

Tutorial 7

keywords: level log interpretation, dummy variables, hypothesis test, F-test, t-test, p-value, overall significance, individual significance, multiple linear restrictions, reparameterisation

estimated reading time: 35 minutes

Quang Bui

April 16, 2018

Question 1

EViews workfile: *vote1.wf1*

vote1.wf1 contains data on election outcomes and campaign expenditures for 173 two-party competitive races (Democrats and Republicans) for the House of Representatives in 1988. Information about each two-party competitive race is held in the following variables:

- votea* – %vote received by Candidate A
- expenda* – Candidate A's campaign expenditure (\$'000)
- expendb* – Candidate B's campaign expenditure (\$'000)
- democa* – 1 if Candidate A was a democrat, 0 otherwise (dummy variable)

Candidate A is chosen to be the candidate whose last name is alphabetically highest.

VOTE A	EXPENDA	EXPENDB	DEMOCA
68	328.296	8.737	1
62	626.377	402.477	0
73	99.607	3.065	1
69	319.69	26.281	0
75	159.221	60.054	0
69	570.155	21.393	1
59	696.748	193.915	0
71	638.688	7.695	1
76	616.936	19.245	1
73	351.687	50.532	1
68	269.887	14.71	1
71	269.51	95.575	1
52	1440.639	1089.57	0
79	252.336	69.563	1
50	1470.674	1548.193	0
64	140.486	100.956	1
72	191.334	15.449	1
68	398.597	15.239	1
60	460.622	382.111	1
67	457.41	20.608	1

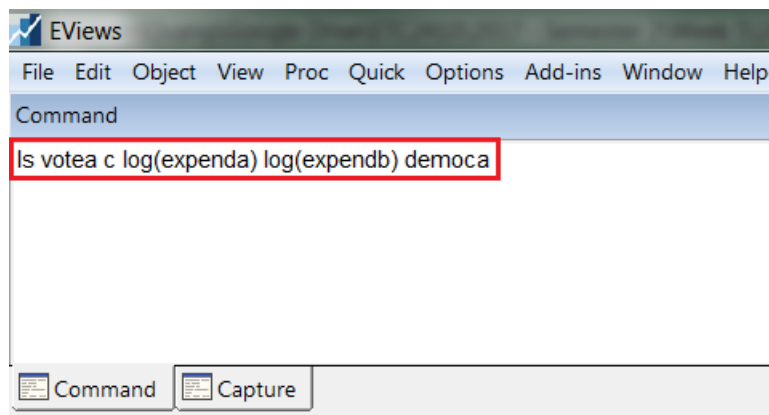
Table 1: Sample of the first 20 two-party competitive race.

Run a regression of *votea* on a constant, $\log(\text{expenda})$, $\log(\text{expendb})$ and *democa*

$$\text{votea} = \beta_0 + \beta_1 \log(\text{expenda}) + \beta_2 \log(\text{expendb}) + \beta_3 \text{democa} + u$$

