Nikolay Laptev
Professor Shapiro
Economics 240B

Final Exam Winter 2007

(1)

We know that $Y_i = \beta X_i + \varepsilon_i$ is in deviations from mean form.

(i) What assumption must be true in order for $\hat{\beta}$ to be a consistent estimator for β ?

a. An estimator is said to be *consistent* $\lim_{n\to\infty} P(|\hat{\beta}_n - \beta| > \varepsilon) = 0$.

b. In other words the probability that $\hat{\beta}$ deviates from the population parameter β by more than some arbitrarily small value ε approaches 0 as the sample size approaches infinity.

c. When we show that $\hat{\beta}$ is consistent we will need to make an important assumption, namely that $Cov(\varepsilon_i, X_i) = 0$. We will show why this assumption is important as we go through the proof below.

(ii) Show that $\hat{\beta}$ is consistent

a. In this case we prove that $\hat{\beta}$ is consistent by applying the OLS (Ordinary Least squares approach).

b. We know that the fitted line is given by the liner equation: $\hat{Y}_i = \hat{\beta} X_i$

c. Then the deviation of the observation from the line would be given by $e_i = Y_i - \hat{Y}_i$, where e_i denotes the estimated error or residual. Note that e_i is only an estimation.

d. So the goal of OLS is to minimize the sum of squared errors from the line. That is select $\hat{\beta}$ to minimize: $S = \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} (Y_i - \hat{Y}_i)^2 = \sum_{i=1}^{n} (Y_i - \hat{\beta} X_i)^2$

e. So to begin we take first partial derivatives and equate them to zero:

$$\frac{\partial S}{\partial \hat{\beta}} = \frac{\partial \sum_{i=1}^{n} (Y_i - \hat{\beta} X_i)^2}{\partial \hat{\beta}_0} = 2\sum_{i=1}^{n} (Y_i - \hat{\beta} X_i)(-X_i) = \sum_{i=1}^{n} (Y_i - \hat{\beta} X_i)(-X_i) =$$

$$\sum_{i=1}^{n} (Y_i - \hat{\beta} X_i)(X_i)$$

f. We can expand out more to get:

$$\sum_{i=1}^{n}(Y_i X_i) = \hat{\beta}\sum_{i=1}^{n}(X_i)^2$$

$$\hat{\beta} = \frac{\sum_{i=1}^{n}(Y_i X_i)}{\sum_{i=1}^{n}(X_i)^2} = \frac{\sum_{i=1}^{n}((\beta X_i + \varepsilon_i)X_i)}{\sum_{i=1}^{n}(X_i)^2} =$$

$$= \frac{\sum_{i=1}^{n}(\beta X_i^2 + X_i\varepsilon_i)}{\sum_{i=1}^{n}(X_i)^2} = \beta + \frac{\sum_{i=1}^{n}\varepsilon_i X_i}{\sum_{i=1}^{n}X_i^2}$$

Now we use the probability limits to show consistency of $\hat{\beta}$.

$$p\lim(\hat{\beta}) = p\lim(\beta) + p\lim\left(\frac{\sum_{i=1}^{n}\varepsilon_i X_i}{\sum_{i=1}^{n}X_i^2}\right) =$$

$$= p\lim(\beta) + p\lim\left(\frac{\frac{\sum_{i=1}^{n}\varepsilon_i X_i}{n}}{\frac{\sum_{i=1}^{n}X_i^2}{n}}\right) = p\lim(\beta) + p\lim\left(\frac{Cov(\varepsilon_i, X_i)}{Var(X)}\right)$$

Since by assumption we have: $Cov(\varepsilon_i, X_i) = 0$ and we know that $p\lim(\beta) = \beta$

Using these assumptions the above probability limit approaches β, in other words $\hat{\beta}$ is *consistent.*

(2)

