Exam #2

Economics 435: Quantitative Methods

Spring 2006

1 Applications

Answer ONE of questions a, b, and c below.

Suppose you want to measure the effect of economics education on savings behavior. You have a panel data set observing a random sample of working-age Canadians in April 2005 and in April 2006 with the following variables:

- ECON: a binary variable that equals one if the respondent has ever taken a university-level course in economics.

- UNIV: a binary variable that equals one if the respondent has a university degree.

- INCOME: The respondent's total income from all sources in the previous calendar year.

- SAVINGS: The respondent's total savings in the previous calendar year (can be negative).

- RRSP: The respondent's total RRSP contributions in the previous calendar year (cannot be negative).

- STOCKS: The proportion of the respondent's previous-year RRSP contributions that were invested in stocks or equity-based mutual funds. If the respondent had no previous-year RRSP contributions, this quantity is reported as -999.

- SEX, AGE: The respondent's sex and current age in years.

For the question you choose, describe how you would use this data set to answer the question. Your description should include enough detail that someone could use your description to estimate the exact model you have in mind using EViews. In particular:

- Specify all explanatory and outcome variables. All variables must be either in the list above or some functional transformation of those variables (for example, you could include the natural logarithm of income, age squared, etc.).

- Specify the type of model (OLS, fixed effect, Tobit, Heckman, probit, etc.)

- Specify the value of any other quantities that must be chosen in order to estimate the model. For example, if you are estimating a Tobit, specify the upper and/or lower bound for y.

- If you are estimating a Tobit model, specify whether the observed outcome (y) or the unobserved outcome (y^*) is the outcome of interest.

In addition to describing your methodological approach, please identify your biggest concern with using this approach to answer the question. For example, your biggest concern might be a particular omitted variable.

Note that there is no requirement that you use every variable or every observation in the dataset. Some of the data I have described above may not be particularly helpful in answering any of the questions.

a) What is the effect of economics education on savings rates?

b) What is the effect of economics education on RRSP[1] contribution choices?

c) What is the effect of economics education on asset allocation (whether an economics education is associated with relatively greater or lesser investment in equities)?

2   A variation on the Heckman correction

In your textbook and in the lecture notes, the Heckman correction model is used to handle *incidental truncation*, i.e. a situation in which the outcome y is only observed if $s = 1$ where s is some binary random variable.

A similar model is used to handle *selection into treatment*, i.e., a situation in which the outcome y is affected by some binary random variable s, but s is itself affected by most factors that also affect y. A common application is in measuring the effect of participation in job-training programs on subsequent labour market outcomes.

Here's the model. Let (y, x, z, s, u, v) be a set of random variables such that:

$$
\begin{aligned}
y &= \beta_0 + \beta_1 x + \beta_2 s + u \\
s &= I\left(\gamma_0 + \gamma_1 x + \gamma_2 z + v\right) \\
\{(y_i, x_i, z_i, s_i)\}_{i=1}^{n} &\quad \text{is a random sample we observe} \\
\left[\begin{array}{c} u \\ v \end{array}\right]\Bigg| x, z &\sim N\left(\left[\begin{array}{c} 0 \\ 0 \end{array}\right], \left[\begin{array}{cc} \sigma^2 & \rho\sigma \\ \rho\sigma & 1 \end{array}\right]\right) \\
v\hat{a}r(x) > 0, &\quad v\hat{a}r(z) > 0
\end{aligned}
$$

where $I(.)$ is the indicator function.

a) Show that:

$$E(y|x, z, s) = \beta_0 + \beta_1 x + \beta_2 s + E(u|x, z, s)$$

Is the CEF $E(y|x, z, s)$ identified in our data?

b) It can be shown (after a substantial amount of integral calculus) that:

$$
\begin{aligned}
E(u|x, z, s = 1) &= \rho \frac{\phi(\gamma_0 + \gamma_1 x + \gamma_2 z)}{\Phi(\gamma_0 + \gamma_1 x + \gamma_2 z)} \\
E(u|x, z, s = 0) &= -\rho \frac{\phi(\gamma_0 + \gamma_1 x + \gamma_2 z)}{1 - \Phi(\gamma_0 + \gamma_1 x + \gamma_2 z)}
\end{aligned}
$$

where $\phi(.)$ is the PDF of the $N(0, 1)$ distribution and $\Phi(.)$ is its CDF.

---

[1]Since you are a student you may or may not be familiar with the RRSP rules. RRSP's are tax-deferred retirement savings accounts. Every dollar that you put into an RRSP reduces this year's taxable income by one dollar. RRSP funds may not be withdrawn until the individual retires (there are a few exceptions we will ignore), and are taxed upon withdrawal. A person's RRSP contributions in a given year have a lower bound of zero and an upper bound ("contribution limit") that varies by person. The contribution limit of an individual in a given year is 18% of the previous year's earned income to a maximum of $16,500 plus any unused contribution limits carried forward from previous years.

Let $\lambda(x, z, s)$ be defined as:

$$\lambda(x, z, s) \equiv s\frac{\phi(\gamma_0 + \gamma_1 x + \gamma_2 z)}{\Phi(\gamma_0 + \gamma_1 x + \gamma_2 z)} - (1 - s)\frac{\phi(\gamma_0 + \gamma_1 x + \gamma_2 z)}{1 - \Phi(\gamma_0 + \gamma_1 x + \gamma_2 z)}$$

Find $E(y|x, z, s)$ in terms of the random variables x, s, and $\lambda(x, z, s)$, and the model's parameters (i.e., β_0, β_1, and ρ).

c) Describe a procedure for consistently estimating the parameters of this model.