Empirical Analysis of Micro Data: Suggested Answer for Final Examination

Fumio Ohtake and Daiji Kawaguchi

January 30, 2003

1 Short Questions

1.1 Panel Data Method (20pts)

Suppose you are interested in the estimation of the following equation:

$$y_{it} = x_{it}\beta + c_i + u_{it}, \tag{1}$$

where x_{it} is the vector of independent variables, c_i is the individual unobserved heterogeneity and u_{it} is the idiosyncratic error at time t for individual i. Assume that all of the columns in vector x are linearly independent and time-variant.

1.1.1 Pooled OLS Estimator (5pts)

What is the additional assumptions needed for pooled OLS estimator to be consistent. Applying the usual OLS estimation, do you obtain the consistent estimator of the variance of $\hat{\beta}$?

`Let` $v_{it} = c_i + u_{it}$`, the pooled OLS estimator is consistent if` $Ex'_{it}v_{it} = 0$ `holds.`

`Alternative Answer`

$E(c_i|x_{it}) = 0$ and $E(u_{it}|c_i, x_{it}) = 0$.

1.1.2 Random Effects Estimator (5pts)

What is the additional assumptions needed for the random effects estimator to be consistent. Write down two assumptions.

$E(c_i|x_i) = 0$ and $E(u_{it}|x_i, c_i) = 0$ where $x_i = [x_{i1}, x_{i2}, ..., x_{iT}]$.

1.1.3 Fixed Effects Estimator (5pts)

What is the assumption needed for the fixed effects estimator to be consistent.

$E(u_{it}|x_i, c_i) = 0$

1.1.4 RE vs. FE (5pts)

Describe the statistical procedure how you would test whether the random effects estimator or the fixed effects estimator is more preferable. What is the merit of using the random effects estimator over the fixed effects estimator given the assumptions for the random effects estimator to be consistent hold?

Implement the Hausmant Test. (Appropriate Explanation for the Hausman test is required.) The random effects estimator is efficient if the composite error structure is such that the random effects estimator assumes.

1.2 Random Trend Models (15pts)

1.2.1 Linear Random Trend (5pts)

In the class, you are taught how to estimate the model

$$y_{it} = c_i + g_i t + x_{it}\beta + u_{it}, \tag{2}$$

in which c_i and g_i are allowed to be correlated with x_{it}. Under the strict exogeneity assumption $E(u_{it}|x_i, c_i, g_i) = 0$, describe the estimation procedure to obtain a consistent estimator of β.

`Take the first difference to get`

$$\Delta y_{it} = g_i + \Delta x_{it} + \Delta u_{it}.$$

`Apply fixed effects or first difference estimation for this equation.`

1.2.2 Quadratic Random Trend (10pts)

Assume the quadratic random trend instead. The equation becomes

$$y_{it} = c_i + g_i t + h_i t^2 + x_{it}\beta + u_{it}, \tag{3}$$

in which c_i, g_i and h_i can be correlated with x_{it}. Suggest the procedure to obtain the consistent estimator of β. Explicitly write the assumption needed for the consistency of β.

`The equation for` $t-1$ `is`

$$y_{it-1} = c_i + g_i(t-1) + h_i(t-1)^2 + x_{it-1}\beta + u_{it-1}.$$

`Take the first difference to get`

$$\Delta y_{it} = g_i + h_i 2t - h_i + \Delta x_{it}\beta + \Delta u_{it}.$$

`This is equivalent to the linear time trend model. Taking one more difference gives`

$$\Delta\Delta y_{it} = h_i 2 + \Delta\Delta x_{it}\beta + \Delta\Delta u_{it}.$$

`Then estimate the model by fixed effects or first difference estimation. The strict exogeneity assumption` $E(u_{it}|x_{i1}, x_{i2}, ..., x_{iT}, c_i) = 0$ `is needed for the consistency of the suggested estimator.`

1.3 Sample Selection Bias (25pts)

Suppose that you are to estimate the wage offer function for the population of all women in the working age, say age between 25 to 50, including those who do not actually work. In particular, you are interested in the estimation of "marriage penalty," which is the measure that how much lower wage married women earn than unmarried women with similar characteristics. To implement the idea, you are considering to estimate

$$\ln w^o = x\beta + married\gamma + u, \tag{4}$$

where x is the usual demographic variables and *married* is the dummy variable that takes 1 if the woman is married and 0 otherwise. In the equation, γ is the marriage penalty (or premium if γ is positive number). Without the sample selection, x and *married* are assumed to be exogenous (i.e. $E(u|x, married) = 0$). However, the sample only includes those women who actually work. The sample selection indicator, s, takes the value 1 for those observations selected into the sample and 0 otherwise.

