Econometrics 710
Final Exam
Spring, 2008
Sample Answers

1. The question was not specific regarding the dimensions of z_i and x_i. Therefore you should presume that the model could be overidentified, which includes just-identified as a special case, so it is sufficient to focus on the overidentified case. The assumptions are minimal, with the only restriction $E(x_i e_i) = 0$. There is no conditional moment restriction. Therefore the appropriate estimation method is GMM. In particular, homoskedasticity is NOT assumed. (The equation $Ee_i^2 = \sigma^2$ is not a restriction and does not imply homoskedasticity as it is an unconditional moment.) The efficient GMM estimator for β is

$$\hat{\beta} = \left(Z'X\hat{\Omega}^{-1}X'Z\right)^{-1} Z'X\hat{\Omega}^{-1}X'Y$$

where

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

and $\tilde{\beta}$ is a preliminary consistent estimator for β. OLS is inconsistent under the assumptions, but 2SLS is feasible and consistent:

$$\tilde{\beta} = \left(Z'X\left(X'X\right)^{-1}X'Z\right)^{-1} Z'X\left(X'X\right)^{-1}X'Y$$

Given $\hat{\beta}$ a moment estimator for σ^2 is

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^n \hat{e}_i^2.$$

where

$$\hat{e}_i = y_i - z_i'\hat{\beta}$$

are residuals from the GMM estimation.

As an alternative, (β, σ^2) could be jointly estimated. The joint moment equations are

$$g_i(\beta, \sigma^2) = \begin{pmatrix} x_i\left(y_i - z_i'\beta\right) \\ \left(y_i - z_i'\beta\right)^2 - \sigma^2 \end{pmatrix}.$$

Setting

$$\bar{g}_n(\beta, \sigma^2) = \frac{1}{n}\sum_{i=1}^n g_i(\beta, \sigma^2) = \begin{pmatrix} \frac{1}{n}\left(X'y - X'Z\beta\right) \\ \frac{1}{n}\left(y - Z\beta\right)'\left(y - Z\beta\right) - \sigma^2 \end{pmatrix},$$

the GMM criterion is

$$J_n(\beta, \sigma^2) = n\bar{g}_n(\beta, \sigma^2)'\hat{\Omega}^{*-1}\bar{g}_n(\beta, \sigma^2)$$

and

$$\begin{aligned} \hat{\Omega}^* &= \frac{1}{n}\sum_{i=1}^n \tilde{g}_i\tilde{g}_i' - \left(\frac{1}{n}\sum_{i=1}^n \tilde{g}_i\right)\left(\frac{1}{n}\sum_{i=1}^n \tilde{g}_i\right)' \\ \tilde{g}_i &= g_i(\tilde{\beta}, \tilde{\sigma}^2) \end{aligned}$$

with $(\tilde{\beta}, \tilde{\sigma}^2)$ consistent estimates, for example the 2SLS or GMM estimator for β described above, and a similar estimator for σ^2. The joint estimator $(\hat{\beta}, \hat{\sigma}^2)$ minimizes $J_n(\beta, \sigma^2)$. An explicit solution is not available.

2. The question asks if the estimate is statistically different than zero. This is asking for a statistical test. (In contrast, model selection asks which model fits better). The specified null hypothesis is that $\beta_2 = 0$. The general Wald statistic would be appropriate, but cannot be calculated from the information. However, the Wald statistic assuming homoskedasticity can be calculated. Thus the key assumption required is that the error is conditionally homoskedastic: $E\left(e_i^2 \mid x_i\right) = \sigma^2$, a constant. The Wald statistic for (1) versus (2) assuming homoskedasticity is

$$
\begin{aligned}
W &= n\frac{\left(\tilde{\sigma}^2 - \hat{\sigma}^2\right)}{\hat{\sigma}^2} \\
&= n\left(\frac{\sum \tilde{e}_i^2 - \sum \hat{e}_i^2}{\sum \hat{e}_i^2}\right)\frac{n-k}{k_2} \\
&= 50\left(\frac{106-100}{100}\right) \\
&= 3
\end{aligned}
$$

