14.662, Spring 2015: Problem Set 3
Due Wednesday 22 April (before class)
Heidi L. Williams
TA: Peter Hull

1 Roy model: Chiswick (1978) and Borjas (1987)

Chiswick (1978) is interested in estimating regressions like the following for a single Census year T (in his case, $T = 1970$):

$$\ln(wage_i(T)) = \mathbf{X_i'}\theta + \delta I_i + \alpha_1 I_i \text{Years}_i + \alpha_2 I_i \text{Years}_i^2 + \beta_1 I_i \text{Arrive}_i + \beta_2 I_i \text{Arrive}_i^2 + \epsilon_i,$$

where I_i is an indicator for foreign-born, Years_i counts the number of years since migration, and Arrive_i is the calendar year of arrival.

1. By substituting $\text{Arrive}_i = T - \text{Years}_i$ into the above regression equation, show mathematically that in a single cross-section β_1, β_2, δ, α_1, and α_2 cannot be separately identified.

2. Re-write your new regression equation from part (1) to let γ_1 represent the coefficient on I_i, γ_2 represent the coefficient on $I_i\text{Years}_i$, and γ_3 represent the coefficient on $I_i\text{Years}_i^2$. What is $\frac{\partial \gamma_1}{\partial T}$, $\frac{\partial \gamma_2}{\partial T}$, and $\frac{\partial \gamma_3}{\partial T}$? Use these expressions to show that with two years of Census data (say, $T = 1970, 1980$) it is possible to identify β_1, β_2, δ, α_1, and α_2.

3. In order to identify both the assimilation effect and the cohort indicators while also controlling for Census year indicators, Borjas (1987) imposed the restriction that time-specific shocks have the same effect on log earnings of natives and immigrants. How might you assess the validity of this restriction?

2 Roy model: Rothschild and Scheuer (2013)

Rothschild and Scheuer (2013) characterize optimal taxation in a Roy model where individuals can self-select into one of multiple sectors based on relative potential skill. In this problem, you will use data from the Current Population Survey to replicate their results.

Consider an economy with a unit mass of individuals who can choose between working in either of two sectors. Each person has a two-dimensional skill endowment $(\theta, \varphi) \in \Theta \times \Phi$. The parameter θ captures an individual's productivity in the Θ-sector and φ captures her ability in the Φ-sector. These endowments are jointly distributed with CDF $F(\theta, \varphi)$. Let $S(\theta, \varphi) \in \{\Theta, \Phi\}$ denote a worker's chosen sector and $P_\Theta = \{(\theta, \varphi) \mid S(\theta, \varphi) = \Theta\}$ denote the set of types who choose the Θ-sector.

Individuals have preferences over consumption c and effort e given by

$$U(c, e) = c - \left(\frac{\epsilon}{1+\epsilon}\right) e^{\frac{1+\epsilon}{\epsilon}}$$

Aggregate effort in the Θ-sector is given by

$$E_\Theta \equiv \int_{P_\Theta} \theta e(\theta, \varphi) \, dF(\theta, \varphi)$$

for effort $e(\theta, \varphi)$, and likewise for aggregate effort in the Φ sector, E_Φ. Output is a Cobb-Douglas function of these aggregate effort levels.

$$Y = E_\Theta^\alpha E_\Phi^{1-\alpha}$$

for $\alpha \in (0, 1)$. Let $E \equiv E_\Theta / E_\Phi$ denote relative aggregate effort.

1. What simplifying assumptions are embedded in the functional form for preferences? In particular, what does the parameter ϵ capture? Use a short derivation from your undergrad micro days to justify your interpretation.

2. Assuming that effort is directly observed by employers, derive an expression for the wage of type (θ, φ) as a function of the equilibrium value of E. Use this result to argue that wages are invariant to the scale of θ and φ.