1.3.1 Ignorability (5pts)

The OLS estimation applied to the selected sample renders consistent estimator of β and γ when the selection is "ignorable." Write down the definition of the ignorability of the selection using conditional expectation of u in terms of s, x and *married*.

$$E(u|x, married, s) = E(u|x, married) = 0.$$

1.3.2 Intuition for Sample Selection Bias (5pts)

Married women are less likely to work because of their involvement in the household production. Thus those women who work although they are married are considered to have advantage in market production or to have have unobserved high wage factor. What kind of correlation between *married* and u conditioned on $s = 1$ is implied from this intuition. What kind of bias would you expect for the OLS estimator of γ using only working women as analysis sample.

`We expect the upward bias for the estimator of` γ `in prior.`

1.3.3 Heckman Correction (10pts)

Labor economists have routinely used Heckman's two-step self selection correction to get the consistent estimator for β and γ. Describe the procedure step-by-step assuming that s is determined by Probit structure. Write down the first stage Probit equation and the second stage linear model including inverse Mill's ratio. To assure that the identification does not depend of the normality of the error term for the first stage sample selection equation, what kind of assumption (additional information) would you need? Give a concrete example for the information for this specific application.

Firstly run the Probit regression $P(s = 1|z) = \Phi(z\gamma)$, where z is the set of exogenous variables for each individual that affects the decision to work or not. Then using the estimated coefficient γ, calculate the inverse Mill's ratio $\phi(z\hat{\gamma})/\Phi(z\hat{\gamma})$. Finally regress y on x, *married* and the inverse Mill's ratio by OLS. The standard error should be adjusted appropriately. For a credible identification, `the variable that enters the first stage Probit but excluded from the second stage OLS is required. In this application, husband's income would serve as the excluded variable.`

1.3.4 Math. Question (5pts)

This question asks the derivation of inverse Mill's ratio. Suppose $v \sim Normal(0,1)$. Show that

$$E(v|v > -z\gamma) = \frac{\phi(z\gamma)}{\Phi(z\gamma)}, \tag{5}$$

where ϕ is the standard normal density function and Φ is the standard normal distribution function.

$$\begin{aligned}
E(v|v > -z\gamma) &= \int_{-z\gamma}^{\infty} v(\phi(v)/\Phi(z\gamma))dv \\
&= \frac{[-\phi(v)]_{-z\gamma}^{\infty}}{\Phi(z\gamma)} \\
&= \frac{\phi(-z\gamma)}{\Phi(z\gamma)} \\
&= \frac{\phi(z\gamma)}{\Phi(z\gamma)}.
\end{aligned}$$

2 Measurement Error in Fixed Effects Model (10pts)

Consider a linear model with a single explanatory variable:

$$y_{it} = \beta x_{it}^* + c_i + u_{it} \tag{6}$$

under the strict exogeneity assumption

$$E(u_{it}|x_i^*, x_i, c_i) = 0 \tag{7}$$

where x_{it} denotes the observed measure of the unobservable x^*_{it}. Denote the measurement error as $v_{it} = x_{it} - x^*_{it}$ and assume that v_{it} is uncorrelated with x^*_{it}. Under the assumption $E(v_{it}|x^*_i, c_i) = 0$ and $Var(v_{it} - \bar{v}_i)$ and $Var(x^*_{it} - \bar{x}_i)$ are constant across t, show that the plim of the FE estimator is

$$plim_{N\to\infty}\hat{\beta}_{FE} = \beta(1 - \frac{Var(v_{it} - \bar{v}_i)}{[Var(x^*_{it} - \bar{x}^*_i) + Var(v_{it} - \bar{v}_i)]}). \tag{8}$$

The equation that is actually estimated is

$$y_{it} = \beta x_{it} + c_i + (\beta v_{it} + u_{it})$$

The probability limit of the fixed effects estimator is

$$plim_{N\to\infty}\hat{\beta}_{FE} = \beta + \frac{cov(x_{it} - \bar{x}_i, \beta(v_{it} - \bar{v}_i) + u_{it} - \bar{u}_i)}{var(x_{it} - \bar{x}_i)}$$

following the usual calculation for the bias in linear models. Then replacing x_{it} with $x^*_{it} + v_{it}$ and using assumptions, the result follows.

