Econometrics I: Final Examination Suggested Answer

Daiji Kawaguchi

March 2 2004

1 Problem 1: Basic Questions

1-1. The additional assumption needed for the consistency is $Ex'u = 0$.

1-2. Multiplying x' from the front of (1) and taking expectation, we obtain,

$$Ex'y = Ex'x\beta + Ex'u,$$

but $Ex'u = 0$ from the assumption given in 1-1. $Ex'x$ is invertible from the assumption $\text{rank}\{Ex'x\} = k$. Thus

$$\beta = [Ex'x]^{-1}Ex'y$$

in the population. Applying the analogy principle to this population moment, we obtain,

$$\hat{\beta} = (\sum_{i=1}^{N} x_i'x_i)^{-1}(\sum_{i=1}^{N} x_i'y_i).$$

1-3.

$$\hat{\beta} = (\sum_{i=1}^{N} x_i'x_i)^{-1}(\sum_{i=1}^{N} x_i'y_i)$$

$$
\begin{aligned}
&= (\sum_{i=1}^{N} x_i' x_i)^{-1} (\sum_{i=1}^{N} x_i'(x_i\beta + u_i)) \\
&= \beta + (\sum_{i=1}^{N} x_i' x_i)^{-1} (\sum_{i=1}^{N} x_i' u_i) \\
&\to \beta + (Ex'x)^{-1}(Ex'u) = \beta \qquad (1)
\end{aligned}
$$

1-4.

$$
\sqrt{n}(\hat{\beta} - \beta) = (\frac{1}{N}\sum_{i=1}^{N} x_i' x_i)^{-1} (\frac{1}{\sqrt{N}}\sum_{i=1}^{N} x_i' u_i) \qquad (2)
$$

The central limit theorem assures that

$$
(\frac{1}{\sqrt{N}}\sum_{i=1}^{N} x_i' u_i) \sim^{asy} N(0, E(u^2 x' x)).
$$

From this result,

$$
(\sqrt{n}(\hat{\beta} - \beta) \sim^{asy} N(0, (Ex'x)^{-1} E(u^2 x' x)(Ex'x)^{-1}).
$$

Without the homoskedasticity assumption, we cannot simplify this formula any further.

1-5. This claim is true for the small sample. The normality of OLS estimator depends on the normality of u. However, in the large sample, with help of the central limit theorem, we can show the asymptotic normality of the OLS estimator without relying on the normality assumption of u as shown in 1-4.

1-6.

$$
E(\hat{\beta}|x) = E(\beta + (\sum_{i=1}^{N} x_i' x_i)^{-1} \sum_{i=1}^{N} x_i' u_i | x)
$$

$$
\begin{aligned}
&= \beta + E((\sum_{i=1}^{N} x_i' x_i)^{-1} \sum_{i=1}^{N} x_i' u_i | x) \\
&= \beta + (\sum_{i=1}^{N} x_i' x_i)^{-1} \sum_{i=1}^{N} x_i' E(u_i | x) \\
&= \beta \qquad (3)
\end{aligned}
$$

1-7. The unbiasedness property applies to the small sample while the consistency property applies only to the large sample. The applicability of the unbiasedness to small sample is the benefit of being able to establish unbiasedness of estimators.

1-8. $E(u^2 x' x) \neq \sigma^2 E(x' x)$. Usual OLS standard error is inconsistent. Heteroscedasticity itself does not cause the inconsistency of the OLS estimator.

1-9. Multicolinearity is the situation when explanatory variables are highly correlated. This itself does not cause the inconsistency of the OLS estimator, but the precise estimation of the coefficients become difficult. This result in the large standard errors for the estimated coefficients.

1-10. OLS estimator is more efficient than IV estimator. Among the candidates of instrumental variables, x itself has the strongest correlation with x.

