Econ 6200: Econometrics II
Prelim, April 8^{th}, 2021

This exam consists of ten questions, not of equal length or difficulty, grouped into three exercises. The questions are only partly cumulative; for example, 1.1-3 are not presupposed thereafter. Each question is worth 10 points. Remember to always explain your answer.

Good luck!

1

1.1 Consider the *Simple Linear Regression* problem of estimating

$$y = \beta_0 + \beta_1 z^* + \varepsilon$$

from i.i.d. data. The catch is that you do not observe realizations of (y, z^*) but of (y, z), where $z = dz^* + (1-d)\tilde{z}$, d is i.i.d. Bernoulli with parameter π, and $\tilde{z}$ is i.i.d. (and independent of all other r.v.'s in this exercise) with the same distribution as z^*. Other than this, the usual OLS assumptions hold; in particular, ε is independent of x.

The motivation for this question is *hot deck imputation*: Imagine that some observations of z^* were missing and were replaced with independent realizations from the marginal distribution of z^*. Our model reflects the relatively benign assumption that the data are *missing at random.* Note that we do not know which data points were imputed.

All regressions in this question are understood to be "on x," i.e. no attempt was made to identify imputed data points.

1.1 Let $\hat{\beta}_1$ denote the standard OLS estimator. Prove: $\hat{\beta}_1 \xrightarrow{p} \pi\beta_1$.

1.2 Let $\hat{\gamma}_1$ be the OLS slope estimator from the *reverse regression* $z = \gamma_0 + \gamma_1 y + \eta$. Give an exact expression for $\gamma_1 \equiv \text{plim}_{n\to\infty} \hat{\gamma}_1$. Explain why your answers up to here can be used to estimate *bounds* on the true β_1.

1.3 Say you also observe r.v. x with $\text{cov}(x, z^*) \neq 0$ but $\text{cov}(x, \varepsilon) = 0$. Let $\hat{\beta}_1^{IV}$ be the slope copefficient from an IV regression of y on z using x as instrument. Give an exact expression for $\text{plim}_{n\to\infty} \hat{\beta}_1^{IV}$. Explain how the asnwer allows you to estimate β_1 and π (as wqell as β_0).

1.4 The answer to **1.3** can be recast as GMM with moment conditions

$$\begin{aligned}\mathbb{E}(y-\beta_0-\pi\beta_1 z) &= 0\\ \mathbb{E}\big(z(y-\beta_0-\pi\beta_1 z)\big) &= 0\\ \mathbb{E}\big(x\big(y-\beta_0-\tfrac{\beta_1}{\pi}z\big)\big) &= 0.\end{aligned}$$

Explain why and verify the expressions. (Also, this question contains a strong hint regarding **1.3**.)

1.5 In this example, is there an efficiency gain from employing two-stage GMM, i.e. initially estimating a weighting matrix? Explain.

1.6 Explain how you would test whether the true probability of imputation $(1-\pi)$ is zero. (Give a test statistic and how you would compute the critical value.)

2 Consider linear GMM estimation under the exact assumptions from lecture, except that (i) the weighting matrix is the identity matrix, (ii) the true data generating process has

$$\mathbb{E}\big(\boldsymbol{x}(y-\boldsymbol{z}\boldsymbol{\beta}_0)\big)=\begin{bmatrix}0\\0\\\vdots\\0\\1\end{bmatrix}.$$

(However, the econometrician analyzing the data assumes that the right-hand side above is $\mathbf{0}$.)

Suppose initially that $\boldsymbol{z}$ and $\boldsymbol{x}$ are of the same length. Also, assume regularity conditions, e.g. that $\mathbb{E}\boldsymbol{z}\boldsymbol{x}'$ has full rank.

2.1 Without resorting to general results about extremum estimation, argue that

$$\hat{\boldsymbol{\beta}}\xrightarrow{p}\boldsymbol{\beta}^*\equiv\arg\min_{\boldsymbol{\beta}}\big\{\mathbb{E}\big(\boldsymbol{x}(y-\boldsymbol{z}\boldsymbol{\beta})\big)'\mathbb{E}\big(\boldsymbol{x}(y-\boldsymbol{z}\boldsymbol{\beta})\big)\big\}.$$

2.2 In this example, can $\hat{\boldsymbol{\beta}}$ be consistent for $\boldsymbol{\beta}_0$?

(The questions now get harder, give it a try but to not overly obsess!)

2.3 Under which, if any, conditions can we show that $\sqrt{n}(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}^*)\xrightarrow{d}N(0,\text{Avar})$ for Avar (more or less) as in the lecture notes??

2.4 How do your answers change if $\boldsymbol{x}$ is longer than $\boldsymbol{z}$, i.e. the model is overidentified?

Answer Key

1.1 Using LLN, Slutsky, and regularity conditions,

$$\hat{\beta}_1 \xrightarrow{p} \frac{\operatorname{cov}(y,z)}{\operatorname{var}(z)} = \frac{\operatorname{cov}(\beta_0+\beta_1 z^*+u, dz^*+(1-d)\tilde{z})}{\operatorname{var}(z)} = \frac{\beta_1 \mathbb{E}d \operatorname{var}(z)}{\operatorname{var}(z)} = \pi\beta_1,$$

because z^* and $\tilde{z}$ are not correlated.

