14.383 Final Exam
October 2004
Jerry Hausman
Answer only 2 Questions

1. **Consider the specification:**

$$y_1 = \beta_{12} y_2 + \gamma_{11} Z_1 + \varepsilon_1$$
$$y_2 = \beta_{21} y_1 + \gamma_{22} Z_2 + \gamma_{23} Z_3 + \varepsilon_2$$

(i) Discuss the rank and order conditions for identification. What changes if $\gamma_{11} = 0$? What changes if both $\gamma_{11} = 0$? and $\sigma_{12} = 0$?

(ii) Calculate the reduced form and demonstrate how the structural coefficients can be estimated from the reduced form coefficients. Discuss how the reduced form coefficients can be used to derive an efficient estimator of the structural coefficients.

(iii) Give tests of over-identification for each equation separately or for both equations together. Is one type of test for over-identification preferred to the other?

(iv) For the first equation, specify White's efficient IV estimator when there is conditional heteroscedasticity. You notice the coefficient estimates seem to change a lot from the 2SLS estimates. What should you do? What happens if you apply White's efficient IV estimator to the second equation?

2. **Consider the specification with m equations:**

$$y_i = Y_i \beta_i + Z_i \gamma_i + \varepsilon_i$$
$$i = 1, \dots m., \quad Var(\varepsilon) = \Sigma \otimes \mathrm{I}$$

(i) Give the 3SLS estimator and derive its asymptotic distribution.

(ii) Give the FIML estimator and compare its asymptotic distribution to that of the 3SLS estimator. Demonstrate that even if the stochastic disturbances are not normal, the asymptotic distribution is still correct.

(iii) Give the GMM estimator and determine its asymptotic distribution. Compare it to the asymptotic distribution of 3SLS and FIML.

(iv) Suppose in equation (1) that plim $\frac{1}{T} Z_1' \varepsilon_1 \neq 0$. Discuss the effect on the 3SLS, FIML, and GMM estimators. Can you test whether plim $\frac{1}{T} Z_1' \varepsilon_1 \neq 0$?

14.383 Fall 2003

Final

Answer two out of three questions (50 points each).

1. Consider the specification:

$$y_1 = \beta_{11} y_2 + Z_1 \gamma_1 + \epsilon_1$$
$$y_2 = \beta_{21} y_1 + Z_2 \gamma_2 + \epsilon_2 \qquad var(\epsilon) = \Sigma \otimes I$$

 (i) Give the rank and order conditions for identification for each equation. Calculate the reduced form and discuss the rank and order conditions in terms of the reduced form parameters.
 (ii) Suppose you know that $\beta_{11} = 0$. Would ordinary least squares (OLS) on the second equation be unbiased or consistent? If not, give a condition under which it would be consistent.
 (iii) Again assume that $\beta_{11} = 0$. Determine the test of overid restriction on the second equation using 2SLS. Suppose you did 3SLS on both equations and you do a test of overid, give the form of the test and compare it to the 2SLS test of overid.
 (iv) Again assume that $\beta_{11} = 0$. You do maximum likelihood on both equations together assuming normality but not taking into account that y_1 is a jointly endogenous variable. What is the property of the ML estimator?

2. Consider the equation:

$$y_1 = Y_1 \beta_1 + Z_1 \gamma_1 + \epsilon_1$$

with instruments Z. You think you may have conditional heteroscedasticity in ϵ_1. The equation is overidentified.

 (i) Specify the optimal IV (OIV) estimator and prove that its choice of instruments is optimal.
 (ii) Suppose instead you did 2SLS on the equation. What are the properties of the 2SLS estimator? Compare it to the OIV estimator.
 (iii) Suppose the equation is just identified. Compare the 2SLS and OIV estimators.
 (iv) You take the 2SLS estimates and do a 2SLS test of overid. What are the properties of the test? (assume the equation is overid)

3. Consider the specification:

$$y_1 = \beta_{11} y_2 + Z_1 \gamma_1 + \epsilon_1$$
$$y_2 = Z \pi_2 + v_2$$

We assume homoscedasticity and normality. The dimension of Z_1 is k_1, and the dimension of Z is K with $K > k_1$.