To estimate this model from the Command window,

ls votea c log(expenda) log(expendb) democa

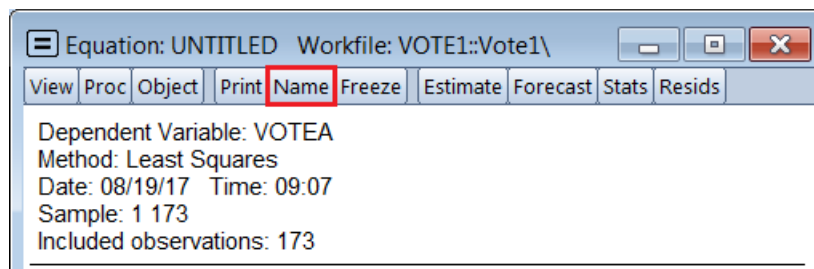


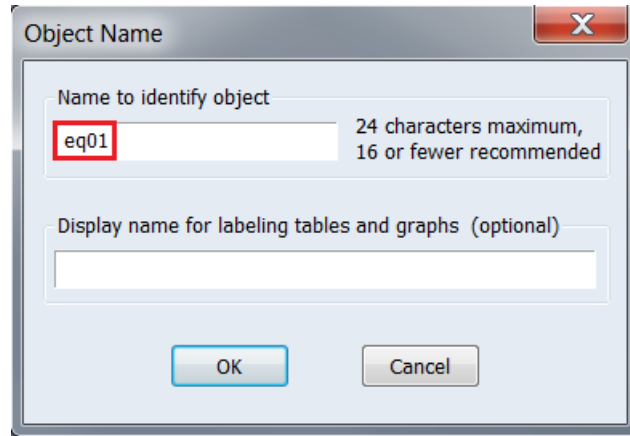
(Press Enter to execute code)

To name (save) the estimated equation,

Name \rightarrow *Name to identify object* : *eq01*

*(This names the equation **eq01**)*





Dependent Variable: VOTEA
Method: Least Squares
Date: 08/19/17 Time: 09:07
Sample: 1 173
Included observations: 173

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	51.13410	2.903327	17.61224	0.0000
LOG(EXPENDA)	6.299279	0.375274	16.78582	0.0000
LOG(EXPENDB)	-6.666045	0.391187	-17.04054	0.0000
DEMOCA	1.208824	1.241612	0.973593	0.3317
R-squared	0.786385	Mean dependent var	50.50289	
Adjusted R-squared	0.782593	S.D. dependent var	16.78476	
S.E. of regression	7.826209	Akaike info criterion	6.975683	
Sum squared resid	10351.17	Schwarz criterion	7.048592	
Log likelihood	-599.3966	Hannan-Quinn criter.	7.005262	
F-statistic	207.3816	Durbin-Watson stat	1.652138	
Prob(F-statistic)	0.000000			

Table 2: Regression output of *votea* on a constant, $\log(\text{expenda})$, $\log(\text{expendb})$ and *democa*.

$$\widehat{\text{votea}} = 51.1341 + 6.2993\log(\text{expenda}) - 6.6660\log(\text{expendb}) + 1.2088\text{democa}$$

(2.9033)
(0.3753)
(0.3912)
(1.2416)

$$R^2 = 0.7864$$

(a) Interpreting the regression results when explanatory variables are logarithms of the original variables and also interpreting the coefficient of dummy variables.

Explain what each parameter estimate shows.

Background

Logarithms for approximating percentage change

Since,

$$\frac{d\log(x)}{dx} = \frac{1}{x}$$

replacing infinitesimally small change d with finite change Δ gives the approximation of $\frac{1}{x}$,

$$\frac{\Delta\log(x)}{\Delta x} \approx \frac{1}{x}$$

and by multiplying Δx on both sides, we have the approximate proportional change in x ,

$$\Delta\log(x) \approx \frac{\Delta x}{x}$$

Multiplying 100 on both sides gives us the approximate percentage change in x ,

$$\begin{aligned} 100\Delta\log(x) &\approx 100\frac{\Delta x}{x} \\ &= \% \Delta x \end{aligned}$$

Since the approximation comes from replacing infinitesimally small change d with finite change Δ , this means that when the finite change Δ is large, the approximation will be less precise.

Level-log interpretation

For the following estimated model, which is level in the dependent variable and log in x_1 ,

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 \log(x_1) + \hat{\beta}_2 x_2$$

the change in \hat{y} depends on change in $\log(x_1)$ and x_2 ,

$$\Delta\hat{y} = \hat{\beta}_1 \Delta\log(x_1) + \hat{\beta}_2 \Delta x_2$$

Multiplying $\Delta\log(x_1)$ with $\frac{100}{100}$ gives us,

$$\Delta\hat{y} = \hat{\beta}_1 \frac{100}{100} \Delta\log(x_1) + \hat{\beta}_2 \Delta x_2$$

$$\begin{aligned}
\Delta \hat{y} &= \hat{\beta}_1 \frac{100}{100} \Delta \log(x_1) + \hat{\beta}_2 \Delta x_2 \\
&= \frac{\hat{\beta}_1}{100} 100 \Delta \log(x_1) + \hat{\beta}_2 \Delta x_2 \\
&= \frac{\hat{\beta}_1}{100} \% \Delta x_1 + \hat{\beta}_2 \Delta x_2
\end{aligned}$$

