied whenever $\eta_L \geq \eta_H - c(1, \eta_L)$. IC_H holds if $\eta_H - c(1, \eta_H) \geq \eta_L$. Since a separating equilibrium exists if and only if both incentive compatibility constraints can be satisfied simultaneously, we have to require that

$$
c(1, \eta_L) \geq \eta_H - \eta_L \geq c(1, \eta_H)
$$

and this is where the single-crossing property $c(1, \eta_L) > c(1, \eta_H)$ comes in (necessary but not sufficient for existence of separating equilibrium). Remember that under "first-best contract with reporting of types", only the low types had an incentive to lie, and as a general rule, if e were continuous, the incentive compatibility constraints for the types who had that incentive to disguise themselves in that artificial situation will typically be "binding" and pin down the amount of signaling in a PBE (in a very broad sense - if as in our example educational levels are discrete, it will generically not hold with equality, but it will still introduce a distortion vis-à-vis the first-best), if other IC constraints are also violated, this will typically kill the separating equilibrium.

Of course this is all a far too pedantic and complicated way of solving our really simple model, but once the problem gets more complex, we will essentially do these same steps, and it's good to see how it works out in this simple example.

1.5 Pooling Equilibrium

In a pooling equilibrium, both types choose the same educational level $\bar{e}$, implying beliefs $\mu(\eta_H|\bar{e}) = \lambda$ and $\mu(\eta_L|\bar{e}) = 1 - \lambda$ (note that for now, beliefs conditional on any other educational level could in principle be anything). Therefore, the firm offers a wage $w(\bar{e}) = w(1) = w(0) = \lambda\eta_H + (1 - \lambda)\eta_L =: \bar{\eta}$ (notice that I also have to specify a wage for an educational level which isn't going to be chosen in equilibrium in order to give a full characterization of the PBE). This choice of $w(e)$ ensures that nobody has any incentive whatsoever to get education, so that $\bar{e} = 0$. I would like to point out that it is an artifact of this simple model that the pooling equilibrium is actually Pareto-efficient - this will in general not be true. However, one big problem with this equilibrium is that $w(1) = \bar{\eta}$ means that it is supported by unreasonable beliefs in the following sense: Assume that the payoffs are such that the separating equilibrium exists as well, and after everyone agreed to play a pooling equilibrium, an individual with education $e = 1$ shows up (and all competing firms observe that educational level). The firm would only pay the wage $w(1) = \bar{\eta}$ prescribed by this equilibrium if it believed that all types were equally likely to get this educational level. However the worker could make a "speech" saying that regardless of which equilibrium everyone wanted to play, low types would never deviate to the high education level no matter what. This means that even in the pooling equilibrium the firm should believe that a person choosing to get education should be a high-productivity type for sure (which is different from the beliefs $\mu(\eta_H|e = 1) = \lambda$ that induce firms to set a wage $\bar{\eta}$). There are equilibrium "refinements" on PBE which only keep the equilibria based on "reasonable" beliefs.