Econometrics 710
Final Exam
Spring 2003
Answer Key

1. The parameter is the $m \times 1$ θ, so the model is overidentified. The efficient estimator is efficient GMM. We can rewrite the equation as $y_i = z_i'\theta + e_i$ where $z_i = Q'x_i$, and then we know that the two-step efficient GMM estimator takes the form

$$\begin{aligned}\hat{\theta} &= \left(Z'X\hat{\Omega}^{-1}X'Z\right)^{-1}\left(Z'X\hat{\Omega}^{-1}X'Y\right) \\ &= \left(Q'X'X\hat{\Omega}^{-1}X'XQ\right)^{-1}\left(Q'X'X\hat{\Omega}^{-1}X'Y\right)\end{aligned}$$

where

$$\begin{aligned}\hat{\Omega} &= \frac{1}{n}\sum_{i=1}^{n} x_i x_i' \hat{e}_i^2 \\ \hat{e}_i &= y_i - x_i'\left(X'X\right)^{-1}X'Y.\end{aligned}$$

2. The efficient estimator for β is LS, $\hat{\beta} = \left(X'X\right)^{-1}X'Y$ and that for θ is $\hat{\theta} = \hat{\beta}_1/\hat{\beta}_2$. The natural test statistic for hypotheses on θ is the t-ratio $T_n = \left(\hat{\theta} - \theta\right)/s\left(\hat{\theta}\right)$ where

$$\begin{aligned}s\left(\hat{\theta}\right) &= \sqrt{H_{\hat{\beta}}'\hat{V}H_{\hat{\beta}}} \\ \hat{V} &= \left(X'X\right)^{-1}\left(\sum_{i=1}^{n} x_i x_i'\hat{e}_i^2\right)\left(X'X\right)^{-1} \\ H_{\beta} &= \frac{\partial}{\partial\beta}\left(\frac{\beta_1}{\beta_2}\right) = \begin{pmatrix} \frac{1}{\beta_2} \\ \\ -\frac{\beta_1}{\beta_2^2}\end{pmatrix}.\end{aligned}$$

The percentile-t interval is constructed using the quantiles of T_n^*, the bootstrap version of T_n. It is constructed as follows. We simulate samples of size n by sampling iid (with replacement) from the observed sample $\{y_i, x_i : i = 1, ..., n\}$. On each sample we compute $\hat{\beta}^*$, $\hat{\theta}^*$ and $s\left(\hat{\theta}^*\right)$ just as done on the original sample. Then we set $T_n^* = \left(\hat{\theta}^* - \hat{\theta}\right)/s\left(\hat{\theta}^*\right)$. We store this simulated statistic T_n^* and repeat B times. We compute the $\alpha/2$ and $1-\alpha/2$ quantiles $q^*\left(\alpha/2\right)$ and $q^*\left(1-\alpha/2\right)$ (by sorting the statistics T_n^*). The percentile-t interval is

$$\left[\theta - s\left(\hat{\theta}\right)q^*\left(1-\alpha/2\right), \quad \theta - s\left(\hat{\theta}\right)q^*\left(\alpha/2\right)\right]$$

3. $Var\left(y^*\right) = n^{-1}\sum_{i=1}^{n}\left(y_i - \overline{y}\right)^2$, the sample variance.

4.

(a) For the μ_i, there are n parameters and n moment conditions. This is a just-identified linear model, so the efficient GMM estimator is least squares. The regressors are dummy variables (indicating the individual), so the LS estimates are the sample means for the individuals. Thus $\hat{\mu}_i = \overline{y}_i$. (This can be shown directly by solving the n moment conditions. (However, it is important to be clear about the notation and definitions. Many exam answers were incorrectly summing over i rather than t, or not defining the argument of the summation, etc.)

(b) In addition to the n moment conditions for the μ_i, we have the moment condition $E\left((y_{it} - \mu_i)^2 - \sigma^2\right) = 0$, which holds for all nT observations $i = 1, ..., n$ and $t = 1, ..., T$. We have added one moment condition and one parameter, so we remain just-identified. The GMM estimator for σ^2 solves the sample moment

$$\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\left((y_{it} - \hat{\mu}_i)^2 - \hat{\sigma}^2\right) = 0$$

so that

$$\hat{\sigma}^2 = \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}(y_{it} - \hat{\mu}_i)^2$$

Note that the average is taken over all observations.

