Exam 2

Dependent Variable: WHY
Method: Least Squares
Sample(adjusted): 1965 2006
Included observations: 30 after adjusting endpoints

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-19.60486	123.5064	-0.158736	0.8747
EX1	20.04112	12.56987	1.594377	0.1189
EX2	0.517140	0.007738	66.83540	0.0000

R-squared	0.991948	Mean dependent var	2775.633
Adjusted R-squared	0.991536	S.D. dependent var	1915.813
S.E. of regression	176.2591	Akaike info criterion	13.25054
Sum squared resid	1211624.	Schwarz criterion	13.37466
Log likelihood	-275.2613	F-statistic	2402.403
Durbin-Watson stat	0.289425	Prob(F-statistic)	0.000000

1. Use the Eviews output above to help with A) through I) below. (8 pts. each)

A) Interpret the coefficient on EX1.
If EX1 increases 1 unit, then WHY is expected to increase 20.04 units holding EX2 constant.

B) Interpret R^2 for this regression.
99% of the variation in WHY is explained by EX1 and EX2.

C) Replace the "?"'s with the appropriate values. An 80% confidence interval around the coefficient on EX1 is: 20.04112 +/- (?)(?) **12.57 and 1.314**

D) Perform a positive sign test on the coefficient attached to EX2. Show the 5-step procedure.

1) Ho: $\beta_2 = 0$ **Ha:** $\beta_2 > 0$ **(minus 2 for messing up step 1)**
2) 5%
3) If t-ratio > t^c, then reject Ho
4) 66.84 > 1.703 (d.f. = n- k = 30 – 3 =27)
5) Reject Ho → β_1 **is not significantly positive**

E) IF you performed a test of significance on the constant term at the 87 percent critical level, would you reject the null hypothesis or not? Explain.
Not reject because the P-value > .87

F) Perform a test to determine if the coefficient attached to EX2 is greater than 7.47. Show the 5-step procedure.

1. Ho: $\beta_2 = 7.47$ **Ha:** $\beta_2 > 7.47$
2. 5%
3. If t-ratio > t^c, then reject Ho
4. (20.04 - 7.47)/12.57 = 1 < 1.703 (d.f. = n - k = 30 – 3 =27)
5. Do not reject Ho ; β_2 **is not statistically greater than 7.47.**

G) Based on the information in the table above (and ignoring Ramsey's Reset test which we will perform in a moment) do you think the regression above is under or over specified?

Explain. **All we have to go on is significance and EX1 is insignificant at the 5% critical level. So the equation may be over specified.**

H) Perform Ramsey's Reset Test on the regression above using the Eviews printout below. Show the 5-step procedure.

Ramsey RESET Test:

F-statistic	7.975242	Probability	0.001319
Log likelihood ratio	15.05445	Probability	0.000538

$$F = \frac{SSR_U - SSR_R/2}{(1 - SSR_U)/(n-k)}$$

Test Equation:
Dependent Variable: WHY
Method: Least Squares
Sample: 1965 2006
Included observations: 30 after adjusting endpoints

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	1556.100	846.2011	1.838925	0.0740
EX1	0.680971	0.371065	1.835177	0.0745
EX2	-3.069166	1.944468	-1.578410	0.1230
FITTED^2	0.000103	2.62E-05	3.923561	0.0004
FITTED^3	-4.87E-09	1.22E-09	-3.977147	0.0003

R-squared	0.994212	Mean dependent var	5087.281
Adjusted R-squared	0.993586	S.D. dependent var	3714.167
S.E. of regression	297.4593	Akaike info criterion	14.33978
Sum squared resid	3273836.	Schwarz criterion	14.54664
Log likelihood	-296.1353	F-statistic	1588.804
Durbin-Watson stat	0.419302	Prob(F-statistic)	0.000000

1. Ho: B3 = B4 = 0 **Ha: B3 and/or B4 ≠ 0**
(No variables excluded) **(Variables excluded)**
2. 5%
3. If F > F^c , then reject Ho
4. 7.98 > 3.39; where F^c has d.f. NUM = 2 ; d.f. DEN = n-k = 30-5=25
5. Reject Ho: Variables are excluded from this regression.