3. By your argument above, we will proceed as if the distribution of wages and skills coincide. Assume now that potential skills/wages are drawn from a bivariate lognormal distribution with means μ_θ and μ_φ, variances σ_θ^2 and σ_φ^2, and correlation coefficient ρ. We want to estimate these parameters from the observed distribution of wages. To do so, we will take advantage of a useful fact about the bivariate normal distribution (derived in Basu and Ghosh (1978)):

   Let X and Y be distributed bivariate normal with means μ_x and μ_y, variances σ_x^2 and σ_y^2, and correlation coefficient ρ. Let $Z = \max\{X, Y\}$. Then the density of Z is

   $$g(z) = \left(\frac{1}{\sigma_x}\right) \phi\left(\frac{z - \mu_x}{\sigma_x}\right) \Phi\left(\frac{z - \tilde{\mu}_y}{\tilde{\sigma}_y}\right) + \left(\frac{1}{\sigma_y}\right) \phi\left(\frac{z - \mu_y}{\sigma_y}\right) \Phi\left(\frac{z - \tilde{\mu}_x}{\tilde{\sigma}_x}\right)$$

   where

   $$\tilde{\mu}_x = \begin{cases} \frac{1}{\gamma_x}\left[\mu_x - (1 - \gamma_x)\mu_y\right] & \gamma_x \neq 0 \\ \mu_x - \mu_y & \gamma_x = 0 \end{cases}$$

   $$\tilde{\sigma}_x = \begin{cases} \frac{\sigma_x}{|\gamma_x|}\sqrt{1 - \rho^2} & \gamma_x \neq 0 \\ \sigma_x \sqrt{1 - \rho^2} & \gamma_x = 0 \end{cases}$$

and

$$\gamma_x = 1 - \rho\left(\frac{\sigma_x}{\sigma_y}\right)$$
$$\gamma_y = 1 - \rho\left(\frac{\sigma_y}{\sigma_x}\right)$$

Download the March 2011 CPS earnings and hours data from the NBER website. Generate a sample of log hourly wages from the weekly earnings and weekly hours data.[1] Please note that there is no extensive margin for labor force participation in this model, so you can restrict your attention to the subset of respondents with positive, non-missing wages. Use the fact above to estimate the parameters of the bivariate wage distribution.[2] Use your estimates to generate a predicted wage distribution and plot your prediction against the distribution observed in the CPS.

4. Assume that the elasticity of labor supply is 0.5 and that all workers face a marginal tax rate of 0.25 on their wages. Use these values and your estimates from part (3) to determine:

   - the effort supplied by each worker in her chosen sector
   - the share of income paid to each sector
   - the parameter α that governs the aggregate production function.

   Report and interpret your estimate of α here. You do not need to report anything for the first two results; they're simply intermediate steps.

5. Plot the optimal tax schedule derived by Rothschild and Scheuer (2013) and provided in MTR.mat. Interpret the shape of this schedule. Taking the schedule as given, show how the share of workers in the Θ-sector varies with the sector's offered wage. How does the average effort of Θ-sector workers move with wages? Plot both of these results and interpret.

[1] The relevant sample weight for the earnings variables is earnwt.

[2] You may wish to experiment with different optimization packages and starting values in running this MLE; you might also find that trimming the distribution of raw log wages to drop extreme outliers improves the stability of your estimators.

3 Compensating Differences: Lucas (1977) and Brown (1980)

Suppose true earnings are described by:

$$\ln(wage_{it}) = \beta_0 + \beta_1 Z^*_{it} + \beta_2 X_{it} + \beta_3 A_i + \epsilon_{it}$$

where Z^*_{it} measures working conditions, X_{it} measures observed time-varying worker characteristics, A_i measures unobserved fixed worker characteristics, and ϵ_{it} measures other unobserved factors that affect earnings (such as unmeasured job characteristics). Assume ϵ_{it} is orthogonal to Z^*_{it}, X_{it}, and A_i.

You would like to estimate β_1 - the compensating wage differential paid to workers to offset the disutility of working in jobs with higher levels of the disamenity Z^*_{it}. In practice, you face two estimation problems:

- "Ability" A_i is unobserved; suppose A_i is negatively correlated with Z^*_{it} conditional on X_{it}.
- Working conditions Z^*_{it} are measured with error. For example, measurement error could arise if you assigned job characteristics to a survey of workers using a match to the Dictionary of Occupational Titles data based only on occupation and industry, and if that occupation-industry match of job characteristics does not perfectly correspond to characteristics in the worker's specific job. In particular, suppose you observe a noisy measure of working conditions $Z_{it} = Z^*_{it} + \eta_{it}$.