(i) What two conditions must be true for Z_i to be an instrument for X_i

a. Due to omitted variables or measurement error the $\operatorname{cov}(X, \varepsilon) \neq 0$ will cause our ordinary least squares estimator to be biased.
b. We can use an instrumental variable which would help us deal with the biasness. Suppose we have an instrumental variable Z. Z must satisfy the following two conditions:
    i. Z is correlated with X. That is $\operatorname{cov}(Z, X) \neq 0$
    ii. Z is uncorrelated with the error term in the regression. That is, $\operatorname{cov}(z, \varepsilon) = 0$

(ii) Derive the IV estimator for β and show that it is consistent

a. So we have: $Y_i = \beta X_i + \varepsilon_i$. We can multiply by Z and sum over all n observations and dividing by n, giving us: $\frac{\sum_{i=1}^{n} Z_i Y_i}{n} = \frac{\beta \sum_{i=1}^{n} Z_i X_i}{n} + \frac{\sum_{i=1}^{n} Z_i \varepsilon_i}{n}$

b. Notice that as $n \to \infty$ the last term would denote that covariance between Z and ε which is assumed to be 0. Therefore, as n becomes large the last term goes to 0, which proves our estimator to be **consistent**. (i.e. for large n our estimator approaches true value)

c. Thus, we can rewrite our instrumental variable to be: $\hat{\beta}_1 = \frac{\sum_{i=1}^{n} Z_i Y_i}{\sum_{i=1}^{n} Z_i X_i}$

d. (Another way to prove that it is consistent is to adapt similar strategy of plim as in the previous question)

(3)

We have the following two models:

Model I:
$Y_1 = \alpha_1 + \alpha_2 Y_2 + \alpha_3 X_1 + \alpha_4 X_2 + u_1$
$Y_2 = \beta_1 + \beta_2 Y_3 + \beta_3 X_2 + u_2$
$Y_3 = \gamma_1 + \gamma_2 Y_2 + u_3$

Model II:
$Y_1 = \alpha_0 + \alpha_1 Y_2 + \alpha_2 X_2 + u_1$
$Y_2 = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + u_2$

For each system we must determine which equations are under identified, just identified, and over identified.
Let's start with some definitions:
Just Identified model: A model with zero degrees of freedom (i.e. An identified Model in which the number of free parameters exactly equals the number of known values)
Under identified Model: A model for which it is not possible to estimate all of the parameters (i.e. Only some parameters are identified)
Over identified Model: A model for which all the parameters are identified and for which there are more known than free parameters.

The book States order conditions which can be used to determine which equation is identified, over identified, or under identified.

The outline of the order condition is as follows:
X= Number of exogenous variables excluded for equation
N=Number of endogenous variables on the right hand side of equation
Furthermore,
If X<N The equation is not identified
If X=N the equation is exactly identified
If X>N the equation is over identified

(i) Under identified equations:

a. $Y_1 = \alpha_1 + \alpha_2 Y_2 + \alpha_3 X_1 + \alpha_4 X_2 + u$ (X=0, N=1)

(ii) Over-identified equations:

a. $Y_3 = \gamma_1 + \gamma_2 Y_2 + u_3$ (X=2, N = 1)

b. $Y_1 = \alpha_0 + \alpha_1 Y_2 + \alpha_2 X_2 + u_1$ (X=2, N = 1)

(iii) Just identified equations:

a. $Y_2 = \beta_1 + \beta_2 Y_3 + \beta_3 X_2 + u_2$ (X =1, N =1)

b. $Y_2 = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + u_2$ ()X=0, Y=0)

(4)

(i) *Derive the Maximum Likelihood estimator for* $\hat{\beta}$

a. We know that the Maximum Likelihood estimator is designed to maximize the probability of observing the sample we actually see. Our assumption is that $\varepsilon_i \sim N(0,\sigma^2)$. Also note that this implies that $Y_i \sim N(\hat{\beta} X_i, \sigma^2)$, therefore the probability of observing an individual observation, y^i, is given by the particular functional form of the normal distribution: $f(y_i) = \frac{1}{\sqrt{2\pi\sigma^2}} e\left\{-\frac{1}{2}\frac{(Y_i - \beta X_i)^2}{\sigma^2}\right\}$

b. The probability of observing all of the n observations in our sample is what is established by our likelihood function: $L(\beta,\sigma^2) = f(y_1, y_2, y_3, ..., y_n)$. With the assumptions that our observations are independent we can rewrite the above as follows: $L(\beta,\sigma^2) = f(y_1)f(y_2)f(y_3)....f(y_n)$.

