Econometrics 2: Problem set 1

Problem 1

Part 1.1

Question: Show that under assumptions $3.1 - 3.4$

$$\hat{\boldsymbol{\delta}}(\hat{\boldsymbol{W}}) \xrightarrow{p} \boldsymbol{\delta}, \quad n \to \infty$$

Proof: The GMM estimator is

$$\hat{\boldsymbol{\delta}}(\hat{\boldsymbol{W}}) = (\boldsymbol{S_{xz}}'\hat{\boldsymbol{W}}\boldsymbol{S_{xz}})^{-1}\boldsymbol{S_{xz}}'\hat{\boldsymbol{W}}\hat{\boldsymbol{s_{xy}}} \tag{1}$$

Furthermore, assumption 3.1 provides us with the equation

$$y_i = \boldsymbol{z_i}'\boldsymbol{\delta} + \epsilon_i$$

Let's multiply both sides from the left by $\boldsymbol{x_i}$, sum over i, and then multiply both sides again by $\frac{1}{n}$. We then obtain the expression for the sampling error

$$\begin{aligned} \frac{1}{n}\sum_i^n \boldsymbol{x_i} y_i &= \frac{1}{n}\sum_i^n \boldsymbol{x_i}\boldsymbol{z_i}\boldsymbol{\delta} + \frac{1}{n}\sum_i^n \boldsymbol{x_i}\epsilon_i \\ \iff \boldsymbol{s_{xy}} &= \boldsymbol{S_{xz}}\boldsymbol{\delta} + \bar{\boldsymbol{g}} \end{aligned} \tag{2}$$

Substituting this into (1), we arrive at

$$\begin{aligned} \hat{\boldsymbol{\delta}}(\hat{\boldsymbol{W}}) &= (\boldsymbol{S_{xz}}'\hat{\boldsymbol{W}}\boldsymbol{S_{xz}})^{-1}\boldsymbol{S_{xz}}'\hat{\boldsymbol{W}}\boldsymbol{s_{xy}} \\ \iff \hat{\boldsymbol{\delta}}(\hat{\boldsymbol{W}}) &= (\boldsymbol{S_{xz}}'\hat{\boldsymbol{W}}\boldsymbol{S_{xz}})^{-1}\boldsymbol{S_{xz}}'\hat{\boldsymbol{W}}[\boldsymbol{S_{xz}}\boldsymbol{\delta} + \bar{\boldsymbol{g}}] \\ \iff \hat{\boldsymbol{\delta}}(\hat{\boldsymbol{W}}) &= \boldsymbol{\delta} + (\boldsymbol{S_{xz}}'\hat{\boldsymbol{W}}\boldsymbol{S_{xz}})^{-1}\boldsymbol{S_{xz}}'\hat{\boldsymbol{W}}\bar{\boldsymbol{g}} \\ \iff \hat{\boldsymbol{\delta}}(\hat{\boldsymbol{W}}) - \boldsymbol{\delta} &= (\boldsymbol{S_{xz}}'\hat{\boldsymbol{W}}\boldsymbol{S_{xz}})^{-1}\boldsymbol{S_{xz}}'\hat{\boldsymbol{W}}\bar{\boldsymbol{g}} \end{aligned} \tag{3}$$

Now, by assumption A3.2 (ergodic stationarity),

$$\boldsymbol{S_{xz}} \xrightarrow{p} \boldsymbol{\Sigma_{xz}} \tag{4}$$

and by both assumptions A3.2 and A3.3

$$\bar{\boldsymbol{g}} \xrightarrow{p} \boldsymbol{0} \tag{5}$$

To show that the GMM estimator converges in probability to the true parameter, it is now enough to show that the RHS in (3) converges in probability to zero.

Now, because each $\boldsymbol{S_{xz}}$ converges in probability to something finite ($\boldsymbol{\Sigma_{xz}}$), and because $\hat{\boldsymbol{W}} \xrightarrow{p} \boldsymbol{W}$ by definition, their product *also* converges in probability to something finite. This, along with the fact

$$\bar{\boldsymbol{g}} \xrightarrow{p} \boldsymbol{0} \tag{6}$$

and Lemma 2.3(a) is enough to show that

$$(\boldsymbol{S_{xz}}'\hat{\boldsymbol{W}}\boldsymbol{S_{xz}})^{-1}\boldsymbol{S_{xz}}'\hat{\boldsymbol{W}}\bar{\boldsymbol{g}} \quad \xrightarrow{p} \quad \boldsymbol{0}$$

and thus

$$\begin{aligned} \hat{\boldsymbol{\delta}}(\hat{\boldsymbol{W}}) - \boldsymbol{\delta} \quad &\xrightarrow{p} \quad \boldsymbol{0} \\ \iff \hat{\boldsymbol{\delta}}(\hat{\boldsymbol{W}}) \quad &\xrightarrow{p} \quad \boldsymbol{\delta} \end{aligned}$$