1.2

$$\hat{\gamma}_1 \xrightarrow{p} \frac{\operatorname{cov}(y,z)}{\operatorname{var}(y)} = \frac{\beta_1\pi \operatorname{var}(z)}{\beta_1^2 \operatorname{var}(z)+\operatorname{var}(\varepsilon)}$$

and therefore

$$\hat{\gamma}_1^{-1} \xrightarrow{p} \frac{\beta_1^2 \operatorname{var}(z)+\operatorname{var}(\varepsilon)}{\beta_1\pi \operatorname{var}(z)} > \beta_1.$$

We can therefore bound $\operatorname{plim}\hat{\beta}_1 \leq \beta_1 < \operatorname{plim}\hat{\gamma}_1^{-1}$.

1.3

$$\hat{\beta}_1^{IV} \xrightarrow{p} \frac{\operatorname{cov}(x,y)}{\operatorname{cov}(x,z)} = \frac{\operatorname{cov}(x,\beta_0+\beta_1 z^*+u)}{\operatorname{cov}(x,dz^*+(1-d)\tilde{z})} = \frac{\beta_1 \operatorname{cov}(x,z^*)}{\pi \operatorname{cov}(x,z^*))} = \frac{\beta_1}{\pi}.$$

We can therefore estimate β_1 by $(\hat{\beta}_1\hat{\beta}_1^{IV})^{1/2}$ and π by $(\hat{\beta}_1/\hat{\beta}_1^{IV})^{1/2}$.

1.4 I show this for the second condition, the others are similar and of course related to algebra above.

$$\begin{aligned}\mathbb{E}\big(z(y-\beta_0-\beta_1 z)\big) &= \mathbb{E}\big(z(\beta_0+\beta_1 z^*+\varepsilon)-\beta_0 z-\pi\beta_1 z^2\big) = \mathbb{E}\big(\beta_1 zz^*+z\varepsilon-\pi\beta_1 z^2\big)\\ &= \mathbb{E}\big(\beta_1(dz^*+(1-d)\tilde{z})z^*-\pi\beta_1(z^*)^2\big) = \beta_1\mathbb{E}d\mathbb{E}(z^*)^2-\beta_1\pi\mathbb{E}(z^*)^2 = 0.\end{aligned}$$

1.5 No because the model is just identified.

1.6 This testing problem (i.e., $H_0 : \pi = 1$) can be embedded in limear GMM with a little trick: Redefine the parameters of interest to be $\rho \equiv \beta_1/\pi$ and $\delta \equiv \pi\beta_1$, then moment condsitions are linear in parameters and the null hypothesis is $H_0 : \delta = \rho$.

With hindsight, since at first glance the conditions are nonlinear, I should have given a hint to this effect! AS is, you dug deep in the lecture notes. Nice ideas that came up in several answers are a Hausman test for the same or a specification test after making the system overidentified by forcing $\pi = 1$.

2.1 From the first-order condition of the population respectively sample minimization problerms, we have

$$\begin{aligned}\hat{\boldsymbol{\beta}} &= \boldsymbol{S}_{\boldsymbol{xz}}^{-1}\boldsymbol{s}_{\boldsymbol{xy}}\\ \boldsymbol{\beta}^* &= \boldsymbol{\Sigma}_{\boldsymbol{xz}}^{-1}\boldsymbol{\sigma}_{\boldsymbol{xy}}.\end{aligned}$$

Thus $\hat{\boldsymbol{\beta}} \xrightarrow{p} \boldsymbol{\beta}^*$ by LLN and Slutsky (plus invertibility of certain matrices etc.).

2.2 No. This is not only true in the sense that consistency cannot be shown, but in the stronger sense that $\operatorname{plim} \hat{\boldsymbol{\beta}} = \boldsymbol{\beta}^* \neq \boldsymbol{\beta}_0$. Informally, this is because an invertible matrix defines a one-to-one mapping. More formally, letting $\boldsymbol{\rho} \equiv (0, 0, \ldots, 0, 1)'$, write

$$\begin{aligned} \boldsymbol{\sigma}_{\boldsymbol{xy}} - \boldsymbol{\Sigma}_{\boldsymbol{xz}} \boldsymbol{\beta}_0 &= \boldsymbol{\rho} \\ \boldsymbol{\sigma}_{\boldsymbol{xy}} - \boldsymbol{\Sigma}_{\boldsymbol{xz}} \boldsymbol{\beta}^* &= \mathbf{0} \\ \implies \boldsymbol{\Sigma}_{\boldsymbol{xz}} (\boldsymbol{\beta}^* - \boldsymbol{\beta}_0) &= \boldsymbol{\rho} \\ \implies \boldsymbol{\beta}^* - \boldsymbol{\beta}_0 &= \boldsymbol{\Sigma}_{\boldsymbol{xz}}^{-1} \boldsymbol{\rho} \neq \mathbf{0} \end{aligned}$$

because a nonzero vector premultiplied by a nonsingular matrix never yields the zero vector.

2.3 Some good attempts here. The correct answer is that asymptotic normality about $\boldsymbol{\beta}^*$ goes through but the expression for Avar does not.

2.4 Here I really only wanted to hear that the model is now overidentified and we can therefore detect misspecification, whereas before we couldn't.