(c) Note that $\hat{\mu}_i = \overline{y}_i = \mu_i + \overline{e}_i$. The random variable $\overline{e}_i$ is independent across i and $E\overline{e}_i^2 = \sigma^2/T$. Thus

$$\begin{aligned}
\hat{\sigma}^2 &= \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}(y_{it} - \hat{\mu}_i)^2 \\
&= \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}(e_{it} - \overline{e}_i)^2 \\
&= \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}e_{it}^2 - \frac{1}{n}\sum_{i=1}^{n}\overline{e}_i^2 \\
&\rightarrow_p \sigma^2 - \frac{\sigma^2}{T} = \left(1 - \frac{1}{T}\right)\sigma^2
\end{aligned}$$

(d) As $n \to \infty$, $\hat{\sigma}^2$ is inconsistent for σ^2.
(If n and T go to infinity together, it is possible to show that $\hat{\sigma}^2$ is consistent, but this is actually a more difficult demonstration.)

5. Under general dependence conditions, if the observations are stationary, then

$$\sqrt{T}(\hat{\rho} - \rho) = \frac{\frac{1}{\sqrt{T}}\sum_{t=1}^{T} y_{t-1}e_t}{\frac{1}{T}\sum_{t=1}^{T} y_{t-1}^2} \rightarrow_d \frac{N(0, \Omega)}{E\left(y_{t-1}^2\right)} = N\left(0, \frac{\Omega}{\left(E\left(y_{t-1}^2\right)\right)^2}\right) \tag{1}$$

where

$$\Omega = \lim_{T\to\infty} E\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T} y_{t-1}e_t\right)^2 \tag{2}$$
$$= \sum_{k=-\infty}^{\infty} E\left(y_{t-1+k}e_{t+k}y_{t-1}e_t\right).$$

When e_t is a MDS, then so is $y_{t-1}e_t$, so $\Omega = E\left(y_{t-1}e_t^2\right)$. The reason why the MDS assumption is critical is that it ensures that the variable $y_{t-1}e_t$ is serially uncorrelated so that this simplification occurs. Instead, if the error e_t is, say, uncorrelated, then it is not necessarily the case the $y_{t-1}e_t$ is uncorrelated, so this simplication is not guaranteed.

In particular, under the minimal assumption $E\left(y_{t-1}e_t\right) = 0$, e_t is not even necessarily uncorrelated, and is certainly not necessarily a MDS. So the simplification $\Omega = E\left(y_{t-1}e_t^2\right)$ cannot be guaranteed. Instead, we have the limit distribution (1) above where Ω is defined in (2).

6.

(a) We typically say that z_i is "endogenous" for β since it is correlated with e_i. This statement only makes sense if β is defined structurally.

(b) No, it is inconsistent, since $\hat{\beta} \to_p \beta + E\left(z_i e_i\right)/E\left(z_i^2\right) \neq \beta$.

(c) Since

$$\tilde{\beta} = \frac{\sum_{i=1}^{n} y_i}{\sum_{i=1}^{n} z_i} = \beta + \frac{\frac{1}{n}\sum_{i=1}^{n} e_i}{\frac{1}{n}\sum_{i=1}^{n} z_i},$$

we can apply the WLLN to show that $\tilde{\beta} \to_p \beta$ when

$$E\left(e_i\right) = 0 \tag{3}$$

and

$$E\left(z_i\right) \neq 0. \tag{4}$$

As (3) is standard, the key extra condition is (4).

(d) Let $d_i = 1$. Then $\tilde{\beta}$ is an IV estimator with instrument d_i for z_i. The needed conditions for consistent estimation using IV are the orthogonality condition $E\left(d_i e_i\right) = 0$, which is (3), and the relevance condition $E\left(d_i z_i\right) \neq 0$, which is (4). The identifying restriction is the exclusion restriction – that the instrument can be excluded from the structural equation. In the present case, this means that it is valid to exclude the intercept d_i from the structural equation. That is, identification using this IV estimator rests on the exclusion of an intercept from an equation of interest – a very dubious identifying restriciton!