ECONOMETRICS II (ECO 2401S)

University of Toronto. Department of Economics. Spring 2008
Professor Victor Aguirregabiria

SOLUTION TO TEST (April 9, 2008)

QUESTION 1 (20 points): Let θ_0 be a $K \times 1$ vector of parameters. $\hat{\theta}_A$ and $\hat{\theta}_B$ are two consistent estimators of θ_0 with asymptotic variances V_A and V_B, respectively, such that $V_B - V_A$ is a positive definite matrix. Let β_0 be a scalar parameter such that $\beta_0 = f(\theta_0)$, where $f(.)$ is a function from $\mathbb{R}^K$ into $\mathbb{R}$ that is continuous and differentiable. The gradient vector $G_0 \equiv \partial f(\theta_0)/\partial\theta$ has elements different than zero. Consider the following estimators of β_0: $\hat{\beta}_A \equiv f(\hat{\theta}_A)$ and $\hat{\beta}_B \equiv f(\hat{\theta}_B)$.

(a) Use the delta method to obtain the expression of the asymptotic variances of $\hat{\beta}_A$ and $\hat{\beta}_B$. [10 POINTS]

The Delta Method is based on the application of the Mean Value Theorem (MVT) to the function $f(\theta)$ between the true value θ_0 and a consistent estimator $\hat{\theta}$. By the MVT we have that:

$$f(\hat{\theta}) = f(\theta_0) + \frac{\partial f(\theta^*)}{\partial\theta}\left(\hat{\theta} - \theta_0\right)$$

where θ^* is such that $\|\theta^* - \theta_0\| \leq \left\|\hat{\theta} - \theta_0\right\|$. Since $\hat{\theta}$ is a consistent estimator of θ_0, it is clear that θ^* should be also a consistent estimator of θ_0. Therefore, $\partial f(\theta^*)/\partial\theta = \partial f(\theta_0)/\partial\theta + o_p(1)$. Given these results, we have that:

$$\left(\hat{\beta} - \beta_0\right) = G_0'\left(\hat{\theta} - \theta_0\right) + o_p(1)$$

And this implies that:

$$AVar\left(\hat{\beta}\right) = G_0' \; AVar\left(\hat{\theta}\right) G_0$$

Then, $AVar\left(\hat{\beta}_A\right) = G_0' V_A G_0$ and $AVar\left(\hat{\beta}_B\right) = G_0' V_B G_0$.

(b) Show that $Var(\hat{\beta}_B) > Var(\hat{\beta}_A)$. [10 POINTS]

$Var(\hat{\beta}_B) - Var(\hat{\beta}_A) = G_0'\left[V_B - V_A\right] G_0$. By definition of positive-definite matrix, for any vector $\lambda \neq 0$, $\lambda'\left[V_B - V_A\right]\lambda > 0$. Since $G_0 \neq 0$, we have that $G_0'\left[V_B - V_A\right] G_0 > 0$.

QUESTION 2 (40 points): Consider the linear regression model $Y = X'\theta_0 + \varepsilon$, where X is a $K \times 1$ vector of regressors. The regressors are potentially endogenous: $E(X\varepsilon) \neq 0$. There is a $q \times 1$ vector of instruments such that $E(W\varepsilon) = 0$, where $q > K$. The error term is heteroskedastic: i.e., $E(\varepsilon^2|X, W)$ depends on (X, W). We have a random sample $\{y_i, x_i, w_i : i = 1, 2, ..., n\}$.

(a) Define the optimal GMM estimator of θ_0 in this model. [15 POINTS]

The model implies the moment conditions $E(W[Y - X'\theta_0]) = 0$. For arbitrary θ, let $m_n(\theta)$ be the sample moment conditions:

$$m_n(\theta) \equiv \frac{1}{n}\sum_{i=1}^{n} w_i (y_i - x_i'\theta)$$

Define also Ω_0 as the variance matrix of the variable $W\varepsilon$: that is, $\Omega_0 \equiv E(W\varepsilon^2 W')$. The optimal GMM estimator of θ_0 is defined as:

$$\hat{\theta} = \arg\min_{\theta} \ m_n(\theta)'\hat{\Omega}_n^{-1} m_n(\theta)$$

where $\hat{\Omega}_n$ is a consistent estimator of Ω_0. For instance, suppose that $\hat{\varepsilon}_i = y_i - x_i'\theta^*$ where θ^* is a consistent estimator of θ_0. Then, a valid $\hat{\Omega}_n$ is:

$$\hat{\Omega}_n = \frac{1}{n}\sum_{i=1}^{n} \hat{\varepsilon}_i^2 w_i w_i'$$

(b) Derive the analytical closed-form expression of this optimal GMM estimator. [15 POINTS]

