UNIVERSITY OF CAMBRIDGE
FACULTY OF ECONOMICS

M.Phil. in Economics
M.Phil. in Economic Research

Subject M300 Econometric Methods

Suggested solutions to Exercise Sheet 2

1. Consider the textbook example of IV regression

$$\ln\left(Q_i^{cigarettes}\right) = \beta_0 + \beta_1 \ln\left(P_i^{cigarettes}\right) + u_i$$

with instrument $Z_{1i} = SalesTax_i$. Here $\ln\left(Q_i^{cigarettes}\right)$ is the logarithm of the number of packs of cigarettes sold per capita in state i, $\ln\left(P_i^{cigarettes}\right)$ is the logarithm of the average real price per pack of cigarettes in state i including all taxes, $SalesTax_i$ is the portion of the tax on cigarettes arising from the general sales tax in state i.

(a) The results of the first and the second stage regressions of the 2SLS method of estimating β_1 are:

$$\begin{aligned} \ln\left(\widehat{P_i^{cigarettes}}\right) &= \underset{(0.03)}{4.63} + \underset{(0.005)}{0.031}\,\text{SalesTax}_i \\ \ln\left(\widehat{Q_i^{cigarettes}}\right) &= 9.72 - 1.08\ln\left(\widehat{P_i^{cigarettes}}\right), \end{aligned}$$

where the standard errors of the estimates in the first stage regression are given below the coefficient estimates. What is the value of $\hat{\beta}_1^{2SLS}$?

Suggested solution to (a) $\hat{\beta}_1^{2SLS} = -1.08$

(b) Check whether we have a weak instrument problem in the 2SLS regressions from a).

Suggested solution to (b) F-statistic from the first stage regression equals the square of t-statistic because there is only one instrument. Therefore, $F = \left(\frac{0.031}{0.005}\right)^2 = 38.44 > 10$. It does not seem that there is a weak instrument problem.

(c) Given the 2SLS results from a), what is the value of the J statistic for testing the instrument's exogeneity?

Suggested solution to (c) The value of J statistic is zero because we have a just-identified case here.

(d) Discuss why SalesTax_i is a good or bad instrument for $\ln\left(P_i^{cigarettes}\right)$ from the point of view of the two requirements for the instrument validity.

Suggested answer to (d) Answers here can vary. Students should discuss both instrument relevance and exogeneity. The relevance is more or less obvious. Sales taxes vary from state to state and passed on to consumers, so that they represent a part of the price. The exogeneity is somewhat more tricky. In this particular example, income is an omitted variable, and it may well be correlated with sales taxes violating exogeneity.

2. This problem is the empirical exercise E12.2 from the Stock and Watson's textbook. We reproduce it here for convenience. How does fertility affect labour supply? That is, how much does a woman's labour supply fall when she has an additional child? In this exercise you will estimate this effect using data for married women from the 1980 U.S. Census. The data are available on the textbook Web site www.pearsonhighered.com/stock_watson in the file Fertility and described in the file Fertility_Description (from the page with the above address, you will need to go to the Companion Website for the 3-rd edition of the textbook. From there, you'll go to Student Resources/Data for Empirical Exercises). The data set contains information on married women aged 21-35 with two or more children.

(a) Regress weeksworked on the indicator variable morekids using OLS. On average, do women with more than two children work less than women with two children? How much less?

Suggested solutions to (a) As can be seen from Table 1, the coefficient is -5.387, which indicates that women with more than 2 children work 5.387 fewer weeks per year than women with 2 or fewer children.

(b) Explain why the OLS regression estimated in (a) is inappropriate for estimating the causal effect of fertility (morekids) on labour supply (weeksworked).

Suggested solution to (b) Both fertility and weeks worked are choice variables. A woman with a positive labor supply regression error (a

	Estimation method		
Regressor	OLS	IV	IV
Morekids	-5.387 (0.087)	-6.313 (1.275)	-5.821 (1.246)
Additional regressors	$const$	$const$	$const, agem1, black,$ $hispan, othrace$
First stage F		1238.2	1280.9

Table 1: Results using full dataset

woman who works more than average) may also be a woman who is less likely to have an additional child. This would imply that Morekids is negatively correlated with the regression error, so that the OLS estimator of $\beta_{Morekids}$ is negatively biased.

(c) The data set contains the variable samesex, which is equal to 1 if the first two children are of the same sex (boy-boy or girl-girl) and equal to 0 otherwise. Are couples whose first two children are of the same sex more likely to have a third child? Is the effect large? Is it statistically significant?

Suggested solution to (c) The linear regression of morekids on samesex (a linear probability model) yields

$$\widehat{Morekids} = \underset{(0.001)}{0.346} + \underset{(0.002)}{0.066} * samesex$$

so that couples with samesex=1 are 6.6% more likely to have an additional child that couples with samesex=0. The effect is highly significant (t-statistic=35.2).

(d) Explain why samesex is a valid instrument for the instrumental variable regression of weeksworked on morekids.

Suggested solution to (d) Samesex is random and is unrelated to any of the other variables in the model including the error term in the labor supply equation. Thus, the instrument is exogenous. From (c), the first stage F-statistic is large (F=1238) so the instrument is relevant. Together, these imply that samesex is a valid instrument

(e) Is samesex a weak instrument?