 (i) Assume you do 2SLS on the first equation. What is the finite sample bias (to second order) of the estimator? Use the formulas to demonstrate that 2SLS is consistent.
 (ii) Calculate the bias of the OLS estimator of the first equation. Compare to the bias of the 2SLS estimate and determine conditions under which the OLS has less bias.
 (iii) Suppose the instruments are invalid. Determine the maximum inconsistency of the 2SLS estimator. In terms of the omitted variable formula determine why this direction maximizes the inconsistency.
 (iv) Suppose $\pi_2 = 0$. What happens to estimation of the first equation? Prof. Q (aYankees fan) says that you can use the 2SLS bias from (i) to get a consistent estimate of β in this situation. See if Prof. Q is correct.

2002 final

Answer only two questions.

1. You have the specification:

$$y_1 = Y_1\beta_1 + Z_1\gamma_1 + u_1 \qquad (1)$$

under the usual assumptions where there are r_1 included jointly endogenous variables, s_1 included predetermined variables, and K predetermined variables overall. The sample is a time series $t = 1, ..., T$.

(i) Discuss the rank and order condition and discuss how they relate to IV estimation of equation (1).

(ii) Assume equation (1) is just identified. However, you decide at time $T_1 < T$ that one of the excluded Z's has a coefficient that changes its values in a reduced form equation for one of the Y_1's. Does this change equation (1) to an overidentified equation?

(iii) Again assume equation (1) is just identified. Now at time $T_1 < T$ you decide that one of the included Y_1's has a coefficient that changes its value. Does this change equation (1) to an underidentified equation?

(iv) Now assume you have conditional heterscedasticity. Discuss an efficient estimator and how it compares to 2SLS under the following situations: (a) equation (1) just identified; (b) one of the excluded Z's has a coefficient that changes value at time $T_1 < T$; (c) one of the included Y_1's has a coefficient that changes its value at time $T_1 < T$.

2. You have the simultaneous equation system with m equations in stacked form

$$y = X\delta + u$$

(i) Derive the asymptotic distribution of the 3SLS estimator.

(ii) Derive the test of the overidentifying restrictions.

(iii) Suppose you do not reject the overidentifying restrictions in part (ii). However, you decide to test the over id restrictions separately for each of the m equations where you use 2SLS. You reject the over id restrictions for the first equation. Discuss the result in relation to not rejecting the over id restrictions based on 3SLS.

(iv) Suppose you do FIML. Discuss what you expect the relationship of $\hat{\delta}_F$ to $\hat{\delta}_3$, the FIML and 3SLS estimators will be. Derive the FIML estimator for the reduced form coefficients, $\hat{\pi}_F$. You note that they look quite different from the unrestricted reduced form coefficients $\hat{\pi}_u$. What do you conclude? Can you specify a test to see if the estimates of the reduced form coefficients differ significantly from each other?

3. You have the single equation specification

$$y_1 = \beta_1 y_2 + \epsilon_1 \quad (1)$$

where all the predetermined variables have been partialled out of the equation. You have K possible instruments.

(i) Specify mathematically what it means to partial out all the predetermined variables in equation (1). What is the relationship of doing 2SLS on equation (1) compared to 2SLS if the predetermined variables had not been partialled out?

(ii) Derive the approximate finite sample bias of 2SLS for equation (1). Discuss the factors that affect the bias.

(iii) Suppose you have used K_1 instruments to estimate equation (1) and you are trying to decide where to use an additional instrument. Discuss whether it is a good idea based on the following criteria: (a) first order asymptotic variance; (b) second order bias; (c) second order MSE

(iv) Suppose you do IV on the specification

$$y_2 = \gamma_1 y_1 + \tilde{\epsilon}_1 \quad (2)$$

Prove that to first order $\hat{\beta}_1 = \frac{1}{\hat{\gamma}_1}$. If $\hat{\beta}_1$ is substantially different than $\frac{1}{\hat{\gamma}_1}$, what do you conclude?