Part 1.2

Question: Show that if assumption 3.3 is strengthened as assumption 3.5, then

$$\sqrt{n}[\hat{\boldsymbol{\delta}}(\hat{\boldsymbol{W}}) - \boldsymbol{\delta}] \xrightarrow{d} N(\boldsymbol{0}, Avar(\hat{\boldsymbol{\delta}}(\hat{\boldsymbol{W}}))), \quad n \to \infty$$

Proof: We have

$$\hat{\boldsymbol{\delta}}(\hat{\boldsymbol{W}}) - \boldsymbol{\delta} = (\boldsymbol{S_{xz}}'\hat{\boldsymbol{W}}\boldsymbol{S_{xz}})^{-1}\boldsymbol{S_{xz}}'\hat{\boldsymbol{W}}\bar{\boldsymbol{g}} \tag{7}$$

Let's multiply both sides by $\sqrt{n}$. We obtain

$$\sqrt{n}(\hat{\boldsymbol{\delta}}(\hat{\boldsymbol{W}}) - \boldsymbol{\delta}) = (\boldsymbol{S_{xz}}'\hat{\boldsymbol{W}}\boldsymbol{S_{xz}})^{-1}\boldsymbol{S_{xz}}'\hat{\boldsymbol{W}}\sqrt{n}\bar{\boldsymbol{g}} \tag{8}$$

Now, by assumption 3.5,

$$\sqrt{n}\bar{\boldsymbol{g}} \xrightarrow{d} N(\boldsymbol{0}, \boldsymbol{S}) \tag{9}$$

We find exactly the above term on the RHS of (8). We also know by that everything else on the RHS converges in probability to something finite ($S \xrightarrow{p} \Sigma$, $\hat{W} \xrightarrow{p} W$), and so the product also converges in probability to something finite. For a random variable whose expected value is zero, multiplying it by anything does not change the asymptotic mean, but the product changes the variance. By lemma 2.4(c), we know that variance. Thus when (9) holds, then

$$\begin{aligned} &\sqrt{n}(\hat{\boldsymbol{\delta}}(\hat{\boldsymbol{W}}) - \boldsymbol{\delta}) \\ &= (\boldsymbol{S_{xz}}'\hat{\boldsymbol{W}}\boldsymbol{S_{xz}})^{-1}\boldsymbol{S_{xz}}'\hat{\boldsymbol{W}}\sqrt{n}\bar{\boldsymbol{g}} \\ &\xrightarrow{d} \\ &N(0, (\boldsymbol{\Sigma_{xz}}'\boldsymbol{W}\boldsymbol{S_{xz}})^{-1}\boldsymbol{\Sigma_{xz}}'\boldsymbol{W}\boldsymbol{S}\boldsymbol{W}'\boldsymbol{\Sigma_{xz}}(\boldsymbol{\Sigma_{xz}}'\boldsymbol{W}\boldsymbol{S_{xz}})^{-1}) \end{aligned} \tag{10}$$

where the latter term is, by definition, $N(\boldsymbol{0}, Avar(\hat{\boldsymbol{\delta}}(\hat{\boldsymbol{W}})))$.

Part 1.3

Question: Suppose there is available a consistent estimator, $\hat{\boldsymbol{S}}$, of $\boldsymbol{S}$. Show that then, under assumption 3.2, $Avar(\hat{\delta}(\hat{W}))$ is consistently estimated by

$$\widehat{Avar(\hat{\boldsymbol{\delta}}(\hat{\boldsymbol{W}}))} = (\boldsymbol{S_{xx}}'\hat{\boldsymbol{W}}\boldsymbol{S_{xz}})^{-1}\boldsymbol{S_{xz}}'\hat{\boldsymbol{W}}\hat{\boldsymbol{S}}\hat{\boldsymbol{W}}\boldsymbol{S_{xz}}(\boldsymbol{S_{xz}}'\hat{\boldsymbol{W}}\boldsymbol{S_{xz}})^{-1}$$