Suggested solution to (e) No, see the answer to (d).

(f) Estimate the regression of weeksworked on morekids using samesex as an instrument. How large is the fertility effect on labour supply?

Suggested solution to (f) See column (2) of Table 1. The estimated value of $\beta_{Morekids}$ is -6.313. This is somewhat surprising because we expected a negative bias of OLS, and IV is supposed to correct for this.

(g) Do the results change when you include the variables agem1, black, hispan, and othrace in the labour supply regression (treating these variables as exogenous)? Explain why or why not.

Suggested solution to (g) See column (3) of Table 1. The results do not change in an important way. The reason is that samesex is unrelated to agem1, black, hispan, othrace, so that there is no omitted variable bias in IV regression in column (2).

3. Consider a regression

$$y_i = \beta_0 + \beta_1 x_i + u_i$$

with one endogenous regressor x_i. Let z_i be a valid instrument for x_i. Assume that (y_i, x_i, z_i) form an i.i.d. sequence. Denote the fitted value $\hat{\pi}_0 + \hat{\pi}_1 z_i$ from the first stage regression

$$x_i = \pi_0 + \pi_1 z_i + e_i$$

as $\hat{x}_i$.

(a) Denote the vector $(1, \hat{x}_i)'$ as V_i, and let $\beta = (\beta_0, \beta_1)'$. Show that

$$\hat{\beta}_{2SLS} = \beta + \left(\frac{1}{n}\sum_{i=1}^{n} V_i V_i'\right)^{-1} \left(\frac{1}{n}\sum_{i=1}^{n} V_i u_i\right).$$

Suggested solution to (a) Since $\hat{\beta}_{2SLS}$ is an OLS estimator from the second stage regression, we have

$$\hat{\beta}_{2SLS} = \left(\frac{1}{n}\sum_{i=1}^{n} V_i V_i'\right)^{-1} \left(\frac{1}{n}\sum_{i=1}^{n} V_i y_i\right).$$

On the other hand, $y_i = \beta_0 + \beta_1 x_i + u_i = V_i'\beta + u_i + \beta_1 (x_i - \hat{x}_i)$. Using this in the above formula, we get

$$\hat{\beta}_{2SLS} = \beta + \left(\frac{1}{n}\sum_{i=1}^{n} V_i V_i'\right)^{-1} \left(\frac{1}{n}\sum_{i=1}^{n} V_i \left(u_i + \beta_1 (x_i - \hat{x}_i)\right)\right).$$

It remains to prove that $\frac{1}{n}\sum_{i=1}^{n} V_i (x_i - \hat{x}_i) = 0$. But this is true because $x_i - \hat{x}_i$ is the residual from the first stage regression, and V_i is the vector of regressors from the first stage regression. Residuals are always orthogonal to regressors, hence, $\frac{1}{n}\sum_{i=1}^{n} V_i (x_i - \hat{x}_i) = 0$.

(b) Let $\tilde{V}_i = (1, \pi_0 + \pi_1 z_i)'$. Show that

$$V_i V_i' = \tilde{V}_i \tilde{V}_i' + \begin{pmatrix} 0 & \hat{\pi}_0 - \pi_0 + (\hat{\pi}_1 - \pi_1) z_i \\ \hat{\pi}_0 - \pi_0 + (\hat{\pi}_1 - \pi_1) z_i & (\hat{\pi}_0 + \hat{\pi}_1 z_i)^2 - (\pi_0 + \pi_1 z_i)^2 \end{pmatrix}.$$

Using the Slutsky theorem, prove that

$$\frac{1}{n}\sum_{i=1}^{n} V_i V_i' \xrightarrow{p} E\left(\tilde{V}_i \tilde{V}_i'\right).$$

Suggested solution to (b) Note that

$$V_i = \tilde{V}_i + \begin{pmatrix} 0 \\ \hat{\pi}_0 - \pi_0 + (\hat{\pi}_1 - \pi_1) z_i \end{pmatrix}. \tag{1}$$

Therefore,

$$\begin{aligned} V_i V_i' &= \tilde{V}_i \tilde{V}_i' + \tilde{V}_i \left(0, \hat{\pi}_0 - \pi_0 + (\hat{\pi}_1 - \pi_1) z_i\right) \\ &+ \begin{pmatrix} 0 \\ \hat{\pi}_0 - \pi_0 + (\hat{\pi}_1 - \pi_1) z_i \end{pmatrix} \tilde{V}_i' \\ &+ \begin{pmatrix} 0 \\ \hat{\pi}_0 - \pi_0 + (\hat{\pi}_1 - \pi_1) z_i \end{pmatrix} \left(0, \hat{\pi}_0 - \pi_0 + (\hat{\pi}_1 - \pi_1) z_i\right). \end{aligned}$$

Opening up the brackets, we get

$$V_i V_i' = \tilde{V}_i \tilde{V}_i' + \begin{pmatrix} 0 & \hat{\pi}_0 - \pi_0 + (\hat{\pi}_1 - \pi_1) z_i \\ \hat{\pi}_0 - \pi_0 + (\hat{\pi}_1 - \pi_1) z_i & (\hat{\pi}_0 + \hat{\pi}_1 z_i)^2 - (\pi_0 + \pi_1 z_i)^2 \end{pmatrix}. \tag{2}$$