\therefore the model estimates that if x_2 is held constant and x_1 increases by 1%,

$$\Delta x_2 = 0$$

$$\% \Delta x_1 = 1$$

on average, y is estimated to change by $\frac{\hat{\beta}_1}{100}$,

$$\begin{aligned}
\Delta \hat{y} &= \frac{\hat{\beta}_1}{100} \% \Delta x_1 + \hat{\beta}_2 \Delta x_2 \\
&= \frac{\hat{\beta}_1}{100} \times 1 + \hat{\beta}_2 \times 0 \\
&= \frac{\hat{\beta}_1}{100}
\end{aligned}$$

(When x_1 increases by 1%, holding x_2 constant, \hat{y} is expected to change by $\frac{\hat{\beta}_1}{100}$ i.e.

$$\Delta \hat{y} = \frac{\hat{\beta}_1}{100}.)$$

Model	Dependent Variable	Independent Variable	Interpretation (population model)	Interpretation (estimated model)
<i>level – level</i>	y	x	$\Delta y = \beta_1 \Delta x$	$\Delta \hat{y} = \hat{\beta}_1 \Delta x$
<i>level – log</i>	y	$\log(x)$	$\Delta y = \left(\frac{\beta_1}{100}\right) \% \Delta x$	$\Delta \hat{y} = \left(\frac{\hat{\beta}_1}{100}\right) \% \Delta x$
<i>log – level</i>	$\log(y)$	x	$\% \Delta y = (100\beta_1) \Delta x$	$\% \Delta \hat{y} = (100\hat{\beta}_1) \Delta x$
<i>log – log</i>	$\log(y)$	$\log(x)$	$\% \Delta y = \beta_1 \% \Delta x$	$\% \Delta \hat{y} = \hat{\beta}_1 \% \Delta x$

$$\widehat{votea} = \underset{(2.9033)}{51.1341} + \underset{(0.3753)}{6.2993\log(expenda)} - \underset{(0.3912)}{6.6660\log(expendb)} + \underset{(1.2416)}{1.2088democa}$$

$$\hat{\beta}_1 = 6.2993$$

The model estimates that regardless of Candidate A's political affiliation (holding *democa* constant) and when there is no change in Candidate B's expenditure (holding *expendb* constant),

$$\Delta\log(expendb) = 0$$

$$\Delta democa = 0$$

a 1% increase in Candidate A's expenditure,

$$\% \Delta expenda = 1\%$$

is expected to increase the share of votes received by Candidate A by 0.063 percentage points,

$$\begin{aligned} \Delta \widehat{votea} &= \hat{\beta}_1 \Delta \log(expenda) + \hat{\beta}_2 \Delta \log(expendb) + \hat{\beta}_3 \Delta democa \\ &= \hat{\beta}_1 \Delta \log(expenda) + \hat{\beta}_2 \times 0 + \hat{\beta}_3 \times 0 \\ &= \hat{\beta}_1 \Delta \log(expenda) \\ &= \frac{\hat{\beta}_1}{100} \% \Delta expenda \\ &= \frac{6.2993}{100} \times 1 = 0.063 \end{aligned}$$

$$\hat{\beta}_2 = -6.6660$$

$$\Delta \widehat{votea} = \frac{\hat{\beta}_2}{100} \% \Delta expendb$$

(holding *expendb* and *democa* constant)

The model estimates that for a 1% increase in Candidate B's expenditure,

$$\% \Delta expendb = 1\%$$

the share of votes received by Candidate A is expected to decrease by 0.0667 percentage points,

$$\Delta \widehat{votea} = \frac{-6.6660}{100} \times 1 = -0.0667$$

regardless of Candidate A's political affiliation (holding *democa* constant) and when there is no change in Candidate A's expenditure (holding *expenda* constant).