Proof: This one is rather straightforward. Because

$$\begin{aligned} \boldsymbol{S_{xz}} &\xrightarrow{p} \boldsymbol{\Sigma_{xz}} \\ \hat{\boldsymbol{W}} &\xrightarrow{p} \boldsymbol{W} \\ \hat{\boldsymbol{S}} &\xrightarrow{p} \boldsymbol{S} \end{aligned} \tag{11}$$

then by Lemma 2.3(a), the product of each term in the LHS will converge to the product of the corresponding terms on the RHS. Thus

$$\begin{aligned} &\widehat{Avar(\hat{\boldsymbol{\delta}}(\hat{\boldsymbol{W}}))} \\ &= (\boldsymbol{S_{xx}}'\hat{\boldsymbol{W}}\boldsymbol{S_{xz}})^{-1}\boldsymbol{S_{xz}}'\hat{\boldsymbol{W}}\hat{\boldsymbol{S}}\hat{\boldsymbol{W}}\boldsymbol{S_{xz}}(\boldsymbol{S_{xz}}'\hat{\boldsymbol{W}}\boldsymbol{S_{xz}})^{-1} \\ &\xrightarrow{p} \\ &(\boldsymbol{\Sigma_{xz}}'\boldsymbol{W}\boldsymbol{S_{xz}})^{-1}\boldsymbol{\Sigma_{xz}}'\boldsymbol{W}\boldsymbol{S}\boldsymbol{W}'\boldsymbol{\Sigma_{xz}}(\boldsymbol{\Sigma_{xz}}'\boldsymbol{W}\boldsymbol{S_{xz}})^{-1} \\ &= Avar(\hat{\boldsymbol{\delta}}(\hat{\boldsymbol{W}})) \end{aligned} \tag{12}$$

Problem 2

Question: Suppose $\hat{\boldsymbol{W}}_{\boldsymbol{1}} - \hat{\boldsymbol{W}}_{\boldsymbol{2}} \xrightarrow{p} \boldsymbol{0}$. Show that

$$\sqrt{n}\hat{\boldsymbol{\delta}}(\hat{\boldsymbol{W}}_{\boldsymbol{1}}) - \sqrt{n}(\hat{\boldsymbol{W}}_{\boldsymbol{2}}) \xrightarrow{p} \boldsymbol{0}$$

Solution: From (3) we have

$$\hat{\boldsymbol{\delta}}(\hat{\boldsymbol{W}}) = \boldsymbol{\delta} + (\boldsymbol{S_{xz}}'\hat{\boldsymbol{W}}\boldsymbol{S_{xz}})^{-1}\boldsymbol{S_{xz}}'\hat{\boldsymbol{W}}\bar{\boldsymbol{g}} \tag{13}$$

and thus with inputs, $\boldsymbol{W_1}, \boldsymbol{W_2}$ the difference of the scaled estimators will appear as

$$\begin{aligned} &\sqrt{n}\hat{\boldsymbol{\delta}}(\hat{\boldsymbol{W}}_{\boldsymbol{1}}) - \sqrt{n}(\hat{\boldsymbol{W}}_{\boldsymbol{2}}) \\ &= \sqrt{n}\boldsymbol{\delta} + \sqrt{n}(\boldsymbol{S_{xz}}'\hat{\boldsymbol{W}}_{\boldsymbol{1}}\boldsymbol{S_{xz}})^{-1}\boldsymbol{S_{xz}}'\hat{\boldsymbol{W}}_{\boldsymbol{1}}\bar{\boldsymbol{g}} - \sqrt{n}\boldsymbol{\delta} + \sqrt{n}(\boldsymbol{S_{xz}}'\hat{\boldsymbol{W}}_{\boldsymbol{2}}\boldsymbol{S_{xz}})^{-1}\boldsymbol{S_{xz}}'\hat{\boldsymbol{W}}_{\boldsymbol{2}}\bar{\boldsymbol{g}} \\ &= \sqrt{n}(\boldsymbol{S_{xz}}'\hat{\boldsymbol{W}}_{\boldsymbol{1}}\boldsymbol{S_{xz}})^{-1}\boldsymbol{S_{xz}}'\hat{\boldsymbol{W}}_{\boldsymbol{1}}\bar{\boldsymbol{g}} - \sqrt{n}(\boldsymbol{S_{xz}}'\hat{\boldsymbol{W}}_{\boldsymbol{2}}\boldsymbol{S_{xz}})^{-1}\boldsymbol{S_{xz}}'\hat{\boldsymbol{W}}_{\boldsymbol{2}}\bar{\boldsymbol{g}} \\ &= \left[(\boldsymbol{S_{xz}}'\hat{\boldsymbol{W}}_{\boldsymbol{1}}\boldsymbol{S_{xz}})^{-1}\boldsymbol{S_{xz}}'\hat{\boldsymbol{W}}_{\boldsymbol{1}} - (\boldsymbol{S_{xz}}'\hat{\boldsymbol{W}}_{\boldsymbol{2}}\boldsymbol{S_{xz}})^{-1}\boldsymbol{S_{xz}}'\hat{\boldsymbol{W}}_{\boldsymbol{2}}\right]\sqrt{n}\bar{\boldsymbol{g}} \end{aligned} \tag{14}$$