2 Problem 2: Standard Questions

2-1. For notational simplicity, we denote rice price as p_{ij} although all villagers face the same rice price within a village. We also introduce a notation:

$\bar{p} \equiv (1/N)\sum_{i=1}^{N} p_{ij}$.

$$\begin{aligned}
\hat{\beta}_1 &= \frac{\sum_{i=1}^{N}(p_{ij}-\bar{p})q_{ij}}{\sum_{i=1}^{N}(p_{ij}-\bar{p})^2} \\
&= \frac{\beta_0\sum_{i=1}^{N}(p_{ij}-\bar{p}) + \sum_{i=1}^{N}(p_{ij}-\bar{p})p_{ij}\beta_1 + \sum_{i=1}^{N}(p_{ij}-\bar{p})u_{ij}}{\sum_{i=1}^{N}(p_{ij}-\bar{p})^2} \\
&= \beta_1 + \frac{\sum_{i=1}^{N}(p_{ij}-\bar{p})u_{ij}}{\sum_{i=1}^{N}(p_{ij}-\bar{p})^2} \\
&\to \beta_1 + \frac{Cov(p,u)}{Var(p)} \\
&= \beta_1 + \frac{Cov(\frac{q_j^s-\beta_0-c_j}{\beta_1}, c+e)}{Var(p)} \\
&= \beta_1 - \frac{1}{\beta_1}\frac{Var(c)}{Var(p)}, \qquad (4)
\end{aligned}$$

I have assumed $Cov(c,e) = 0$. This assumption is natural because the error component that is common across villagers are capture in c and the remaining error component e should not be systematically correlated with e. The systematic correlation between c and e implies that c does not well capture common error component of villagers.

2-2. Assuming downward sloping demand curve (i.e. $\beta_1 < 0$), the bias term has positive sign and the $\hat{\beta}$ is asymptotically upward biased.

2-3. The problem is summarized as following.

$$q^d = \beta_0 + \beta_1 p + u, \quad E(pu) \neq 0,$$

$$p = \frac{q_j^s(z) - \beta_0 - c_j}{\beta_1}, \quad Ez'u = 0.$$

The weather information z determines p but not correlated with u and it serves as IV for the endogenous variable p.

2-4. Regress p on a constant and z. Implement t-test to test whether the coefficient for z is zero or not.

2-5. Estimate

$$q = \beta_0 + \beta_1 p + \delta_1 z_2 + error$$

using z_1 as IV for p and t-test $H_0 : \delta_1 = 0$.

Otherwise, estimate

$$q = \beta_0 + \beta_1 p + \delta_2 z_1 + error$$

using z_2 as IV for p and t-test $H_0 : \delta_2 = 0$.

2-6.

$$plim \hat{\beta}_1^{IV} = \beta + \frac{Cov(z_1, u)}{Cov(p, z_1)}.$$

From the given information, $Cov(z, u) < 0$ and $Cov(p, z) < 0$. The IV estimator is asymptotically upward biased.

3 Problem 3: Advanced Questions

3-1.

Estimated Model is

$$\ln y_{2004} - (\ln y_{1965} - e) = \beta_0 + \beta_1(\ln y_{1965} - e) + u$$

or alternatively,

$$\ln(y_{2004}/y_{1965}) = \beta_0 + \beta_1 \ln y_{1965} + (-e - \beta_1 e + u)$$

$$
\begin{aligned}
plim\hat{\beta}_1 &= \beta_1 + \frac{Cov(\ln y_{1965}, -(1+\beta_1)e + u)}{Var(\ln y_{1965})} \\
&= \beta_1 + \frac{Cov(\ln y^*_{1965} + e, -(1+\beta_1)e + u)}{Var(\ln y_{1965})} \\
&= \beta_1 - \frac{(1+\beta_1)Var(e)}{Var(\ln y_{1965})} \\
&= \beta_1(1 - \frac{Var(e)}{Var(\ln y_{1965})}) - \frac{Var(e)}{Var(\ln y_{1965})} \\
&= \beta_1 \frac{Var(\ln y^*_{1965})}{Var(\ln y_{1965})} - \frac{Var(e)}{Var(\ln y_{1965})}. \qquad (5)
\end{aligned}
$$

The first term in the above expression corresponds to the usual attenuation bias term. This makes the estimate biased toward zero. Given $\beta_1 < 0$, the attenuation bias causes asymptotic upward bias. On the other hand, the second term exhibits a negative bias. Thus we cannot determine the sign of bias a priori.

3-2.

We need to find an instrument that determines GNP in 1965 but does not directly affect the growth rate afterward. If the instrument is measured with error, that measurement error should not be correlated with e in the above expression. Average GNP between 1960 and 1964, for example, could serve as an instrument. By taking average the issue of measurement error can be mitigated because positive and negative measurement error for several years could cancel out each other.