Background

Dummy variable interpretation

For the following estimated model,

$$\widehat{votea} = \hat{\beta}_0 + \hat{\beta}_1 \log(expenda) + \hat{\beta}_2 \log(expendb) + \hat{\beta}_3 democa$$

where *democa* is a dummy variable,

$$democa = \begin{cases} 1 & \text{if Candidate A was a democrat} \\ 0 & \text{otherwise} \end{cases}$$

The change in \widehat{votea} depends on the change in $\log(expenda)$, $\log(expendb)$, and *democa*,

$$\Delta \widehat{votea} = \hat{\beta}_1 \Delta \log(expenda) + \hat{\beta}_2 \Delta \log(expendb) + \hat{\beta}_3 \Delta democa$$

\therefore the model estimates that when Candidate A and B's campaign expenditure are held constant,

$$\Delta \log(expenda) = 0$$

$$\Delta \log(expendb) = 0$$

the share of votes received by Candidate A if Candidate A is from the Democratic party,

$$\Delta democa = 1$$

is expected to be $\hat{\beta}_3$ percentage points higher than Candidate B,

$$\begin{aligned} \Delta \widehat{votea} &= \hat{\beta}_1 \times 0 + \hat{\beta}_2 \times 0 + \hat{\beta}_3 \times 1 \\ &= \hat{\beta}_3 \end{aligned}$$

(The model estimates that when Candidate A and B's campaign expenditure are held constant, the share of votes received by Candidate A if Candidate A is from the Democratic party is expected to be $\hat{\beta}_3$ percentage points higher than Candidate B.)

$$\hat{\beta}_3 = 1.2088$$

Controlling for both candidate's campaign expenditure, the model estimates that the share of votes received by Candidate A (*votea*) if Candidate A is from the Democratic party (*democa* = 1) is expected to be 1.2088 percentage points higher than if Candidate A were from the Republican party (*democa* = 0).

(b) Test the overall significance of a regression.

Test the overall significance of the model at the 1% significance level (ignore the fact that EViews produces the F statistic, compute it using R^2). Explain in words the hypothesis that you are testing.

$$votea = \beta_0 + \beta_1 \log(expenda) + \beta_2 \log(expendb) + \beta_3 democa + u$$

A test of overall significance is a test of whether the regressors in our model jointly help explain the dependent variable. If none of the regressors jointly help to explain *votea* then,

$$\beta_1 = \beta_2 = \beta_3 = 0$$

but if at least one of the regressors helps to explain *votea* then,

$$at\ least\ one\ of\ \beta_1, \beta_2, \beta_3\ is\ not\ 0$$

State the null and alternative hypothesis

$$H_0 : \beta_1 = \beta_2 = \beta_3 = 0$$

$$H_1 : at\ least\ one\ of\ \beta_1, \beta_2, \beta_3\ is\ not\ 0$$

The test statistic and its distribution under H_0

$$F = \frac{R^2/k}{(1 - R^2)/(n - k - 1)} = \frac{R^2/3}{(1 - R^2)/(173 - 3 - 1)} \sim F_{q, n-k-1} \quad under\ H_0$$

n = sample size

k = number of regressors in the model

q = number of restrictions (here it is equal to k)

Note: For a test of overall joint significance of a regression,

$$F = \frac{(SSR_r - SSR_{ur})/q}{SSR_{ur}/(n - k - 1)} = \frac{R^2/k}{(1 - R^2)/(n - k - 1)}$$

q = number of restrictions

Calculate the test statistic

$$F_{calc} = \frac{0.786385/3}{(1 - 0.786385)/(173 - 3 - 1)} = 207.38$$

Dependent Variable: VOTEA
Method: Least Squares
Date: 08/19/17 Time: 09:07
Sample: 1 173
Included observations: 173

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	51.13410	2.903327	17.61224	0.0000
LOG(EXPENDA)	6.299279	0.375274	16.78582	0.0000
LOG(EXPENDB)	-6.666045	0.391187	-17.04054	0.0000
DEMOCA	1.208824	1.241612	0.973593	0.3317
R-squared	0.786385	Mean dependent var	50.50289	
Adjusted R-squared	0.782593	S.D. dependent var	16.78476	
S.E. of regression	7.826209	Akaike info criterion	6.975683	
Sum squared resid	10351.17	Schwarz criterion	7.048592	
Log likelihood	-599.3966	Hannan-Quinn criter.	7.005262	
F-statistic	207.3816	Durbin-Watson stat	1.652138	
Prob(F-statistic)	0.000000			

Critical value and rejection region

1% *significance level* $\rightarrow \alpha = 0.01$

To obtain the critical value using the Stats Table, locate the F distribution table at the 1% significance level,

$$\text{Numerator } d.o.f = 3$$

$$\text{Denominator } d.o.f = 169$$

Since 169 is not in the table, we take a conservative approach and choose the closest available degrees of freedom less than 169 i.e. $d.o.f = 120$.