By assumption, $\hat{\boldsymbol{W}}_{\boldsymbol{1}} - \hat{\boldsymbol{W}}_{\boldsymbol{2}} \xrightarrow{p} \boldsymbol{0}$, and by Lemma 2.3(a) also the expression inside the square brackers will converge in probability to zero, because it is a continuous transformation of the two weighting matrices. As before, by assumption 3.5,

$$\sqrt{n}\bar{\boldsymbol{g}} \quad \xrightarrow{d} \quad N(\mathbf{0}, \boldsymbol{S}) \tag{15}$$

and since

$$\left[(\boldsymbol{S_{xz}}'\hat{\boldsymbol{W}}_{\mathbf{1}}\boldsymbol{S_{xz}})^{-1}\boldsymbol{S_{xz}}'\hat{\boldsymbol{W}}_{\mathbf{1}} - (\boldsymbol{S_{xz}}'\hat{\boldsymbol{W}}_{\mathbf{2}}\boldsymbol{S_{xz}})^{-1}\boldsymbol{S_{xz}}'\hat{\boldsymbol{W}}_{\mathbf{2}}\right] \quad \xrightarrow{p} \quad \mathbf{0} \tag{16}$$

then by Lemma 2.4(b),

$$\begin{aligned}
&\sqrt{n}\hat{\boldsymbol{\delta}}(\hat{\boldsymbol{W}}_{\mathbf{1}}) - \sqrt{n}(\hat{\boldsymbol{W}}_{\mathbf{2}}) \\
&= \left[(\boldsymbol{S_{xz}}'\hat{\boldsymbol{W}}_{\mathbf{1}}\boldsymbol{S_{xz}})^{-1}\boldsymbol{S_{xz}}'\hat{\boldsymbol{W}}_{\mathbf{1}} - (\boldsymbol{S_{xz}}'\hat{\boldsymbol{W}}_{\mathbf{2}}\boldsymbol{S_{xz}})^{-1}\boldsymbol{S_{xz}}'\hat{\boldsymbol{W}}_{\mathbf{2}}\right]\sqrt{n}\bar{\boldsymbol{g}} \\
&\xrightarrow{p} \mathbf{0}
\end{aligned} \tag{17}$$

which concludes the proof for the second problem.

Problem 3

Question: Suppose assumptions 3.1 – 3.5 hold, and suppose there is available a consistent estimate $\hat{S}$ of S. Let

$$\widehat{Avar(\hat{\boldsymbol{\delta}}(\hat{\boldsymbol{W}}))} = (\boldsymbol{S_{xx}}'\hat{\boldsymbol{W}}\boldsymbol{S_{xz}})^{-1}\boldsymbol{S_{xz}}'\hat{\boldsymbol{W}}\hat{\boldsymbol{S}}\hat{\boldsymbol{W}}\boldsymbol{S_{xz}}(\boldsymbol{S_{xz}}'\hat{\boldsymbol{W}}\boldsymbol{S_{xz}})^{-1}$$

Then, prove that under the null hypothesis $H_0 : \delta_l = \bar{\delta}_l$,

$$t_l = \frac{\sqrt{n}(\hat{\delta}_l(\hat{\boldsymbol{W}}) - \bar{\delta}_l)}{\sqrt{\widehat{Avar(\hat{\boldsymbol{\delta}}(\hat{\boldsymbol{W}}))}_{ll}}} = \frac{\hat{\delta}_l(\hat{\boldsymbol{W}}) - \bar{\delta}_l}{SE_l^*} \xrightarrow{d} N(0,1) \tag{18}$$