c. Combining it all together we get the following likelihood function:

$$L(\beta,\sigma^2) = f(y_1)f(y_2)f(y_3)....f(y_n) =$$

$$= \frac{1}{\sqrt{2\pi\sigma^2}} e\left\{-\frac{1}{2}\frac{(Y_1 - \beta X_1)^2}{\sigma^2}\right\} \frac{1}{\sqrt{2\pi\sigma^2}} e\left\{-\frac{1}{2}\frac{(Y_n - \beta X_n)^2}{\sigma^2}\right\}...$$

$$..\frac{1}{\sqrt{2\pi\sigma^2}} e\left\{-\frac{1}{2}\frac{(Y_n - \beta X_n)^2}{\sigma^2}\right\} = \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n e\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(Y_i - \beta X_i)^2\right\}$$

Taking the logs of both sides yields:

$$\ln L(\beta,\sigma^2) = \ln(f(y_1)f(y_2)f(y_3)....f(y_n)) =$$

$$= n\ln\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(Y_i - \beta X_i)^2$$

Now in order to actually get the maximum likelihood estimator we need to take derivative of the log likelihood with respect to β.

$$\frac{\partial \ln L(\beta,\sigma^2)}{\partial\beta} = \frac{\partial\left(\sum_{i=1}^{n}(Y_i - \beta X_i)^2\right)}{\partial\beta} =$$

$$= 2\sum_{i=1}^{n}(Y_i - \beta X_i)(-1) = 0$$

$$\Rightarrow \sum_{i=1}^{n}(Y_i - \beta X_i) = 0$$ (Dividing both sides by n)

$$\Rightarrow \sum_{i=1}^{n} Y_i = n\hat{\beta}\sum_{i=1}^{n} X_i)$$

$$\Rightarrow \overline{Y} = \hat{\beta}\,\overline{X}$$

$$\Rightarrow \hat{\beta} = \frac{\overline{Y}}{\overline{X}}$$

(ii) *Is it the same as the OLS estimator?*

Notice that this approach must give us the same result as the OLS approach because when we are taking these derivatives we are simply minimizing: $\sum_{i=1}^{n}(Y_i - \beta X_i)^2$

Thus, both the method of maximum likelihood (assuming the errors are normal) and the method of moments yield the same estimators as ordinary least squares.

(5)

Suppose that the true model is $Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \varepsilon_i$ but erroneously we estimate $Y_i = \beta_0 + \beta_1 X_{1i} + \varepsilon_i$ (i.e. we omit X_{2i}), now can we use $\hat{\beta}$ with any confidence? In order to examine this question we need to see if there is any correlation between X_{1i} and the variable we left out X_{2i}. By looking at chapter 9 from the book we can write the slope coefficient in a simple regression model as: $\hat{\beta}_1 = \dfrac{\sum_{i=1}^{n} Y_i(X_{1i} - \overline{X}_1)}{\sum_{i=1}^{n}(X_{1i} - \overline{X}_1)^2}$

We can substitute the correct model, namely $Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \varepsilon_i$ and take the expected value which yields:

$$E(\hat{\beta}_1) = E\left(\frac{\sum_{i=1}^{n}(\beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \varepsilon)(X_{1i} - \overline{X}_1)}{\sum_{i=1}^{n}(X_{1i} - \overline{X}_1)^2}\right) = \beta_0 \underbrace{\frac{\sum_{i=1}^{n}(X_{1i} - \overline{X}_1)}{\sum_{i=1}^{n}(X_{1i} - \overline{X}_1)^2}}_{=0} + \beta_1 \underbrace{\frac{\sum_{i=1}^{n}X_{1i}(X_{1i} - \overline{X}_1)}{\sum_{i=1}^{n}(X_{1i} - \overline{X}_1)^2}}_{=1}$$