I) What is the difference between "FITTED^2" and "FITTED^3" in the Eviews printout above? Do the two terms appear in the restricted or unrestricted regression?

FITTED^2 are the $\hat{Y}_i^2$ from the restricted regression; whereas FITTED^3 are the $\hat{Y}_i^3$. They appear in the unrestricted regression.

2. A) Criticize stepwise regression.* (4pts.) **need two separate criticisms**
It ignores theory.
Perhaps the pool stinks.
Includes or excludes variables based on biased statistics.
Ignores Ramsey's reset test, Akaike, Schwarz, Hannan-Quinn
preconceived bias of researcher?

B) What criteria are important for determining if a particular explanatory variable belongs in a regression. (4pts.) **Theory, significance, measures of goodness-of-fit, altered estimators, Ramsey's reset test.**

3. Mark each statement TRUE or FALSE. (2 pts. each)

A) **T** If a regression is underspecified, then the estimators will not average out to equal the true values of the structural parameters in repeated sampling.
B) **F** If a regression is over specified then the estimators should change markedly once the extraneous variables are removed.
C) **T** The constant term is biased in a regression missing an important explanatory variable.
D) **F** A 95% confidence interval for a given coefficient can never be larger than a 90% confidence interval.
E) **F** The standard error of the regression (SER) is monotonically decreasing with respect to the number of explanatory variables.
F) **T** Raising the critical level of a test from 5% to 10% decreases the probability that the test will result in a TYPE II error.
G) **F** The Schwarz Criterion is monotonically decreasing with respect to the number of explanatory variables.
H) **T** If the prob-value is less than .05, then the hypothesis test will reject Ho at the 5% critical level.
I) **T** You can conduct a positive sign test on a negative coefficient, but the results will be the same every time: Do not reject Ho.
J) **F** If a test of significance results in "Do not reject Ho", then there is absolutely no chance that a TYPE II error has been committed.

* "Stepwise regression" is a technique for determining what explanatory variables belong in a regression when a researcher is unsure. Basically, this technique runs a regression with one explanatory variable. Then a second regression is run adding an explanatory variable from a pool of variables selected by the researcher. If the added variable is statistically significant and improves adjusted R2 it is kept. Otherwise it is omitted. The next variable from the pool is added and a third regression is run. This process continues until all the variables from the pool have been tried. Only those that are significant and improve the fit of the regression are kept.

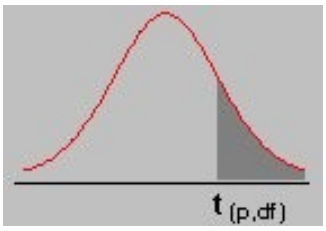
Critical Values of the F-Distribution (5%)

The table below shows critical values of the F distribution at the 5% level of significance. For example, the critical F (Fc) for 3 degrees of freedom in the numerator and 25 degrees of freedom in the denominator is 2.99.

df	df numerator										
denom	1	2	3	4	5	6	7	8	12	24	∞
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.68	4.53	4.37
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.00	3.84	3.67
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.57	3.41	3.23
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.28	3.12	2.93
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.07	2.90	2.71
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	2.91	2.74	2.54
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.79	2.61	2.41
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.69	2.51	2.30
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.60	2.42	2.21
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.53	2.35	2.13
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.48	2.29	2.07
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.42	2.24	2.01
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.38	2.19	1.96
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.34	2.15	1.92
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.31	2.11	1.88
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.28	2.08	1.84
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.25	2.05	1.81
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.23	2.03	1.78
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.20	2.01	1.76
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.18	1.98	1.73
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.16	1.96	1.71
26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.15	1.95	1.69
27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.13	1.93	1.67
28	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.12	1.91	1.66
29	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.10	1.90	1.64
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.09	1.89	1.62
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.00	1.79	1.51
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	1.92	1.70	1.39
80	3.96	3.11	2.72	2.49	2.33	2.21	2.13	2.06	1.88	1.65	1.33
100	3.94	3.09	2.70	2.46	2.31	2.19	2.10	2.03	1.85	1.63	1.28
120	3.92	3.07	2.68	2.45	2.29	2.18	2.09	2.02	1.83	1.61	1.26
∞	3.84	3.00	2.61	2.37	2.22	2.10	2.01	1.94	1.75	1.52	1.00