Solution: This one is also rather straightforward. In **part 1.2** we showed that

$$\sqrt{n}[\hat{\boldsymbol{\delta}}(\hat{\boldsymbol{W}}) - \boldsymbol{\delta}] \quad \xrightarrow{d} \quad N(\mathbf{0}, Avar(\hat{\boldsymbol{\delta}}(\hat{\boldsymbol{W}}))) \tag{19}$$

Because of the assumption provided by the null hypothesis, the term on the LHS is of same form as the numerator of (18). In **part 1.3** we on the other hand showed that

$$\widehat{Avar(\hat{\boldsymbol{\delta}}(\hat{\boldsymbol{W}}))} \quad \xrightarrow{p} \quad Avar(\hat{\boldsymbol{\delta}}(\hat{\boldsymbol{W}})) \tag{20}$$

Now, variance has the property

$$\begin{aligned}
\mathbb{V}[aX] &= \mathbb{E}[(aX - \mathbb{E}[aX])(aX - \mathbb{E}[aX])] \\
&= \mathbb{E}[(aX - a\,\mathbb{E}[X])(aX - a\,\mathbb{E}[X])] \\
&= \mathbb{E}[a(X - \mathbb{E}[X])a(X - \mathbb{E}[X])] \\
&= \mathbb{E}[a^2(X - \mathbb{E}[X])(X - \mathbb{E}[X])] \\
&= a^2\,\mathbb{E}[(X - \mathbb{E}[X])(X - \mathbb{E}[X])] \\
&= a^2\,\mathbb{V}[X]
\end{aligned}$$

This means that we can manipulate the variance of random variables as we please by scaling them with appropriate magnitudes. Assume, for example, that$X \sim N(0, \sigma^2)$. Then

$$\mathbb{V}\, X = \mathbb{E}[(X - \mathbb{E}\, X)^2] = \sigma^2$$

And if we now wish the variance to equal 1, we can work backwards to find a scaling magnitude that achieves that

$$\begin{aligned}
1 &= \frac{\sigma^2}{\sigma^2} \\
&= \frac{1}{\sigma^2}\sigma^2 \\
&= \frac{1}{\sigma^2}\, \mathbb{E}[(X - \mathbb{E}\, X)^2] \\
&= \mathbb{E}[(\frac{1}{\sigma^2})(X - \mathbb{E}\, X)^2] \\
&= \mathbb{E}[(\frac{1}{\sigma})(X - \mathbb{E}\, X)(\frac{1}{\sigma})(X - \mathbb{E}\, X)] \\
&= \mathbb{E}[(\frac{1}{\sigma}X - \frac{1}{\sigma}\,\mathbb{E}\, X)(\frac{1}{\sigma}X - \frac{1}{\sigma}\,\mathbb{E}\, X)] \\
&= \mathbb{E}[(\frac{1}{\sigma}X - \mathbb{E}\,\frac{1}{\sigma}X)(\frac{1}{\sigma}X - \mathbb{E}\,\frac{1}{\sigma}X)] \\
&= \mathbb{V}[\frac{1}{\sigma}X]
\end{aligned}$$

This means that since in (19) converges to a normally distributed random variable with variance $Avar(\hat{\boldsymbol{\delta}}(\hat{\boldsymbol{W}}))$, we can obtain a standard normal random variable by scaling the LHS of (19) by the reciprocal of the square root of $Avar(\hat{\boldsymbol{\delta}}(\hat{\boldsymbol{W}}))$. By lemma 2.3(a) this is possible. Thus we have

$$\sqrt{n}[\hat{\boldsymbol{\delta}}(\hat{\boldsymbol{W}}) - \boldsymbol{\delta}] \quad \xrightarrow{d} \quad N(\mathbf{0}, Avar(\hat{\boldsymbol{\delta}}(\hat{\boldsymbol{W}}))) \tag{21}$$

and as a consequence of (20), the properties of variance, and lemma 2.4(c),

$$\frac{\sqrt{n}[\hat{\delta}_l(\hat{\boldsymbol{W}}) - \bar{\delta}_l]}{\sqrt{\widehat{Avar(\hat{\boldsymbol{\delta}}(\hat{\boldsymbol{W}}))}_{ll}}} \quad \xrightarrow{d} \quad N(0, 1) \tag{22}$$

Which concludes the proof of the third part.