Now,

$$\frac{1}{n} \sum_{i=1}^{n} \left(\hat{\pi}_0 - \pi_0 + (\hat{\pi}_1 - \pi_1) z_i\right) = (\hat{\pi}_0 - \pi_0) + (\hat{\pi}_1 - \pi_1) \frac{1}{n} \sum_{i=1}^{n} z_i$$

since $\hat{\pi}_0 \xrightarrow{p} \pi_0$, $\hat{\pi}_1 \xrightarrow{p} \pi_1$, and $\frac{1}{n}\sum_{i=1}^{n} z_i \xrightarrow{p} Ez_i$, we have, by Slutsky's theorem

$$\frac{1}{n} \sum_{i=1}^{n} \left(\hat{\pi}_0 - \pi_0 + (\hat{\pi}_1 - \pi_1) z_i\right) \xrightarrow{p} 0. \tag{3}$$

Further,

$$\begin{aligned} \frac{1}{n} \sum_{i=1}^{n} \left[(\hat{\pi}_0 + \hat{\pi}_1 z_i)^2 - (\pi_0 + \pi_1 z_i)^2\right] &= \hat{\pi}_0^2 + 2\hat{\pi}_0 \hat{\pi}_1 \frac{1}{n} \sum_{i=1}^{n} z_i + \hat{\pi}_1^2 \frac{1}{n} \sum_{i=1}^{n} z_i^2 \\ &\quad - \pi_0^2 - 2\pi_0 \pi_1 \frac{1}{n} \sum_{i=1}^{n} z_i - \pi_1^2 \frac{1}{n} \sum_{i=1}^{n} z_i^2 \\ &= \left(\hat{\pi}_0^2 - \pi_0^2\right) + \left(2\hat{\pi}_0 \hat{\pi}_1 - 2\pi_0 \pi_1\right) \frac{1}{n} \sum_{i=1}^{n} z_i \\ &\quad + \left(\hat{\pi}_1^2 - \pi_1^2\right) \frac{1}{n} \sum_{i=1}^{n} z_i^2. \end{aligned}$$

By Slutsky's theorem, $\hat{\pi}_0^2 - \pi_0^2 \xrightarrow{p} 0$, $2\hat{\pi}_0\hat{\pi}_1 - 2\pi_0\pi_1 \xrightarrow{p} 0$, and $\hat{\pi}_1^2 - \pi_1^2 \xrightarrow{p} 0$. Therefore, using Slutsky once again, we conclude that

$$\frac{1}{n} \sum_{i=1}^{n} \left[(\hat{\pi}_0 + \hat{\pi}_1 z_i)^2 - (\pi_0 + \pi_1 z_i)^2\right] \xrightarrow{p} 0. \tag{4}$$

Finally, by LLN,

$$\frac{1}{n} \sum_{i=1}^{n} \tilde{V}_i \tilde{V}_i' \xrightarrow{p} E\left(\tilde{V}_i \tilde{V}_i'\right). \tag{5}$$

Combining (2, 3, 4, and 5), we obtain

$$\frac{1}{n}\sum_{i=1}^{n} V_i V_i' \xrightarrow{p} E\left(\tilde{V}_i \tilde{V}_i'\right).$$

(c) Show that

$$V_i u_i = \tilde{V}_i u_i + \begin{pmatrix} 0 \\ (\hat{\pi}_0 - \pi_0)\, u_i + (\hat{\pi}_1 - \pi_1)\, z_i u_i \end{pmatrix}.$$

Using the Slutsky theorem, prove that

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n} V_i u_i \xrightarrow{d} N\left(0, E\left(u_i^2 \tilde{V}_i \tilde{V}_i'\right)\right).$$

Suggested solution to (c) Equality

$$V_i u_i = \tilde{V}_i u_i + \begin{pmatrix} 0 \\ (\hat{\pi}_0 - \pi_0)\, u_i + (\hat{\pi}_1 - \pi_1)\, z_i u_i \end{pmatrix}$$

follows from (1). Next,

$$\begin{aligned} & \frac{1}{\sqrt{n}}\sum_{i=1}^{n} (\hat{\pi}_0 - \pi_0)\, u_i + (\hat{\pi}_1 - \pi_1)\, z_i u_i \\ = \quad & (\hat{\pi}_0 - \pi_0)\, \frac{1}{\sqrt{n}}\sum_{i=1}^{n} u_i + (\hat{\pi}_1 - \pi_1)\, \frac{1}{\sqrt{n}}\sum_{i=1}^{n} z_i u_i. \end{aligned}$$

On the other hand, $\hat{\pi}_0 - \pi_0 \xrightarrow{p} 0$, $\hat{\pi}_1 - \pi_1 \xrightarrow{p} 0$, whereas, by CLT, $\frac{1}{\sqrt{n}}\sum_{i=1}^{n} u_i \xrightarrow{d} N\left(0, Var\left(u_i\right)\right)$ and $\frac{1}{\sqrt{n}}\sum_{i=1}^{n} z_i u_i \xrightarrow{d} N\left(0, Var\left(z_i u_i\right)\right)$. By the Slutsky theorem,

$$(\hat{\pi}_0 - \pi_0)\, \frac{1}{\sqrt{n}}\sum_{i=1}^{n} u_i + (\hat{\pi}_1 - \pi_1)\, \frac{1}{\sqrt{n}}\sum_{i=1}^{n} z_i u_i \xrightarrow{p} 0,$$

and $V_i u_i$ converges in distribution to the same limit as $\tilde{V}_i u_i$. Specifically,