$$+ \beta_2 \underbrace{\frac{\sum_{i=1}^{n}X_{2i}(X_{1i} - \overline{X}_1)}{\sum_{i=1}^{n}(X_{1i} - \overline{X}_1)^2}}_{Slope\_coefficient\_of\_a\_regression\_of\_X_2\_on\_X_1} + \underbrace{\frac{E\left(\sum_{i=1}^{n}\varepsilon_i(X_{1i} - \overline{X}_1)\right)}{\sum_{i=1}^{n}(X_{1i} - \overline{X}_1)^2}}_{=0} = \beta_1 + \beta_2 \frac{\sum_{i=1}^{n}X_{2i}(X_{1i} - \overline{X}_1)}{\sum_{i=1}^{n}(X_{1i} - \overline{X}_1)^2}$$

The book provides a simple interpretation of the above result.

Note that in a simple model where the dependent variable was X_2 and independent variable was X_1, we could write a regression equation of the form $X_{2i} = b_{12}X_{1i} + v_i$, where, for simplicity, we have assumed that the intercept equals zero and where the notation for the slope coefficient, b_{12}, indicates the impact of X_1 on X_2.

Using this specification, applying the same formula for a regression coefficient that we used just above we can see that $E(\hat{\beta}_1) = \beta_1 + \beta_2 b_{12}$ thus,

(i) *$\hat{\beta}_1$ will be a biased estimator of β_1 if $\beta_2 b_{12} \neq 0$*

(ii) *There will be no bias if $b_{12} = 0$ (i.e. if the omitted variable is unrelated to the included variable, an OLS regression of the miss specified model will still yield unbiased estimates of β_1*

(iii) *if b_{12} is nonzero then the direction and magnitude of the bias will depend on β_2 and b_{12}.*

(iv) *If b_{12} and β_2 are of the same sign, then their prod6yc will be positive and on average $\hat{\beta}_1$ will be larger than β_1*

(v) *If b_{12} and β_2 are of the opposite sign then their product will be negative and on average $\hat{\beta}_1$ will be smaller than β_1.*

Thus, our estimator will only be *consistent* if (i) holds.

(6)

Revised Question:
The underlying relationship between Y and X is as follows:

$Y_i = \beta X_i + \varepsilon_i$

where the density function of ε_i is $f(\varepsilon_i) = \exp(-\varepsilon_i)$ for ε_i non-negative and zero otherwise. The values of X are observed, by Y is an unobserved latent variable. The only thing know is the value of an indicator variable Z that is 1 when Y is greater than 1.5 and 0 when it is less than 1.5. Find the maximum likelihood estimate for β and test the hypothesis that β = 0.1 using a likelihood ratio test. **The value of the parameter β cannot exceed 0.5.**

Solution
Based on the information given above we can conclude the following:

When $Z_i = 0$
$BX_i + \varepsilon_i \le 1.5$
$\therefore \varepsilon_i \le 1.5 - BX_i$

When $Z_i = 1$
$BX_i + \varepsilon_i > 1.5$
$\therefore \varepsilon_i > 1.5 - BX_i$

Using the liklihood function we have:

$$L = \prod_{i=1}^{n} P(Z_i = 0 \mid X_i)^{Z_i} \; P(Z_i = 1 \mid X_i)^{1-Z_i}$$

$$P(Z_i = 0 \mid X_i) = P(\varepsilon_i < 1.5 \; - \beta_i X_i = \frac{1}{1 + e^{(1.5 \; - \beta_i X_i)}}$$

$$P(Z_i = 1 \mid X_i) = 1 - P(Z_i = 0 \mid X_i) = 1 - \frac{1}{1 + e^{(1.5 - \beta_i X_i)}} = \frac{e^{(1.5 - \beta_i X_i)}}{1 + e^{(1.5 - \beta_i X_i)}}$$

