Econ 6200: Econometrics II
Prelim, April 14^{th}, 2022

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

Throughout, you may invoke theorems from class without proof. Recall also that, in OLS regression of Y on scalar X and if OLS assumptions hold (not true for all instances of Y and X below!), then

$$\sqrt{n}(\hat{\beta} - \beta) \xrightarrow{d} N\left(0, \tfrac{\sigma^2}{\text{var}(X)}\right).$$

Good luck!

1 An econometrician wants to use multiple-equation linear GMM with positive definite weighting matrix $\hat{\boldsymbol{W}} \xrightarrow{p} \boldsymbol{W}$ to estimate the three equation

$$\begin{aligned} Y_1 &= \alpha_1 + X'\beta_1 + \varepsilon_1 \\ Y_2 &= \alpha_2 + X'\beta_2 + \varepsilon_2 \\ Y_3 &= \alpha_3 + X'\beta_3 + \varepsilon_3, \end{aligned}$$

where X includes no constant, $\mathbb{E}(XX')$ is invertible, and $\mathbb{E}(X\varepsilon_1) = \mathbb{E}(X\varepsilon_2) = \mathbb{E}(X\varepsilon_3) = 0$. Furthermore, for parts of this question we will assume that $\beta_3 = 0$ and use that information.

1.1 Does the specific choice of $\hat{\boldsymbol{W}}$ matter if $\beta_3 = 0$ is used? What if it is not used?

1.2 Explain how to compute an estimator of $\boldsymbol{\beta} \equiv (\beta_1', \beta_2', \beta_3')'$ that minimizes asymptotic variance among GMM estimators, assuming that the information $\beta_3 = 0$ *is* used.

1.3 Assuming that $\beta_3 = 0$ is imposed, explain how to conduct a specification test of the overall model.

1.4 Provide a test of $H_0 : \beta_3 = 0$ using only (Y_3, X_3) that uses the same asymptotic distribution (hence, critical values) as the test from 1.3. Do you think the tests are asymptotically equivalent? (A reasoned conjecture suffices.)

2 Consider the equations

$$\begin{aligned} Y &= \alpha + \beta X + \varepsilon \\ X &= \gamma + \delta Z + \eta, \end{aligned}$$

where all r.v.'s are scalars, where $\mathbb{E}\eta = \mathbb{E}\varepsilon = \mathbb{E}(Z\eta) = \mathbb{E}(Z\varepsilon) = 0$, the variances $(\sigma_z^2, \sigma_\eta^2, \sigma_\varepsilon^2)$ are nonzero, and data are i.i.d. Please use the notation $\sigma_{\eta\varepsilon} \equiv \mathrm{cov}(\eta, \varepsilon)$.

Hint: The combined expression $Y = \alpha + \beta(\gamma + \delta Z + \eta) + \varepsilon$ may be useful.

2.1 Give additional restrictions on the model's parameters under which:

- OLS of Y on X is consistent for β,
- IV of Y on X, using Z as instrument, is consistent for β.

2.2 You use OLS to regress Y on Z. Call the slope estimator from that regression $\hat{\rho}$. Characterize its plim and asymptotic distribution.

2.3 The IV estimator of β can be expressed as $\hat{\beta}_{IV} = \hat{\rho}/\hat{\delta}$, where $\hat{\delta}$ comes from OLS regression of X on Z. Verify this and use it to directly argue consistency of $\hat{\beta}_{IV}$ under conditions corresponding to the second bullet in 2.1. (Note: As sanity check or hint for 2.1, you should need just those conditions.)

3 Consider the model

$$Y = \alpha + \beta X + \varepsilon,$$

where:

- the scalars (Y, X) are i.i.d.,
- X is supported on the entire real line, and
- ε is distributed independently from X according to the Poisson distribution with parameter $\lambda = 1$ (thus $\mathbb{E}\varepsilon = \mathrm{var}\,\varepsilon = \lambda$).

A researcher attempts to estimate (α, β) by computing $(\hat{\alpha}, \hat{\beta})$, the OLS regression of Y on X.

3.1 Argue that $\hat{\beta}$ is unbiased. Is $\hat{\alpha}$ unbiased?

3.2 Is $\hat{\beta}$ the best (=minimum variance) linear unbiased estimator of β?

3.3 Is $\hat{\beta}$ the best (=minimum variance) unbiased estimator of β?

Answer Key

1 This system is just identified (and just stacks 3 OLS equations) if all parameters are free; it turns into an overidentified system (namely, SURE) when $\beta_3 = 0$ is imposed. This insight informs the following answers:

1.1 Yes; no.

1.2 Two-stage (efficient) GMM.

1.3 Overidentification ("Sargan") J-statistic.

1.4 Test jolint signicicance of all coefficients on the 3rd equation as estimated by OLS; equivalently, test for overall significance of said equation. This test is asymptotically (and even in finite sample, due to linearity of everything; but this is very involved to show) equivalent to the preceding one.

2.1 For OLS, we need $\mathbb{E}X\varepsilon = 0$, for IV, we need $\text{cov}(Z, x) \neq 0 \Leftrightarrow \delta \neq 0$.

2.2 Substitute in to find

$$Y = \alpha + \beta(\gamma + \delta Z + \eta) = \varepsilon = \alpha + \beta\gamma + \underbrace{\beta\delta}_{\equiv\rho} X + \beta\eta + \varepsilon,$$

and assumptions suffice to conlude that standard OLS asymptotics apply, therefore

$$\begin{aligned} \sqrt{n}(\hat{\rho} - \rho) &\xrightarrow{d} N(0, \boldsymbol{V}_\rho) \\ \boldsymbol{V}_\rho &= \frac{\beta^2\sigma_\eta^2 + \sigma_\varepsilon^2 + 2\beta\sigma_{\eta\varepsilon}}{\text{var}\, Z}. \end{aligned}$$

(A minor but frequent mistake was to drop the cross-product term in the asymptotic variance. Hint: If you don't use a hint, you probably missed something!)

2.3 The verification is easy and consistency then follows by Slutsky. Note that $\delta \neq 0$ is used because we must not divide by 0.

3 This is in principle just OLS with a funky error term. However, we want to be careful in that it strictly is OLS only after writing

$$Y = \underbrace{\alpha + \lambda}_{\equiv\tilde{\alpha}} + \beta X + \underbrace{\varepsilon - \lambda}_{\equiv\tilde{\varepsilon}},$$

since we need the error term to be mean zero. This rewrite will not affect the slope estimator.

3.1 In view of the above, OLS will consistently estimate $(\tilde{\alpha}, \beta)$ and therefore β but not α.

3.2 Yes, by Gauss-Markov applied to the slightly rewritten equation above.

3.3 Probably not, given the amount of structure we have here. That said, my intended answer was incorrectly reasoned, which is why everybody got full marks here.

The intended answer was that if we can find 3 perfectly collinear observations, we can conclude that they correspond to identical realizations of ε and back out β, quite like in an early homework. We cannot back out α because we do not know the value that ε took.

However, this is not quite right. We can indeed learn β superconsistently, but we also need to specify what we do if no linearly dependent observations were found. My intended answer was to stick with OLS then, but contrary to the homework, we cannot argue that this estimator is unbiased conditionally on the event on which it is computed! Indeed, I don't think it is, because a data set without collinear observations will have "unexpectedly many" realizations of rare and therefore high values of ε.