(k_2 is the dimension of x_{2i}, which is 3 since (2) has 3 more coefficients than (1).) A 5% asymptotic Wald test compares this with the 5% critical value of the χ_3^2 distribution, which is about 7.8. Since 3 is less than 7.8, you don't reject. While you may not have memorized this, the mean of χ_3^2 is 3, so the observed value of 3 is certainly less than the 5% quantile. Alternatively, the 5% critical value of the χ_1^2 is $1.96^2 = 3.86$, which must be smaller than the critical value of the χ_3^2 distribution, so it is easy to conclude that the observed value of 3 is smaller than the critical value.

As alternative to the Wald statistic, you could compute the F statistic, and reject the hypothesis if the F statistic exceeds the 5% critical value from the F distribution with degrees of freedom (3,42). This is appropriate if you add the additional assumption that the error is independent of the regressors and Gaussian.

3. No. This is the problem of censoring. In this example, the dependent variable is capped from above. It is identical to the problem of censoring at zero. The y_i variables which equal τ are too small, in that they are less than "correct" value of y_i^*. This biases OLS estimation, which is inconsistent for β. Drawing a picture of the impact of censoring helps, but I won't do that here in the typed answers.

   OLS estimates the linear projection of y_i on x_i, which is not β.

   A rigorous derivation of the projection coefficient is difficult, and I didn't expect anyone to find its value. Some of you tried, with varying degrees of success. Here is the calculation. Write

$$
\begin{aligned}
y_i &= y_i^* + (\tau - y_i^*)\,1\,(y_i^* > \tau) \\
&= y_i^* + \left(\tau - x_i'\beta - e_i\right)1\left(e_i > \tau - x_i'\beta\right)
\end{aligned}
$$

   Let $\hat{\beta} = (X'X)^{-1}X'Y$ and $\hat{\beta}^* = (X'X)^{-1}X'Y^*$. Then

$$
\begin{aligned}
\hat{\beta} &= \hat{\beta}^* + \hat{\delta} \\
\hat{\delta} &= \left(\frac{1}{n}X'X\right)^{-1}\left(\frac{1}{n}\sum_i x_i\left(\tau - x_i'\beta - e_i\right)1\left(e_i > \tau - x_i'\beta\right)\right).
\end{aligned}
$$

   Since $\hat{\beta}^* \to_p \beta$, we examine the plim of $\hat{\delta}$. By the WLLN, $\frac{1}{n}X'X \to_p Q = Ex_ix_i'$ and

$$
\frac{1}{n}\sum_i x_i\left(\tau - x_i'\beta - e_i\right)1\left(e_i > \tau - x_i'\beta\right) \to_p E\left(x_i\left(\tau - x_i'\beta - e_i\right)1\left(e_i > \tau - x_i'\beta\right)\right)
$$

   To make further progress, lets assume that e_i is independent of x_i and distributed $N(0,\sigma^2)$. Then this expectation is

$$\begin{aligned}
& E\left(x_i\left(\tau - x_i'\beta - e_i\right)1\left(e_i > \tau - x_i'\beta\right)\right) \\
= \ & E\left(x_i\left(\tau - x_i'\beta\right)E\left(1\left(e_i > \tau - x_i'\beta\right) \mid x_i\right)\right) - E\left(x_i E\left(e_i 1\left(e_i > \tau - x_i'\beta\right) \mid x_i\right)\right) \\
= \ & E\left(x_i\left(\tau - x_i'\beta\right)\Phi\left(\frac{x_i'\beta - \tau}{\sigma}\right)\right) - \frac{1}{\sigma}E\left(x_i\lambda\left(\frac{x_i'\beta - \tau}{\sigma}\right)\right)
\end{aligned}$$

where Φ is the normal CDF and λ is the inverse Mill's ratio. Together, we have shown that

$$\hat{\beta} \rightarrow_p \beta + Q^{-1}\left(E\left(x_i\left(\tau - x_i'\beta\right)\Phi\left(\frac{x_i'\beta - \tau}{\sigma}\right)\right) - \frac{1}{\sigma}E\left(x_i\lambda\left(\frac{x_i'\beta - \tau}{\sigma}\right)\right)\right)$$

4.