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n} V_i u_i \xrightarrow{d} N\left(0, E\left(u_i^2 \tilde{V}_i \tilde{V}_i'\right)\right)$$

(d) Using your results in (a), (b), and (c), show that

$$\sqrt{n}\left(\hat{\beta}_{2SLS} - \beta\right) \xrightarrow{d} N\left(0, \left[E\left(\tilde{V}_i \tilde{V}_i'\right)\right]^{-1} E\left(u_i^2 \tilde{V}_i \tilde{V}_i'\right) \left[E\left(\tilde{V}_i \tilde{V}_i'\right)\right]^{-1}\right)$$

Suggested solution to (d) From (a),

$$\sqrt{n}\left(\hat{\beta}_{2SLS}-\beta\right)=\left(\frac{1}{n}\sum_{i=1}^{n}V_iV_i'\right)^{-1}\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n}V_iu_i\right).$$

From (b),

$$\frac{1}{n}\sum_{i=1}^{n}V_iV_i'\xrightarrow{p}E\left(\tilde{V}_i\tilde{V}_i'\right).$$

From (c),

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}V_iu_i\xrightarrow{d}N\left(0,E\left(u_i^2\tilde{V}_i\tilde{V}_i'\right)\right).$$

Combining this (using the Slutsky theorem and the continuous mapping theorem), we get

$$\sqrt{n}\left(\hat{\beta}_{2SLS}-\beta\right)\xrightarrow{d}N\left(0,\left[E\left(\tilde{V}_i\tilde{V}_i'\right)\right]^{-1}E\left(u_i^2\tilde{V}_i\tilde{V}_i'\right)\left[E\left(\tilde{V}_i\tilde{V}_i'\right)\right]^{-1}\right).$$

4. Consider a general IV regression model

$$Y_i=\beta_0+\beta_1X_{1i}+...+\beta_kX_{ki}+\beta_{k+1}W_{1i}+...+\beta_{k+r}W_{ri}+u_i$$

with instruments $Z_{1i},...,Z_{mi}$.

(a) Let $\hat{X}_{1i},...,\hat{X}_{ki}$ be fitted values from the first stage regressions, let $\tilde{u}_i$ be the residuals from the second stage regression, and let

$$\hat{u}_i=Y_i-\hat{\beta}_0^{2SLS}-\hat{\beta}_1^{2SLS}X_{1i}-...-\hat{\beta}_k^{2SLS}X_{ki}-\hat{\beta}_{k+1}^{2SLS}W_{1i}-...-\hat{\beta}_{k+r}^{2SLS}W_{ri}.$$

Show that

$$\hat{u}_i=\tilde{u}_i-\hat{\beta}_1^{2SLS}\left(X_{1i}-\hat{X}_{1i}\right)-...-\hat{\beta}_k^{2SLS}\left(X_{ki}-\hat{X}_{ki}\right)$$

Suggested solution to (a) By definition,

$$\tilde{u}_i=Y_i-\hat{\beta}_0^{2SLS}-\hat{\beta}_1^{2SLS}\hat{X}_{1i}-...-\hat{\beta}_k^{2SLS}\hat{X}_{ki}-\hat{\beta}_{k+1}^{2SLS}W_{1i}-...-\hat{\beta}_{k+r}^{2SLS}W_{ri}.$$

It follows that

$$\hat{u}_i=\tilde{u}_i-\hat{\beta}_1^{2SLS}\left(X_{1i}-\hat{X}_{1i}\right)-...-\hat{\beta}_k^{2SLS}\left(X_{ki}-\hat{X}_{ki}\right).$$

(b) Show that, for any $j=1,...,k$, the OLS estimates of all the coefficients in the regression of $X_{1i}-\hat{X}_{1i}$ on constant and $Z_{1i},...,Z_{mi},W_{1i},...,W_{ri}$ equal zero.

Suggested solution to (b) $X_{1i} - \hat{X}_{1i}$ are residuals from the first stage regression. $Z_{1i}, ..., Z_{mi}, W_{1i}, ..., W_{ri}$ are explanatory variables in the first stage regression. OLS residuals are always orthogonal to the corresponding regressors.

(c) Show that the OLS estimates of all the coefficients in the regression of $\tilde{u}_i$ on constant and $\hat{X}_{1i}, ..., \hat{X}_{ki}, W_{1i}, ..., W_{ri}$ equal zero.

Suggested solution to (c) $\tilde{u}_i$ is the OLS residual from a regression of Y_i on constant and $\hat{X}_{1i}, ..., \hat{X}_{ki}, W_{1i}, ..., W_{ri}$. Since OLS residuals are orthogonal to regressors, all the coefficients in the regression of $\tilde{u}_i$ on constant and $\hat{X}_{1i}, ..., \hat{X}_{ki}, W_{1i}, ..., W_{ri}$ equal zero.