TABLE G.3c 1% Critical Values of the F Distribution

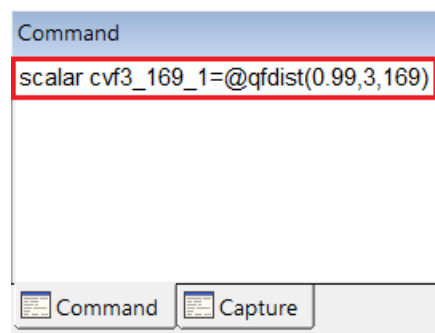
		Numerator Degrees of Freedom									
		1	2	3	4	5	6	7	8	9	10
D e n o m i n a t o r	10	10.04	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.94	4.85
	11	9.65	7.21	6.22	5.67	5.32	5.07	4.89	4.74	4.63	4.54
	12	9.33	6.93	5.95	5.41	5.06	4.82	4.64	4.50	4.39	4.30
	13	9.07	6.70	5.74	5.21	4.86	4.62	4.44	4.30	4.19	4.10
	14	8.86	6.51	5.56	5.04	4.69	4.46	4.28	4.14	4.03	3.94
	15	8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.89	3.80
	16	8.53	6.23	5.29	4.77	4.44	4.20	4.03	3.89	3.78	3.69
	17	8.40	6.11	5.18	4.67	4.34	4.10	3.93	3.79	3.68	3.59
	18	8.29	6.01	5.09	4.58	4.25	4.01	3.84	3.71	3.60	3.51
	19	8.18	5.93	5.01	4.50	4.17	3.94	3.77	3.63	3.52	3.43
	20	8.10	5.85	4.94	4.43	4.10	3.87	3.70	3.56	3.46	3.37
	21	8.02	5.78	4.87	4.37	4.04	3.81	3.64	3.51	3.40	3.31
	22	7.95	5.72	4.82	4.31	3.99	3.76	3.59	3.45	3.35	3.26
	23	7.88	5.66	4.76	4.26	3.94	3.71	3.54	3.41	3.30	3.21
	24	7.82	5.61	4.72	4.22	3.90	3.67	3.50	3.36	3.26	3.17
	25	7.77	5.57	4.68	4.18	3.85	3.63	3.46	3.32	3.22	3.13
	26	7.72	5.53	4.64	4.14	3.82	3.59	3.42	3.29	3.18	3.09
	27	7.68	5.49	4.60	4.11	3.78	3.56	3.39	3.26	3.15	3.06
	28	7.64	5.45	4.57	4.07	3.75	3.53	3.36	3.23	3.12	3.03
	29	7.60	5.42	4.54	4.04	3.73	3.50	3.33	3.20	3.09	3.00
D e g r e e s o f F r e e d o m	30	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.07	2.98
	40	7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	2.89	2.80
	60	7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.72	2.63
	90	6.93	4.85	4.01	3.54	3.23	3.01	2.84	2.72	2.61	2.52
	120	6.85	4.79	3.95	3.48	3.17	2.96	2.79	2.66	2.56	2.47
∞		6.63	4.61	3.78	3.32	3.02	2.80	2.64	2.51	2.41	2.32