(a) Using matrix notation

$$\hat{\beta} = \left(X'WX\right)^{-1}X'WY$$

where $W = \operatorname{diag}\left(w_1, ..., w_n\right)$.Alternatively

$$\hat{\beta} = \left(\sum_{i=1}^{n} w_i x_i x_i'\right)^{-1} \sum_{i=1}^{n} w_i x_i y_i$$

(b) It appears to be estimating

$$\beta = \left(E\left(w_i x_i x_i'\right)\right)^{-1} E\left(w_i x_i y_i\right),$$

a weighted projection. This imposes no assumptions beyond the existence of moments and the invertibility of $E\left(w_i x_i x_i'\right)$.

As an alternative, you might state that $\hat{\beta}$ is estimating the slope parameter in the regression model

$$\begin{aligned}
y_i &= x_i'\beta + e_i \\
E\left(e_i \mid x_i, w_i\right) &= 0
\end{aligned}$$

but this is more restrictive than the simple weighted projection. As another alternative, you might state that you are estimating the model

$$\begin{aligned}
y_i &= x_i'\beta + e_i \\
E\left(e_i \mid x_i\right) &= 0 \\
w_i &= \sigma_i^{-2} \\
\sigma_i^2 &= E\left(e_i^2 \mid x_i\right)
\end{aligned}$$

but this is the most restrictive and narrow assumption.

(c) By the WLLN, $\frac{1}{n}\sum_{i=1}^{n} w_i x_i x_i' \rightarrow_p E\left(w_i x_i x_i'\right)$ and $\frac{1}{n}\sum_{i=1}^{n} w_i x_i y_i \rightarrow_p E\left(w_i x_i y_i\right)$. By the CMT,

$$\hat{\beta} = \left(\frac{1}{n}\sum_{i=1}^{n} w_i x_i x_i'\right)^{-1} \frac{1}{n}\sum_{i=1}^{n} w_i x_i y_i \rightarrow_p \beta$$

as defined in part (b).

(d) With β defined in part (b), we then **define** the error by the equation

$$y_i = x_i'\beta + e_i$$

It is important that e_i be defined, as it is not given in the question! It is also important that the definition be consistent with your answer in part (b). Then

$$\begin{aligned}\hat{\beta} &= \left(\sum_{i=1}^{n} w_i x_i x_i'\right)^{-1} \sum_{i=1}^{n} w_i x_i y_i \\ &= \beta + \left(\sum_{i=1}^{n} w_i x_i x_i'\right)^{-1} \sum_{i=1}^{n} w_i x_i e_i\end{aligned}$$

so

$$\sqrt{n}\left(\hat{\beta} - \beta\right) = \left(\frac{1}{n}\sum_{i=1}^{n} w_i x_i x_i'\right)^{-1} \frac{1}{\sqrt{n}}\sum_{i=1}^{n} w_i x_i e_i$$

As $n \to \infty$

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n} w_i x_i e_i \to_d N(0, \Omega)$$

where

$$\Omega = E\left(w_i^2 x_i x_i' e_i^2\right)$$

Then letting $Q = E(w_i x_i x_i')$,

$$\sqrt{n}\left(\hat{\beta} - \beta\right) \to_d N\left(0, Q^{-1}\Omega Q^{-1}\right)$$

5. There is a major error in reasoning. It is in the construction of the bootstrap t-ratio. He uses $t_n^* = \hat{\alpha}^*/s(\hat{\alpha}^*)$.The correct bootstrap t-ratio is $t_n^* = (\hat{\alpha}^* - \hat{\alpha})/s(\hat{\alpha}^*)$. What is of interest is the 95% quantile of $(\hat{\alpha}^* - \hat{\alpha})/s(\hat{\alpha}^*)$, not $\hat{\alpha}^*/s(\hat{\alpha}^*)$. The reason why the student's estimated quantile $\hat{q}^*_{.95}$ is so high is very likely because of this mis-centering. There are no other errors in his method.