(d) Let Z be an $n \times m$ matrix with i-th row $(Z_{1i}, ..., Z_{mi})'$ and let X be an $n \times k$ matrix with i-th row $(X_{1i}, ..., X_{ki})'$. Assume that the number of instruments m equals the number of the endogenous regressors k, and that both $Z'Z$ and $Z'X$ are invertible matrices. Argue that the fitted value from the regression of $\tilde{u}_i$ on constant and $\hat{X}_{1i}, ..., \hat{X}_{ki}, W_{1i}, ..., W_{ri}$ must be the same as the fitted value from the regression of $\tilde{u}_i$ on constant and $Z_{1i}, ..., Z_{mi}, W_{1i}, ..., W_{ri}$. Using this and the result from (c), show that the OLS estimates of all the coefficients in the regression of $\tilde{u}_i$ on constant and $Z_{1i}, ..., Z_{mi}, W_{1i}, ..., W_{ri}$ equal zero.

Suggested solution to (d) Let W be the $n \times r$ matrix with i-th row $(W_{1i}, ..., W_{ri})'$, $\hat{X}$ the $n \times k$ matrix with i-th row $\left(\hat{X}_{1i}, ..., \hat{X}_{ki}\right)'$, $\tilde{u} = (\tilde{u}_1, ..., \tilde{u}_n)$, and $\mathbf{1} = (1, ..., 1)$ (an n-dimensional vector of ones). Further, let $\hat{\alpha}, \hat{\beta}$, and $\hat{\gamma}$ be the vectors of the OLS coefficient estimates in the regression of $\tilde{u}_i$ on constant and $Z_{1i}, ..., Z_{mi}, W_{1i}, ..., W_{ri}$, and let $\tilde{\alpha}, \tilde{\beta}$, and $\tilde{\gamma}$ be the OLS coefficient estimates in the regression of $\tilde{u}_i$ on constant and $\hat{X}_{1i}, ..., \hat{X}_{ki}, W_{1i}, ..., W_{ri}$ That is,

$$\left\{\hat{\alpha}, \hat{\beta}, \hat{\gamma}\right\} = \arg\min_{\alpha,\beta,\gamma} \left\{(\tilde{u} - \alpha\mathbf{1} - Z\beta - W\gamma)' (\tilde{u} - \alpha\mathbf{1} - Z\beta - W\gamma)\right\}, \tag{6}$$

and

$$\left\{\tilde{\alpha}, \tilde{\beta}, \tilde{\gamma}\right\} = \arg\min_{\alpha,\beta,\gamma} \left\{\left(\tilde{u} - \alpha\mathbf{1} - \hat{X}\beta - W\gamma\right)' \left(\tilde{u} - \alpha\mathbf{1} - \hat{X}\beta - W\gamma\right)\right\}.$$

In particular, the fitted values from the regression of $\tilde{u}_i$ on constant and $Z_{1i}, ..., Z_{mi}, W_{1i}, ..., W_{ri}$ equal

$$\hat{\alpha}\mathbf{1} + Z\hat{\beta} + W\hat{\gamma},$$

and that from the regression of $\tilde{u}_i$ on constant and $\hat{X}_{1i}, ..., \hat{X}_{ki}, W_{1i}, ..., W_{ri}$ equal

$$\tilde{\alpha}\mathbf{1} + \hat{X}\tilde{\beta} + W\tilde{\gamma}.$$

Now, by definition, $\hat{X} = Z(Z'Z)^{-1} Z'X$. Since $Z'X$ is assumed to be invertible, we can write

$$Z = \hat{X}(Z'X)^{-1} Z'Z = \hat{X}A,$$

where $A = (Z'X)^{-1} Z'Z$ is an invertible k-dimensional matrix. Replacing Z in (6) by $\hat{X}A$, we get

$$\left\{\hat{\alpha}, \hat{\beta}, \hat{\gamma}\right\} = \arg\min_{\alpha,\beta,\gamma} \left\{ \left(\tilde{u} - \alpha\mathbf{1} - \hat{X}A\beta - W\gamma\right)' \left(\tilde{u} - \alpha\mathbf{1} - \hat{X}A\beta - W\gamma\right) \right\}.$$

This implies that

$$\left\{\hat{\alpha}, A\hat{\beta}, \hat{\gamma}\right\} = \arg\min_{\alpha,\delta,\gamma} \left\{ \left(\tilde{u} - \alpha\mathbf{1} - \hat{X}\delta - W\gamma\right)' \left(\tilde{u} - \alpha\mathbf{1} - \hat{X}\delta - W\gamma\right) \right\},$$

and hence

$$\tilde{\alpha} = \hat{\alpha}, \tilde{\beta} = A\hat{\beta}, \text{ and } \tilde{\gamma} = \hat{\gamma}.$$

This, and the equality $Z = \hat{X}A$ imply that

$$\hat{\alpha}\mathbf{1} + Z\hat{\beta} + W\hat{\gamma} = \tilde{\alpha}\mathbf{1} + \hat{X}\tilde{\beta} + W\tilde{\gamma}.$$

Furthermore, from (c), $\hat{\alpha} = 0, \hat{\beta} = 0$, and $\hat{\gamma} = 0$. Therefore, $\tilde{\alpha} = 0, \tilde{\beta} = 0$, and $\tilde{\gamma} = 0$ too.