We'd like to consider the net effect of these two potential sources of bias as well as possible solutions to estimating the compensating differential.

1. Say that you estimate a cross-sectional model as in Lucas (1977):

   $$\ln(wage_{it}) = b_0 + b_1 Z_{it} + b_2 X_{it} + e_{it}$$

   For a given t. Suppose in the cross section η_{it} is distributed as independent white noise. Derive an expression for the population regression coefficient b_1 in terms of structural parameters. Can you sign the overall direction of bias?

2. Say that, like Brown (1980), you find a panel dataset that allows you to estimate a model with individual fixed effects and you estimate

   $$\Delta \ln(wage_{it}) = b_1 \Delta Z_{it} + b_2 \Delta X_{it} + \Delta e_{it}$$

   Suppose within-individual measurement error is persistent, so that $\eta_{it} = \rho \eta_{it-1} + \nu_{it}$ where ν_{it} is independent (across both time and individuals) white noise. Derive an expression for the population regression coefficient b_1 in terms of structural parameters.

3. Briefly discuss what problem(s) are solved by moving to the panel model, and what problem(s) are introduced. Is it always the case that b_1 from the panel model will be an attenuated estimate of β_1, so that we can at least consistently estimate the sign of the compensating differential?

4 Compensating differences: Gruber and Krueger (1991) and Gruber (1997)

Consider the formalization of the Summers (1989) model from Gruber and Krueger (1991). Suppose that labor demand (L_d) is given by:

$$L_d = f_d(W + C)$$

and suppose labor supply (L_s) is given by:

$$L_s = f_s(W + \alpha C)$$

where C is the cost of mandated health insurance, αC is the monetary value that employees place on health insurance, and W is the wage rate.

1. Derive an expression for how wages change under a mandate ($\frac{dW}{dC}$) in terms of α, the labor demand elasticity η^d, and the labor supply elasticity η^s. Derive an analogous expression for how employment changes under a mandate. Give an intuition for the cases where $\alpha = 0$ and $\alpha = 1$.

2. Draw a graph of employment (x-axis) against wages (y-axis) with labor supply and labor demand curves before and after the mandated benefit regime. Give some intuition for how to interpret the graph.

3. Describe how the effects of a payroll tax on wages and employment might differ from the effects of a mandated benefit. Would it matter whether the payroll tax collections were used to finance a public health insurance program? What if the public health insurance program had enrollment that was restricted to workers only?

4. Read over the Gruber (1997) paper. He discusses three potential explanations for his results: inelastic labor supply, perfectly elastic labor demand, and full employee valuation of benefits. How might you distinguish between these three potential explanations?

References

Basu, Asit and Jayanta Ghosh, "Identifiability of the Multinormal and other Distributions under Competing Risks Model," *Journal of Multivariate Analysis*, 1978, *8* (3), 413–429. 3

Borjas, George, "Self-selection and the earnings of immigrants," *American Economic Review*, 1987, *77* (4), 531–553. 3

Brown, Charles, "Equalizing differences in the labor market," *Quarterly Journal of Economics*, 1980, *94* (1), 113–134. 2

Chiswick, Barry, "The effect of Americanization on the earnings of foreign-born men," *Journal of Political Economy*, 1978, *86* (5), 897–921. 1

Gruber, Jonathan, "The incidence of payroll taxation: Evidence from Chile," *Journal of Labor Economics*, 1997, *15* (S3), S72–S101. 4

_ **and Alan Krueger**, "The incidence of mandated employer-provided insurance: Lessons from workers' compensation insurance," in David Bradford, ed., *Tax Policy and the Economy Volume 5*, 1991, pp. 111–143. 4

Lucas, Robert, "Hedonic wage equations and psychic wages in the returns to schooling," *American Economic Review*, 1977, *67* (4), 549–558. 1

Rothschild, Casey and Florian Scheuer, "Redistributive Taxation in the Roy Model," *Quarterly Journal of Economics*, 2013, *128* (2), 623–668. 2

Summers, Lawrence, "Some simple economics of mandated benefits," *American Economic Review Papers and Proceedings*, 1989, *79* (2), 177–183. 4

14.662 Labor Economics II
Spring 2015