Plugging this into the liklihood function we get:

$$L = \prod_{i=1}^{n} \left(\frac{1}{1 + e^{(1.5 - \beta_i X_i)}} \right)^{Z_i} \left(\frac{e^{(1.5 - \beta_i X_i)}}{1 + e^{(1.5 - \beta_i X_i)}} \right)^{1-Z_i}$$

Taking the log we get:

$$\ln(L(\beta)) = \sum_{i=1}^{n} (-Z \ln(1 + e^{1.5 - \beta X_i}) - (1 - Z)(\ln(1 + 1 + e^{1.5 - \beta X_i}) - (1.5 \; - \beta X_i))$$

Running this equation on the dataset provided we get:

X	Z	B=0.1	B=0.2	B=0.3	B=0.4	B=0.5
1.877394	0	0.730789198	0.675192022	0.608112968	0.527180806	0.429534607
2.736984	0	0.706624406	0.614264432	0.492827857	0.33316084	0.12322774
0.283546	0	0.77045252	0.763850636	0.757058879	0.750071788	0.742883745
2.943069	1	0.299484342	0.401966596	0.539517837	0.724138521	0.97193561
2.037355	0	0.726448273	0.664632753	0.588848546	0.495939094	0.382034543
1.239817	0	0.747417742	0.714078109	0.676337807	0.633615969	0.585255055
0.888521	1	0.243863191	0.266522713	0.291287736	0.3183539	0.347935025
1.11922	0	0.750445515	0.720891873	0.687838323	0.650870386	0.609524499
0.630565	0	0.762346964	0.746878837	0.730403937	0.712856734	0.694167436
2.809156	1	0.295500609	0.391343824	0.518273005	0.686370631	0.908989352
0.037483	0	0.776031911	0.775190836	0.774346602	0.773499198	0.772648611
2.429275	0	0.715514366	0.637287601	0.53755034	0.41038771	0.24825839
1.725589	0	0.734845088	0.684905317	0.625559795	0.555037027	0.471231868
2.154813	0	0.723216233	0.656661145	0.574102302	0.47169146	0.344655032
2.941003	0	0.700577503	0.598199403	0.460816333	0.276459445	0.029067521
2.72467	1	0.293014552	0.384786744	0.505301997	0.663562641	0.871390537
2.077812	1	0.274660676	0.338091843	0.416172041	0.512284373	0.630593248
0.57213	1	0.236268354	0.250180141	0.264911072	0.28050938	0.297026137
0.563253	0	0.763941289	0.750263635	0.735793475	0.720484888	0.704289294
1.103447	1	0.249161183	0.278229062	0.310688085	0.346933874	0.387408203
0.526476	0	0.76480784	0.752093791	0.738692444	0.724566646	0.709677233
0.814503	0	0.757935165	0.737393707	0.715109115	0.690933467	0.664706291
0.102441	0	0.774572319	0.772251141	0.769906063	0.767536837	0.765143217
0.887586	1	0.243840402	0.266472903	0.291206082	0.318234917	0.347772484
1.629014	1	0.262606495	0.309067008	0.363747346	0.428101766	0.503841811
	likelihood	5.73789E-08	1.08104E-07	1.39743E-07	8.71253E-08	2.91634E-09

Thus our estimate for β *is 0.3* because it has the highest liklihood ratio.

Now we run the liklihood ratio test to determine if the hypothesis that β is 0.1 is valid.

In order to do this, we first compute the Chi-Squred with 1 degree of freedom.

Our unrestricted liklihood is: $L_{UR}=$1.39743E-07 and restricted: $L_R=$5.73789E-08.
$\chi^2=-2(L_R-L_{UR})=-2($ (ln(5.73789E-08)-ln(1.39743E-07))=1.78
And our critical value with 5% significance level is 3.84. Therefore we will reject the null hypothesis!