(e) Combining the results from (a), (b), and (d), prove that the value of J statistic in the just-identified case ($m = k$) must be zero.

Suggested solution to (e) From (a), (b) and (d), the OLS estimates of the coefficients in the regression of $\hat{u}_i$ on constant and $Z_{1i}, ..., Z_{mi}, W_{1i}, ..., W_{ri}$ must be zero. Therefore, the F statistic for testing the null that the coefficients on Z are zero must be equal to zero. This implies that $J = 0$.

5. This problem is the empirical exercise E12.3 from the Stock and Watson textbook. We reproduce it here for convenience. On the textbook Web site www.pearsonhighered.com/stock_watson you will find the data set WeakInstrument that contains 200 observations on (Y_i, X_i, Z_i) for the instrumental variable regression $Y_i = \beta_0 + \beta_1 X_i + u_i$.

(a) Construct $\hat{\beta}_1^{2SLS}$, its standard error, and the 95% confidence interval for β_1.

Suggested solution to (a) The 2SLS estimate of beta1 is 1.1577. Assuming homoskedasticity, the standard error of the 2SLS estimate of beta1 equals 0.4269, and the 95 percent confidence interval is [0.3210, 1.9945].

(b) Compute the F-statistic for the regression of X_i on Z_i. Is there evidence of a "weak instrument" problem?

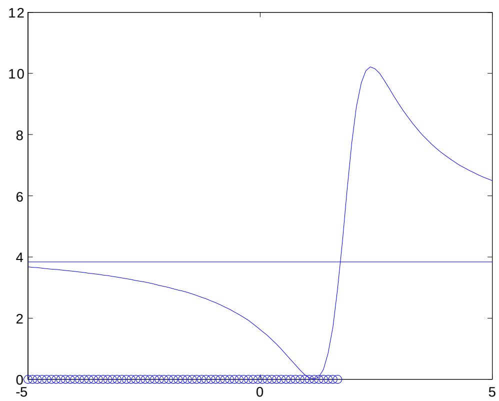

Figure 1: Values of Anderson-Rubin statistic. The horizontal line at the critical level 3.84. Circles on the horizontal axis represent the non-rejected points, which span the confidence set.

Suggested solution to (b) Assuming homoskedasticity, $F = 4.5661 < 10$. This suggests that there may be a weak instrument problem.

(c) Compute a 95% confidence interval for β_1 using the Anderson-Rubin procedure. (To implement the procedure, assume that $-5 \leq \beta_1 \leq 5$.)

Suggested solution to (c) Assuming homoskedasticity, the confidence interval is $[-5, 1.6667]$. Figure 1 shows a plot of the values of the AR-statistic for various $-5 \leq \beta_1 \leq 5$. The non-rejected points are denoted as circles on the horizontal axis. The non-rejection happens when the value of the F-statistic is above the horizontal line with ordinate 3.84.

(d) Comment on the differences in the confidence intervals in (a) and (c). Which is more reliable?

Suggested solution to (d) The confidence interval in (a) is not reliable because of the weak instrument problem. The confidence interval in (c) is reliable even when instruments are weak. Students were welcome to do the problem using a program of their choice, be it Excel, stata, E-views, or anything else. I have done the problem in Matlab. Here is the corresponding code.

Matlab code for problem 5 %first stage regression

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
n=200;%number of observations
c=ones(200,1); %create constant
Z=[c,z]; %create matrix of explanatory variables
pihat=inv(Z'*Z)*Z'*x; %OLS estimates of the coefficients
xhat=Z*pihat; %get fitted values
```

```
ehat=x-xhat;% residuals from the first stage
sigehat=ehat'*ehat/(n-2);%estimate of the variance of the
%error in the first stage regression
varpi=sigehat*inv(Z'*Z); %covariance matrix of the
%first stage estimates
tstat=pihat(2,1)/sqrt(varpi(2,2));%t-statistic from the first stage
Fstat=tstat^2;%F-statistic from the first stage
disp('F-statistic from the first stage regression equals')
disp(Fstat)

%second stage regression
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
Xhat=[c,xhat]; %create matrix of explanatory variables
beta2SLS=inv(Xhat'*Xhat)*Xhat'*y; %obtain 2SLS estimates
beta12SLS=beta2SLS(2,1); %2SLS estimate of beta1
disp('the 2SLS estimate of beta1 is')
disp(beta12SLS);

%asymptotic covariance matrix
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
X=[c,x]; %create matrix of explanatory variables
uhat=y-X*beta2SLS; %compute 2SLS residuals
%assuming homoscedasticity
sigmahat=uhat'*uhat/(n-2); %estimate of the variance of error
Omegahat=sigmahat*inv(Xhat'*Xhat); %estimate of the asymptotic
% variance

%standard error
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
stanerr=sqrt(Omegahat(2,2));%standard error estimate
disp('the standard error of the 2SLS estimate of beta1 equals')
disp(stanerr)

%confidence interval
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
disp('a 95 percent confidence interval is')
disp([beta12SLS-1.96*stanerr,beta12SLS+1.96*stanerr])