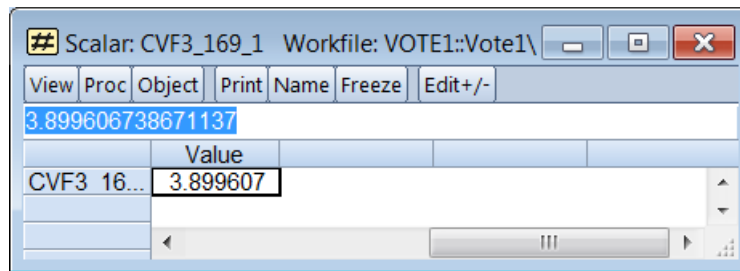
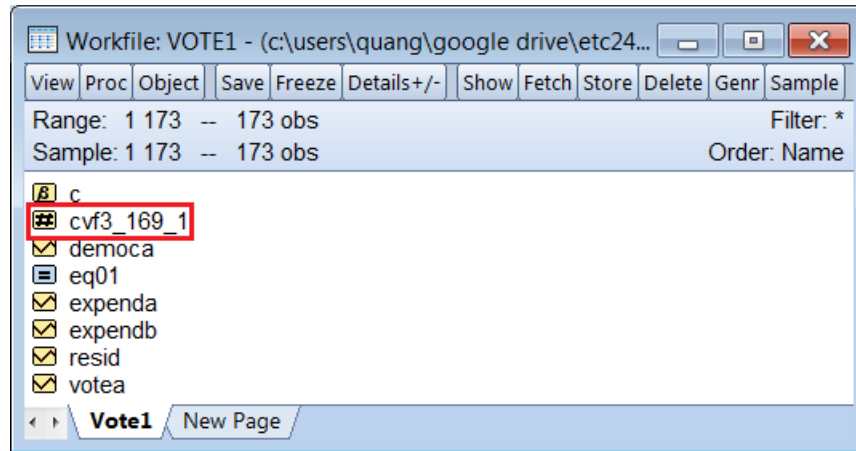
Example: The 1% critical value for numerator $df = 3$ and denominator $df = 60$ is 4.13.

Source: This table was generated using the Stata® function invFtail.

To obtain the critical value using EViews,

Command window : $\text{scalar cvf3_169_1} = @qfdist(0.99, 3, 169)$





$$F_{crit} \text{ (from Stat Table)} = 3.95$$

$$F_{crit} \text{ (from EViews)} = 3.8996$$

Rejection rule:

Comparing the calculated test statistic with the critical value, we reject H_0 if,

$$F_{calc} > F_{crit}$$

Conclusion

Since $F_{calc} = 207.38 > F_{crit} = 3.95$, we reject the null at the 1% significance level and conclude that at least one of the regressors in the model is statistically significant in explaining the share of votes received by Candidate A.

(c) Test of significance of an explanatory variable.

Test the hypothesis that controlling for campaign expenditure, being a democratic candidate is not significant in predicting the % of vote received in competitive races at the 5% level of significance. Perform that test by two methods: (i) comparing the t statistic with an appropriate critical value, and (ii) using the p-value.

If after controlling for campaign expenditure, being a democratic candidate is not significant in predicting the % of vote received then,

$$votea = \beta_0 + \beta_1 \log(expenda) + \beta_2 \log(expendb) + \cancel{\beta_3 \text{democa}} + u$$

that is,

$$\beta_3 = 0$$

but if it does then,

$$\beta_3 \neq 0$$

State the null and alternative hypothesis

$$H_0 : \beta_3 = 0$$

$$H_1 : \beta_3 \neq 0$$

The test statistic and its distribution under H_0

$$t = \frac{\hat{\beta}_3 - \beta_3}{se(\hat{\beta}_3)} = \frac{\hat{\beta}_3}{se(\hat{\beta}_3)} \sim t_{n-k-1} \quad \text{under } H_0$$

$$n = \text{sample size} = 173$$

$$k = \text{number of regressors in the model} = 3$$

Calculate the test statistic

$$t_{calc} = \frac{1.208824}{1.241612} = 0.9736$$

Dependent Variable: VOTEA
Method: Least Squares
Date: 08/19/17 Time: 09:07
Sample: 1 173
Included observations: 173