The GMM estimator solves the first order conditions (f.o.c.):

$$\frac{\partial m_n(\hat{\theta})'}{\partial \theta}\hat{\Omega}_n^{-1} \ m_n(\hat{\theta}) = 0$$

These f.o.c. are necessary conditions. For this model, it is straightforward to show that the criterion function is globally convex. Therefore, the f.o.c. are both necessary and sufficient. The f.o.c. can be written as:

$$\left(\sum_{i=1}^{n} w_i x_i'\right)' \hat{\Omega}_n^{-1} \left(\sum_{i=1}^{n} w_i \left(y_i - x_i'\hat{\theta}\right)\right) = 0$$

And solving for $\hat{\theta}$ we can get the closed-form expression:

$$\hat{\theta} = \left[\left(\sum_{i=1}^{n} w_i x_i'\right)' \hat{\Omega}_n^{-1} \left(\sum_{i=1}^{n} w_i x_i'\right)\right]^{-1} \left[\left(\sum_{i=1}^{n} w_i x_i'\right)' \hat{\Omega}_n^{-1} \left(\sum_{i=1}^{n} w_i y_i\right)\right]$$

When $\hat{\Omega}_n = \frac{1}{n}\sum_{i=1}^{n} \hat{\varepsilon}_i^2 w_i w_i'$, we have:

$$\begin{aligned}\hat{\theta} \ &= \ \left[\left(\sum_{i=1}^{n} w_i x_i'\right)' \left(\sum_{i=1}^{n} \hat{\varepsilon}_i^2 w_i w_i'\right)^{-1} \left(\sum_{i=1}^{n} w_i x_i'\right)\right]^{-1} \\ &\quad \left[\left(\sum_{i=1}^{n} w_i x_i'\right)' \left(\sum_{i=1}^{n} \hat{\varepsilon}_i^2 w_i w_i'\right)^{-1} \left(\sum_{i=1}^{n} w_i y_i\right)\right]\end{aligned}$$

(c) Describe a procedure to implement the feasible optimal GMM estimator. [10 POINTS]

To obtain the feasible optimal GMM estimator we need an initial (or first-step) consistent estimator of θ_0 to construct $\hat{\Omega}_n$. Consider a first-step estimator based on the assumption that ε is homoskedastic: $E(W\varepsilon^2 W') = \sigma_\varepsilon E(WW')$. The estimator is:

$$\begin{aligned}\hat{\theta}^{(1)} &= \left[\left(\sum_{i=1}^{n} w_i x_i'\right)'\left(\sum_{i=1}^{n} w_i w_i'\right)^{-1}\left(\sum_{i=1}^{n} w_i x_i'\right)\right]^{-1} \\ &\quad \left[\left(\sum_{i=1}^{n} w_i x_i'\right)'\left(\sum_{i=1}^{n} w_i w_i'\right)^{-1}\left(\sum_{i=1}^{n} w_i y_i\right)\right]\end{aligned}$$

$\hat{\theta}^{(1)}$ is consistent but not efficient. Then, we define the residuals $\hat{\varepsilon}_i = y_i - x_i'\hat{\theta}^{(1)}$, and obtain the feasible optimal GMM estimator using the expression in Question 2(b).

QUESTION 3 (40 points): Consider the following system of equations:

$$\begin{array}{lrcl} (1) & Y & = & X\,\alpha + \beta\,W + \varepsilon \\ (2) & W & = & X\,\gamma + u \\ (3) & Z & = & X\,\delta + v \end{array}$$

where $E(X\varepsilon) = E(Xu) = E(Xv) = 0$. The error terms ε and u are correlated, and therefore W is an endogenous regressor in equation (1): i.e., $E(W\varepsilon) \neq 0$. We are interested in estimating the parameters α and β in equation (1) given a random sample $\{y_i, x_i, w_i, z_i : i = 1, 2, ..., n\}$.