%Inverting Anderson Rubin test
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
betagrid=linspace(-5,5,100);%creat a grid of equally spaced
%point in between -5 and 5.
keep=zeros(100,1);% create an indicator variable, whose values
% will be changed to one for those points on the
% grid not rejected by AR test
for i=1:100 %start a loop
```

```
ynew=y-x*betagrid(1,i);%transform left-hand variable
betaAR=inv(Z'*Z)*Z'*ynew; %run Anderson-Rubin regression
resid=ynew-Z*betaAR;%get residuals from A-R regression
sigAR=resid'*resid/(n-2);%estimate of the variance of the error term
OmegaAR=sigAR*inv(Z'*Z);%covariance matrix of estimates
tAR=betaAR(2,1)/sqrt(OmegaAR(2,2));%t-statistic
Fstatistic(i,1)=tAR^2;% F-statistic
if Fstatistic(i,1)>3.84
%do nothing
else
%do not reject
keep(i,1)=1;%remember that you've not rejected this value
end
end
plot(betagrid,Fstatistic)%plot the values of AR statistic
hold on%use this to keep previous plot
line([-5 5],[3.84,3.84])%%critical value for AR statistic
plot(betagrid(keep==1),zeros(sum(keep),1),'o')% plot 'o' for non-rejected
```

6. Consider a regression

$$y_i = x_i'\beta + u_i,\ i = 1, ..., n,$$

where $x_i = (x_{1i}, ..., x_{ki})'$ and $E(u_i|x_i) \neq 0$. Suppose you have m instruments $z_i = (z_{1i}, ..., z_{mi})'$, and $m > k$. Assume that (y_i, x_i, z_i) is an i.i.d. sequence. Denote the $n \times m$ matrix with i-th row z_i' as Z, the $n \times k$ matrix with i-th row x_i' as X, and the $n \times 1$ vector with i-th element y_i as y.

(a) State the assumptions that $z_{1i}, ..., z_{mi}$ must satisfy to be valid instruments. Suppose these assumptions hold. Let $\hat{W} = Z(Z'Z)^{-1} Z'X$. Prove that

$$\hat{\beta}_{2SLS} = \left(\hat{W}'X\right)^{-1} \hat{W}'y \xrightarrow{p} \beta$$

Suggested solution to (a) The weakest IV assumptions implying consistency are

$$\frac{1}{n}Z'Z \xrightarrow{p} C > 0,$$
$$\frac{1}{n}Z'X \xrightarrow{p} D,\ \text{which is full column rank (relevance)}$$
$$\frac{1}{n}Z'u \xrightarrow{p} 0\ \text{(exogeneity)}.$$

The following shows consistency

$$
\begin{aligned}
\hat{\beta}_{2SLS} - \beta &= \left(X'Z \left(Z'Z \right)^{-1} Z'X \right)^{-1} X'Z \left(Z'Z \right)^{-1} Z'u \\
&= \left(\frac{X'Z}{n} \left(\frac{Z'Z}{n} \right)^{-1} \frac{Z'X}{n} \right)^{-1} \frac{X'Z}{n} \left(\frac{Z'Z}{n} \right)^{-1} \frac{Z'u}{n} \\
&\xrightarrow{p} \left(D'C^{-1}D \right)^{-1} D'C^{-1} 0 = 0
\end{aligned}
$$

(b) Suppose that heteroskedasticity is present, with $\mathrm{E}\left(u_i^2|z_i\right) = \sigma_i^2$. Prove that the variance of the asymptotic distribution of $\sqrt{n}\left(\hat{\beta}_{2SLS} - \beta\right)$ is given by

$$
\left(D'C^{-1}D \right)^{-1} D'C^{-1}BC^{-1}D \left(D'C^{-1}D \right)^{-1},
$$

where

$$
C = E\left(z_i z_i'\right), \ D = E\left(z_i x_i'\right), \ \text{and } B = E\left(u_i^2 z_i z_i'\right).
$$

Suggested solution to (b)

$$
\sqrt{n}\left(\hat{\beta}_{2SLS} - \beta\right) = \left(\frac{X'Z}{n} \left(\frac{Z'Z}{n} \right)^{-1} \frac{Z'X}{n} \right)^{-1} \frac{X'Z}{n} \left(\frac{Z'Z}{n} \right)^{-1} \frac{Z'u}{\sqrt{n}}
$$

On the other hand,

$$
\left(\frac{X'Z}{n} \left(\frac{Z'Z}{n} \right)^{-1} \frac{Z'X}{n} \right)^{-1} \frac{X'Z}{n} \left(\frac{Z'Z}{n} \right)^{-1} \xrightarrow{p} \left(D'C^{-1}D \right)^{-1} D'C^{-1}
$$

and

$$
\frac{Z'u}{\sqrt{n}} \xrightarrow{d} N\left(0, B\right)
$$

Therefore, by Slutsky's theorem

$$
