Midterm Examination

Economics 522

Posted 19 November 2021

Submitting Your Answers:

Please email your answers to me by 9:00am ET on Monday November 22nd. This is an open-book exam; you are free to use class notes and other resources. Of course, you must *work independently of other students*—collaboration during an exam is a violation of the Honor Code. Be sure to *answer the questions in your own words*: Responses that are simply copied from other sources will not receive credit. Your answers can be provided in a LaTeX or Word document, or submitted as photos of what you've written down on paper (in this case, make sure that the full answer can be seen and that your writing is legible).

Each question counts for one-fourth of the score for the exam as a whole. To be eligible for full credit, you must provide *explanations, in words* along with your mathematical calculations. Define any notation that you use, and be sure to indicate the dimensions of matrices and vectors. If you need to make additional assumptions to answer a question, please state those assumptions and explain why they are needed.

1. Consider a just-identified method-of-moments model for an *independent but not necessarily identically distributed* (i.n.i.d.) data-generating process with parameter θ, which is a $k \times 1$ vector. You have a $k \times 1$ vector of functions $g_i(\theta)$, each element of which depends on the data for observation i as well as the unknown θ parameter. We denote by θ_0 the true value of the θ parameter. Because the number of functions in $g_i(\theta)$ is the same as the number of θ parameters to be estimated, we term this a "just-identified" model and would estimate θ by solving k (nonlinear) equations.

   (a) What is the moment condition for this model?

   (b) Assuming that the moment condition is satisfied, derive the variance of the asymptotic distribution of $\sqrt{n}(\hat{\theta} - \theta_0)$, in which n is the sample size.

   (c) How would you test the null hypothesis that the true θ_0 satisfies the set of linear restrictions $\mathbf{R}\theta_0 = \mathbf{r}$, against the alternative hypothesis that $\mathbf{R}\theta_0 \neq \mathbf{r}$? Here, $\mathbf{R}$ is an $m \times k$ matrix (with $m < k$) composed of known constants, and $\mathbf{r}$ is a known $m \times 1$ vector.

2. Consider a dynamic nonlinear panel-data model with a multiplicative fixed effect,

$$Y_{it} = \phi(\mathbf{X}_{it}, \theta)\, u_i + \alpha Y_{it-1} + \epsilon_{it}.$$

   You know the functional form of $\phi()$ but not the true value of the θ parameter, a $k \times 1$ vector. Let θ_0 denote its true value. You can assume that the ϕ function is differentiable in θ. Each unit i provides T observations over time; the data series are independent over i but within unit i's records the data can be correlated and heteroskedastic.

(a) How would you consistently estimate θ? Please be sure to state all assumptions you make in proving that your approach yields consistency, justify those assumptions to the extent possible, and show exactly how they contribute to your proof.

3. A simple structural model of wages $Y_{i,t}$ for worker i at time t is

$$Y_{it} = \alpha + t \cdot \beta + D_{it}\delta + u_i + \epsilon_{it}.$$

This model contains a constant, a time trend, a dummy variable D_{it} which takes the value 1 if the worker has chosen to participate in a job-training program on or before time t, and an error-components disturbance term. You have $T = 2$ time-points of data for each worker: Time $t = 0$ denotes a period before the training program was offered; time $t = 1$ denotes a period after the program was offered.

If we think of the u_i component as representing the worker's motivation, we would suspect that more motivated workers (those with higher values of u_i and also higher wages, other things being equal) are the kind of people who are more likely to take advantage of opportunities for job training. That is, it seems likely that u_i and D_{it} will be positively correlated.

(a) Describe how to estimate the program effect δ by the method of difference-in-differences. What are the important assumptions of this method?

(b) How would you estimate the *standard errors* of the limiting distribution of $\sqrt{n}(\hat{\delta} - \delta)$ in such a way that the estimates are robust to heteroskedasticity and serial correlation in ϵ_{it}?

(c) Suppose the model is altered to $Y_{it} = \mathbf{X}'_{i,t}\gamma + D_{it}\delta + u_i + \epsilon_{it}$. Discuss how to estimate the program effect δ for this model.

4. Consider a model in which a woman's earnings depend on her participation in a job-training program. If woman i does not participate in the program, her earnings are given by $Y_{i,0}$ and if she participates, her earnings are given by $Y_{i,1}$. That is, $(Y_{i,0}, Y_{i,1})$ are the potential earnings outcomes, only one of which is ever observed (by the researcher). The equations below describe the set-up:

$$\begin{aligned} Y_{i,0} &= \mu_0 + \epsilon_{i,0} \\ Y_{i,1} &= \mu_1 + \epsilon_{i,1}. \end{aligned}$$

We assume that the disturbance terms $(\epsilon_{i,0}, \epsilon_{i,1})$ are distributed as $\mathcal{N}(0, \Sigma)$ with covariance matrix

$$\Sigma = \begin{bmatrix} \sigma_{00} & \sigma_{01} \\ \sigma_{10} & \sigma_{11} \end{bmatrix}.$$

If you need to make any additional assumptions, please be sure to state and explain them.

- Suppose that woman i knows both (μ_0, μ_1) and $(\epsilon_{i,0}, \epsilon_{i,1})$. Equipped with this knowledge, *she decides to participate or not, choosing the option that maximizes her earnings*.

  What is the expected value of earnings among all women who *decide to participate* in the program?

- Consider the group of women who decide to participate in the program. What would have been the expected value of their earnings had they *not* participated? (This is the "counterfactual".)
- Given a set of observations limited to participants only, describe an approach in which you attempt to estimate μ_1 by the method of nonlinear least squares. What problems would you face? Be specific.
- Now suppose that women *do not choose* whether to participate, *but instead are randomly assigned* either to the job-training program or to a control group that is given no training. Assume that all women comply with their assignments.

  How can we consistently estimate the average program effect $\mathrm{E}(Y_{1,i} - Y_{0,i})$? Provide a full proof.