(a) Make the necessary assumptions (in the context of this model) to have a consistent Instrumental Variables estimator (α, β). Explain why the estimator is consistent. Describe a procedure to implement the IV estimator. [20 POINTS]

Solving equation (2) into (1) we have:

$$\begin{array}{lrcl} (1') & Y & = & X\,\alpha + \beta\,\,(X\gamma) + (\varepsilon + \beta u) \\ & & = & X\,\alpha + \beta\,W^* + \varepsilon^* \end{array}$$

where $W^* \equiv X\gamma$ and $\varepsilon^* \equiv \varepsilon + \beta u$. Since W^* depends on X only, and X is not correlated with ε and u, it is clear that $E(W^*\varepsilon^*) = 0$. If the OLS estimator in the regression of Y on (X, W^*) exits, then that estimator is consistent. However, without further assumptions, that OLS estimator does not exist because W^* is perfectly collinear to X. An exclusion restriction can make the OLS estimator of (1') well-defined.

Exclusion Restriction: There is a variable $X_j \subset X$ such that $\gamma_j \neq 0$ and $\alpha_j = 0$.

Under this restriction, equation (1') becomes:

$$Y = X_{(-j)}\,\alpha_{(-j)} + \beta\,W^* + \varepsilon^*$$

where $X_{(-j)}$ is the vector X but excluding X_j. Since $\gamma_j \neq 0$, it is clear that (at least asymptotically) W^* is not perfectly collinear to $X_{(-j)}$, and the OLS estimator in the regression of Y on $(X_{(-j)}, W^*)$ exists.

Finally, note that $W^* \equiv X\gamma$ is not observable because the vector of parameters γ is unknown. Nevertheless, we can obtain a consistent estimator of γ, and therefore of W^*. The IV estimator can be implemented in two-steps.

First Step: OLS regression of W on X to obtain the estimator $\hat{\gamma}$ and the fitted values $\hat{W} = X\hat{\gamma}$.

Second Step: OLS regression of Y on $X_{(-j)}$ and $\hat{W}$.

Note that $W^* = \hat{W} + X(\gamma - \hat{\gamma})$. Therefore, in the regression in the second step we have:

$$Y = X_{(-j)}\,\alpha_{(-j)} + \beta\,\hat{W} + (\varepsilon^* + \beta X(\gamma - \hat{\gamma}))$$

The error term in this regression is $\varepsilon^* + \beta X(\gamma - \hat{\gamma})$. We know that ε^* is not correlated with $(X, \hat{W})$. But, in a finite sample, $(X, \hat{W})$ are correlated with $\beta X(\gamma - \hat{\gamma})$. However, this correlation goes to zero as the sample size goes to infinite because $\hat{\gamma}$ converges to γ. The IV estimator is consistent.

(b) Suppose that the conditions for IV estimation do not hold. Instead, suppose that $\varepsilon = \pi \ v + e$, where π is a parameter and e is a random variable such that $E(ue) = 0$. Propose a consistent estimator of (α, β). Explain why the estimator is consistent. Describe a procedure to implement this estimator. [20 POINTS]

Replacing expression $\varepsilon = \pi v + e$ into equation (1) we have:

$$Y \ = \ X \, \alpha + \beta \, W + \pi \, v + e$$

Given our assumptions, e is not correlated with (X, W, v). Therefore, if the OLS estimator in the regression of Y on (X, W, v) exits, then it is consistent. Though, we do not observe v, we can construct a consistent estimator of this unobservable. We can obtain the residual $\hat{v} = Z - X\hat{\delta}$, where $\hat{\delta}$ is the OLS estimator in the regression of Z on X. Finally, we need the variables $(X, W, \hat{v})$ not to be perfectly collinear. The OLS estimator in the regression of Y on $(X, W, \hat{v})$ is a control function estimator, and it is consistent under our assumptions.