(7)

Note: To calculate the F statistic for the restricted models we use:

$$F=\frac{(R^2_{unrestricted}-R^2_{restricted})/m}{(1-R^2_{unrestricted})/(n-k-1)}$$

(a) H_0 : Y is distributed independently of the X's

To test this hypothesis we run the unrestricted least squares regression model in EViews. By comparing F statistic to the critical value we can reject or accept the hypothesis.

Important variables to keep in mind: (To compute the critical value of the F-Stat)

$m = 6$ (Numerator degrees of freedom)
$n - k - 1 = 18$ (Denominator degrees of freedom)

Based on this, the critical value of F stat is 2.66. If F > 2.66 we will reject the null hypothesis. Now let us run the regression:

Dependent Variable: Y
Method: Least Squares
Date: 03/20/07 Time: 17:49
Sample: 1 25
Included observations: 25

Variable	Coefficient	Std. Error	t-Statistic	Prob.
X1	2.301590	1.512241	1.521973	0.1454
X2	-1.030948	1.636179	-0.630095	0.5366
X3	-5.504074	0.726530	-7.575839	0.0000
X4	24.87606	0.156728	158.7211	0.0000
X5	-8.763135	0.554717	-15.79749	0.0000
X6	4.501716	5.043499	0.892578	0.3839
C	3.847244	6.528777	0.589275	0.5630

R-squared	0.999416	Mean dependent var	-17.07096
Adjusted R-squared	0.999221	S.D. dependent var	132.3258
S.E. of regression	3.692117	Akaike info criterion	5.681773
Sum squared resid	245.3711	Schwarz criterion	6.023058
Log likelihood	-64.02216	F-statistic	5135.049
Durbin-Watson stat	2.453113	Prob(F-statistic)	0.000000

Clearly based on the F statistic we *reject the null*!

(b)

$H_0 : \beta_1 = \beta_2$

$m = 1$ (Numerator degrees of freedom)
$n - k - 1 = 18$ (Denominator degrees of freedom)

Thus our critical F statistic if 4.41

Running the restricted model:

$y = \beta_0 + \beta_1(x_1 + x_2) + \beta_3 x_3 + \beta_4 x_4 + \beta_5 x_5 + \beta_6 x_6$

Dependent Variable: Y
Method: Least Squares
Date: 03/20/07 Time: 17:56
Sample: 1 25
Included observations: 25

Variable	Coefficient	Std. Error	t-Statistic	Prob.
XB	0.775803	1.188978	0.652496	0.5219
X3	-5.472806	0.752522	-7.272617	0.0000
X4	24.85779	0.161937	153.5025	0.0000
X5	-8.868024	0.570486	-15.54468	0.0000
X6	6.445122	5.061686	1.273315	0.2183
C	1.359196	6.557047	0.207288	0.8380

R-squared	0.999338	Mean dependent var	-17.07096
Adjusted R-squared	0.999164	S.D. dependent var	132.3258
S.E. of regression	3.825683	Akaike info criterion	5.726914
Sum squared resid	278.0812	Schwarz criterion	6.019444
Log likelihood	-65.58643	F-statistic	5738.852
Durbin-Watson stat	2.609907	Prob(F-statistic)	0.000000

Calculating the F stat using $F = \dfrac{(R^2_{unrestricted} - R^2_{restricted})/m}{(1 - R^2_{unrestricted})/(n-k-1)}$ we get 2.407. Thus we *accept* the null since the critical value of F stat is greater than computed F stat.

(c)

$H_0 : \beta_4 = -\beta_5 = \beta_6$

Again we use EViews to run a restricted model regression:

$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4(x_4 - x_5 + x_6)$

$m = 2$ (Numerator degrees of freedom)
$n - k - 1 = 18$ (Denominator degrees of freedom)

Thus our critical F statistic if 3.55

Dependent Variable: Y
Method: Least Squares
Date: 03/20/07 Time: 18:08
Sample: 1 25
Included observations: 25

Variable	Coefficient	Std. Error	t-Statistic	Prob.
X1	-1.925138	10.55406	-0.182407	0.8571
X2	8.015632	10.65096	0.752573	0.4605
X3	-9.360259	4.947559	-1.891894	0.0731
X456	24.64863	1.035695	23.79911	0.0000
C	-31.85656	8.791593	-3.623525	0.0017