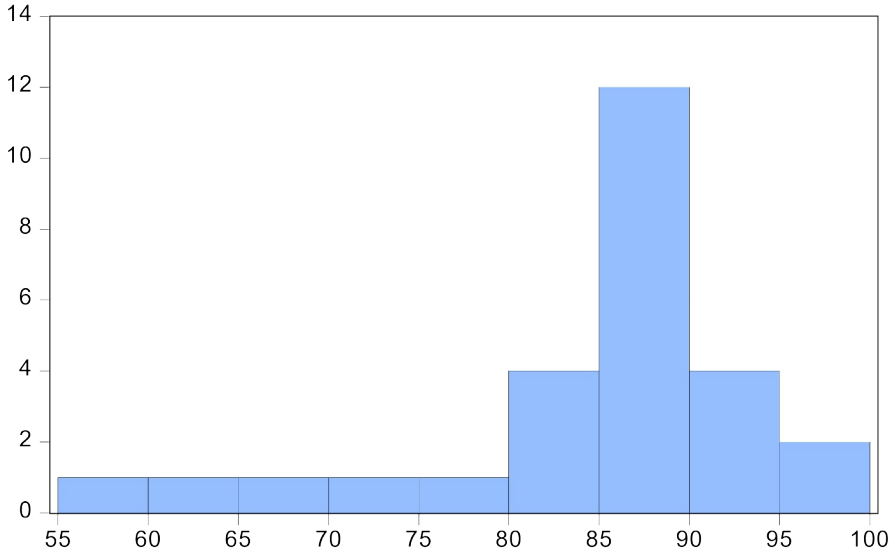
Critical Values of the t-Distribution

As indicated by the chart below, the areas given at the top of this table are the right tail areas for the t-value inside the table. For a one-tailed test at the 5% critical level with 6 degrees of freedom, look in the 5% column at the 6 row to get $t^c = 1.943$. For a two-tailed test at the 5% critical level with 6 degrees of freedom, look in the 2.5% column at the 6 row to get $t^c = 2.447$.



df	10%	5%	2.5%	1%
1	3.078	6.314	12.71	31.82
2	1.886	2.920	4.303	6.965
3	1.638	2.353	3.182	4.541
4	1.533	2.132	2.776	3.747
5	1.476	2.015	2.571	3.365
6	1.440	1.943	2.447	3.143
7	1.415	1.895	2.365	2.998
8	1.397	1.860	2.306	2.896
9	1.383	1.833	2.262	2.821
10	1.372	1.812	2.228	2.764
11	1.363	1.796	2.201	2.718
12	1.356	1.782	2.179	2.681
13	1.350	1.771	2.160	2.650
14	1.345	1.761	2.145	2.624
15	1.341	1.753	2.131	2.602
16	1.337	1.746	2.120	2.583
17	1.333	1.740	2.110	2.567
18	1.330	1.734	2.101	2.552
19	1.328	1.729	2.093	2.539
20	1.325	1.725	2.086	2.528
21	1.323	1.721	2.080	2.518
22	1.321	1.717	2.074	2.508
23	1.319	1.714	2.069	2.500
24	1.318	1.711	2.064	2.492
25	1.316	1.708	2.060	2.485
26	1.315	1.706	2.056	2.479
27	1.314	1.703	2.052	2.473
28	1.313	1.701	2.048	2.467
29	1.311	1.699	2.045	2.462
30	1.310	1.697	2.042	2.457
40	1.303	1.684	2.021	2.423
50	1.295	1.676	2.009	2.403
60	1.296	1.671	2.000	2.390
80	1.292	1.664	1.990	2.374
100	1.290	1.660	1.984	2.364
1000	1.282	1.646	1.962	2.330

∞ 1.282 1.640 1.960 2.326



Series: MID
Sample 1 27
Observations 27

Mean	84.63704
Median	86.35000
Maximum	97.50000
Minimum	57.60000
Std. Dev.	9.353194
Skewness	-1.307304
Kurtosis	4.363603
Jarque-Bera	9.782537
Probability	0.007512