3 Friction Model (30pts)

Considering macroeconomic implications, the nominal rigidity of wages is a very important topic to analyze. Suppose that there is a wage level that would be prevailed if there were no friction in the wage determination at all and call the wage level as "notional wage" at time t for individual i denoted as w^*_{it}. However because of the friction in the wage determination, actual wage level, w_{it} is not always at the level of the notional wage and the current actual wage level tends to stick to the past wage level. To capture the idea above, you wrote down the following econometric model.

$$w^*_{it} = x_{it}\beta + e_{it}, \quad e_{it}|x_{it} \sim Normal(0, \sigma^2), \tag{9}$$

$$
\begin{aligned}
w_{it} &= w_{it}^*, \ if \ w_{it-1} + \alpha_2 < w_{it}^*, && (10)\\
w_{it} &= w_{it-1}, \ if \ w_{it-1} - \alpha_1 \le w_{it}^* \le w_{it-1} + \alpha_2, && (11)\\
w_{it} &= w_{it}^*, \ if \ w_{it}^* < w_{it-1} - \alpha_1, && (12)
\end{aligned}
$$

where x_{it} is the usual demographic variable in the wage equation, α_1 and α_2 are both positive constant. In particular α_2 is the parameter expressing upward "stickiness" and α_1 is the parameter expressing downward "stickiness."

3.1 Probability to fall into each category (10pts)

Calculate $P(w_{it} > w_{it-1}|x), P(w_{it} < w_{it-1}|x)$ and $P(w_{it} = w_{it-1}|x)$ respectively as functions of w_{it-1}, x_{it} and parameters.

$$
\begin{aligned}
P(w_{it} > w_{it-1}|x) &= P(w_{it}^* > w_{it-1} + \alpha_2|x)\\
&= P(x_{it}\beta + e_{it} > w_{it-1} + \alpha_2)\\
&= P(e_{it} > -x_{it}\beta + w_{it-1} + \alpha_2)\\
&= \Phi((x_{it}\beta - w_{it-1} - \alpha_2)/\sigma).
\end{aligned}
$$

$$
\begin{aligned}
P(w_{it} < w_{it-1}|x) &= P(w_{it}^* < w_{it-1} - \alpha_1|x)\\
&= P(x_{it}\beta + e_{it} < w_{it-1} - \alpha_1)\\
&= P(e_{it} < -x_{it}\beta + w_{it-1} - \alpha_1)\\
&= \Phi((-x_{it}\beta + w_{it-1} - \alpha_1)/\sigma).
\end{aligned}
$$

Since the event $w_{it} = w_{it-1}$ is the complement of the two events above,

$$
P(w_{it} = w_{it-1}|x) = 1 - \Phi((x_{it}\beta - w_{it-1} - \alpha_2)/\sigma) - \Phi((-x_{it}\beta + w_{it-1} - \alpha_1)/\sigma). \qquad (13)
$$

or equivalently,

$$P(w_{it} = w_{it-1}|x) \;=\; \Phi((-x_{it}\beta + w_{it-1} + \alpha_2)/\sigma) - \Phi((-x_{it}\beta + w_{it-1} - \alpha_1)/\sigma). \tag{14}$$

3.2 Conditional Density (10pts)

Write down the conditional density function $f(w_{it}|x_{it}, w_{it} > w_{it-1})$ and $f(w_{it}|x_{it}, w_{it} < w_{it-1})$ respectively as functions of w_{it}, x_{it}, w_{it-1} and parameters.

$$f(w_{it}|x_{it}, w_{it} > w_{it-1}) = \frac{\frac{1}{\sigma}\phi((w_{it} - x_{it}\beta)/\sigma)}{\Phi((x_{it}\beta - w_{it-1} - \alpha_2)/\sigma)}, \tag{15}$$

$$f(w_{it}|x_{it}, w_{it} < w_{it-1}) = \frac{\frac{1}{\sigma}\phi((w_{it} - x_{it}\beta)/\sigma)}{\Phi((-x_{it}\beta + w_{it-1} - \alpha_1)/\sigma)}. \tag{16}$$

3.3 Log Likelihood Function (10pts)

Using the solution for the problems above, write down the log likelihood function for each observation i depending on three cases $w_{it} > w_{it-1}, w_{it} = w_{it-1}$, and $w_{it} < w_{it-1}$.

For $w_{it} > w_{it-1}$,

$$\log f(w_{it}|x_{it}) = \log[f(w_{it}|x_{it}, w_{it} > w_{it-1})P(w_{it} > w_{it-1}|x_{it})] = \log[\frac{1}{\sigma}\phi((w_{it} - x_{it}\beta)/\sigma)].$$

For $w_{it} < w_{it-1}$,

$$\log f(w_{it}|x_{it}) = \log[f(w_{it}|x_{it}, w_{it} < w_{it-1})P(w_{it} < w_{it-1}|x_{it})] = \log[\frac{1}{\sigma}\phi((w_{it} - x_{it}\beta)/\sigma)].$$

For $w_{it} = w_{it-1}$,

$$\log f(w_{it} = w_{it-1}|x_{it}) = \log[\Phi(-x_{it}\beta + w_{it-1} + \alpha_2) - \Phi(-x_{it}\beta + w_{it-1} - \alpha_1)].$$