R-squared	0.968108	Mean dependent var	-17.07096
Adjusted R-squared	0.961730	S.D. dependent var	132.3258
S.E. of regression	25.88656	Akaike info criterion	9.522182
Sum squared resid	13402.28	Schwarz criterion	9.765957
Log likelihood	-114.0273	F-statistic	151.7803
Durbin-Watson stat	1.893920	Prob(F-statistic)	0.000000

Calculating the F stat using $F = \dfrac{(R^2_{unrestricted} - R^2_{restricted})/m}{(1 - R^2_{unrestricted})/(n-k-1)}$ we get 482.584. Thus we *reject* the null since the critical value of F stat is less than computed F stat.

(d)

$H_0 : \beta_2 = \beta_3 = 0$

Thus the restricted model becomes:

$y = \beta_0 + \beta_1 x_1 + \beta_4 x_4 + \beta_5 x_5 + \beta_6 x_6$

$m = 2$ (Numerator degrees of freedom)

$n - k - 1 = 18$ (Denominator degrees of freedom)

Thus our critical F statistic if 3.55

Dependent Variable: Y
Method: Least Squares
Date: 03/20/07 Time: 18:18
Sample: 1 25
Included observations: 25

Variable	Coefficient	Std. Error	t-Statistic	Prob.
X1	3.998139	2.909431	1.374199	0.1846
X4	24.56293	0.286974	85.59289	0.0000
X5	-7.591062	1.036192	-7.325919	0.0000
X6	8.479456	9.202850	0.921395	0.3678
C	0.868775	11.42528	0.076040	0.9401

R-squared	0.997537	Mean dependent var	-17.07096
Adjusted R-squared	0.997045	S.D. dependent var	132.3258
S.E. of regression	7.193268	Akaike info criterion	6.961025
Sum squared resid	1034.862	Schwarz criterion	7.204800
Log likelihood	-82.01281	F-statistic	2025.430
Durbin-Watson stat	2.085082	Prob(F-statistic)	0.000000

Calculating the F stat using $F = \frac{(R^2_{unrestricted} - R^2_{restricted})/m}{(1-R^2_{unrestricted})/(n-k-1)}$ we get 28.9578. Thus we *reject* the null since the critical value of F stat is less than computed F stat.

(e)

Now we estimate $y^* = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_5 x_5 + \beta_6 x_6$
Where we denote $y^* = (y - 25x_4)$

$m = 1$ (Numerator degrees of freedom)
$n - k - 1 = 18$ (Denominator degrees of freedom)
Critical value of F = 4.41

Running the regression we get:

Dependent Variable: YMIN25
Method: Least Squares
Date: 03/20/07 Time: 18:27
Sample: 1 25
Included observations: 25

Variable	Coefficient	Std. Error	t-Statistic	Prob.
X1	2.119529	1.479807	1.432302	0.1683
X2	-1.341041	1.572760	-0.852668	0.4045
X3	-5.666159	0.690114	-8.210464	0.0000
X5	-8.895253	0.523720	-16.98474	0.0000
X6	3.279621	4.753344	0.689961	0.4986
C	5.586289	6.086377	0.917835	0.3702

R-squared	0.946916	Mean dependent var	16.92904
Adjusted R-squared	0.932946	S.D. dependent var	14.11691
S.E. of regression	3.655538	Akaike info criterion	5.635927
Sum squared resid	253.8962	Schwarz criterion	5.928457
Log likelihood	-64.44908	F-statistic	67.78433
Durbin-Watson stat	2.591900	Prob(F-statistic)	0.000000

Calculating the F stat using $F = \frac{(R^2_{unrestricted} - R^2_{restricted})/m}{(1 - R^2_{unrestricted})/(n-k-1)}$ we get 1618.5. Thus we *reject* the null since the critical value of F stat is less than computed F stat.