\sqrt{n}\left(\hat{\beta}_{2SLS} - \beta\right) \xrightarrow{d} N\left(0, \left(D'C^{-1}D \right)^{-1} D'C^{-1}BC^{-1}D \left(D'C^{-1}D \right)^{-1}\right)
$$

(c) Take the residuals $\hat{u} = y - X\hat{\beta}_{2SLS}$ and form the $m \times n$ matrix

$$
\hat{Q} = \left[z_1 \hat{u}_1, ..., z_n \hat{u}_n \right].
$$

Then form $\hat{R} = \hat{Q}\hat{Q}'$. Consider

$$
\hat{\beta}^* = \arg\min \left\{ \left(y - X\beta \right)' Z\hat{R}^{-1}Z' \left(y - X\beta \right) \right\}.
$$

Find the first order conditions. Using these conditions, or otherwise, show that

$$
\hat{\beta}^* = \left(X'Z\hat{R}^{-1}Z'X \right)^{-1} X'Z\hat{R}^{-1}Z'y.
$$

Prove that $\hat{\beta}^*$ is consistent.

Suggested solution to (c) The first order conditions are

$$-2X'Z\hat{R}^{-1}Z'\left(y - X\hat{\beta}^*\right) = 0$$

Therefore,

$$\hat{\beta}^* = \left(X'Z\hat{R}^{-1}Z'X\right)^{-1} X'Z\hat{R}^{-1}Z'y.$$

Since $y = X\beta + u$, we have

$$\begin{aligned}\hat{\beta}^* - \beta &= \left(X'Z\hat{R}^{-1}Z'X\right)^{-1} X'Z\hat{R}^{-1}Z'u \\ &= \left(\frac{X'Z}{n}\hat{R}^{-1}\frac{Z'X}{n}\right)^{-1} \frac{X'Z}{n}\hat{R}^{-1}\frac{Z'u}{n} \\ &\xrightarrow{p} \left(D'B^{-1}D\right)^{-1} D'B^{-1}0 = 0\end{aligned}$$

(d) Show that the variance of the asymptotic distribution of $\sqrt{n}\left(\hat{\beta}^* - \beta\right)$ equals $\left(DB^{-1}D\right)^{-1}$

Suggested solution to (d)

$$\sqrt{n}\left(\hat{\beta}^* - \beta\right) = \left(\frac{X'Z}{n}\hat{R}^{-1}\frac{Z'X}{n}\right)^{-1} \frac{X'Z}{n}\hat{R}^{-1}\frac{Z'u}{\sqrt{n}}.$$

On the other hand,

$$\left(\frac{X'Z}{n}\hat{R}^{-1}\frac{Z'X}{n}\right)^{-1} \frac{X'Z}{n}\hat{R}^{-1} \xrightarrow{p} \left(D'B^{-1}D\right)^{-1} D'B^{-1}$$

(to show the convergence of $\hat{R}$ to B one needs to apply Slutsky theorem similarly to how this is done in problem 3) and

$$\frac{Z'u}{\sqrt{n}} \xrightarrow{d} N(0, B).$$

By the Slutsky theorem, we have

$$\sqrt{n}\left(\hat{\beta}^* - \beta\right) \xrightarrow{d} N\left(0, \left(D'B^{-1}D\right)^{-1} D'B^{-1}BB^{-1}D\left(D'B^{-1}D\right)^{-1}\right)$$

Performing the cancellations, we obtain

$$\sqrt{n}\left(\hat{\beta}^* - \beta\right) \xrightarrow{d} N\left(0, \left(D'B^{-1}D\right)^{-1}\right)$$

(e) Bonus question Prove that, under heteroskedasticity, $\hat{\beta}^*$ is asymptotically more efficient than $\hat{\beta}_{2SLS}$.

Suggested solution to (e) Note that

$$\left(D'C^{-1}D\right)^{-1} D'C^{-1}BC^{-1}D\left(D'C^{-1}D\right)^{-1} - \left(D'B^{-1}D\right)^{-1}$$

can be represented in the form

$$\left(D'C^{-1}D\right)^{-1} D'C^{-1}B^{1/2}\left(I - H\left(H'H\right)^{-1}H'\right)B^{1/2}C^{-1}D\left(D'C^{-1}D\right)^{-1},$$

where $H = B^{-1/2}D$. On the other hand,

$$\left(I - H\left(H'H\right)^{-1}H'\right) \geq 0$$

because this is an idempotent matrix (its square is equal to itself). Indeed

$$\begin{aligned}\left(I - H\left(H'H\right)^{-1}H'\right)\left(I - H\left(H'H\right)^{-1}H'\right) &= I - 2H\left(H'H\right)^{-1}H' \\ &\quad + H\left(H'H\right)^{-1}H'H\left(H'H\right)^{-1}H' \\ &= I - 2H\left(H'H\right)^{-1}H' \\ &\quad + H\left(H'H\right)^{-1}H' \\ &= I - H\left(H'H\right)^{-1}H'.\end{aligned}$$

Any matrix of the form $M_1'M_2M_1$, where $M_2 \geq 0$ is a positive semi-definite matrix. In particular,

$$\left(D'C^{-1}D\right)^{-1} D'C^{-1}BC^{-1}D\left(D'C^{-1}D\right)^{-1} - \left(D'B^{-1}D\right)^{-1} \geq 0,$$

which implies that

$$\left(D'C^{-1}D\right)^{-1} D'C^{-1}BC^{-1}D\left(D'C^{-1}D\right)^{-1} \geq \left(D'B^{-1}D\right)^{-1}.$$