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	51.13410	2.903327	17.61224	0.0000
LOG(EXPENDA)	6.299279	0.375274	16.78582	0.0000
LOG(EXPENDB)	-6.666045	0.391187	-17.04054	0.0000
DEMOCA	1.208824	1.241612	0.973593	0.3317
R-squared	0.786385	Mean dependent var	50.50289	
Adjusted R-squared	0.782593	S.D. dependent var	16.78476	
S.E. of regression	7.826209	Akaike info criterion	6.975683	
Sum squared resid	10351.17	Schwarz criterion	7.048592	
Log likelihood	-599.3966	Hannan-Quinn criter.	7.005262	
F-statistic	207.3816	Durbin-Watson stat	1.652138	
Prob(F-statistic)	0.000000			

Note: The t - *Statistics* from the EViews regression output are t_{calc} values for a two-sided t-test to test if a regressor has a statistically significant on the dependent variable, holding the other regressors constant.

p-value from regression output

Dependent Variable: VOTEA
Method: Least Squares
Date: 08/19/17 Time: 09:07
Sample: 1 173
Included observations: 173

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	51.13410	2.903327	17.61224	0.0000
LOG(EXPENDA)	6.299279	0.375274	16.78582	0.0000
LOG(EXPENDB)	-6.666045	0.391187	-17.04054	0.0000
DEMOCA	1.208824	1.241612	0.973593	0.3317
R-squared	0.786385	Mean dependent var	50.50289	
Adjusted R-squared	0.782593	S.D. dependent var	16.78476	
S.E. of regression	7.826209	Akaike info criterion	6.975683	
Sum squared resid	10351.17	Schwarz criterion	7.048592	
Log likelihood	-599.3966	Hannan-Quinn criter.	7.005262	
F-statistic	207.3816	Durbin-Watson stat	1.652138	
Prob(F-statistic)	0.000000			

$$p - value = 0.3317$$

Note: The *Prob.* values from the EViews regression output are p-values for a two-sided t-test to test if a regressor is statistically significant, holding the other regressors constant.

Critical value and rejection region

5% *significance level* $\rightarrow \alpha = 0.05$

To obtain the critical value using the Stats Table, locate the t distribution table,

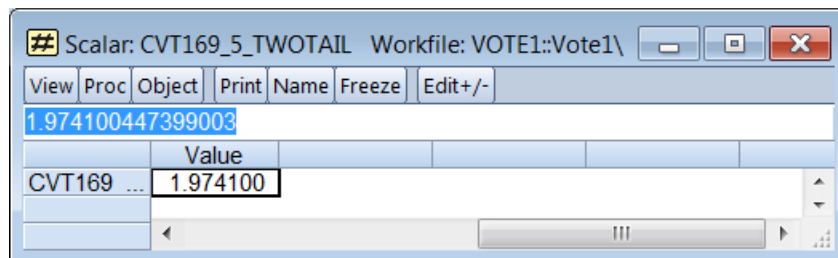
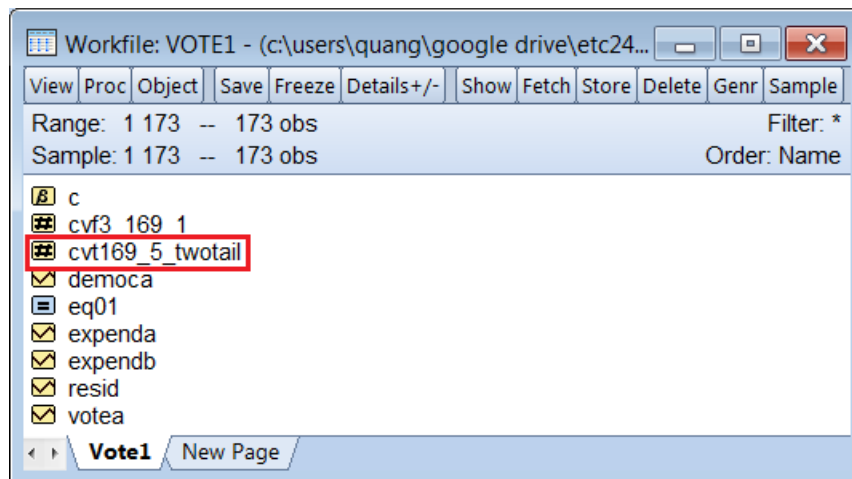
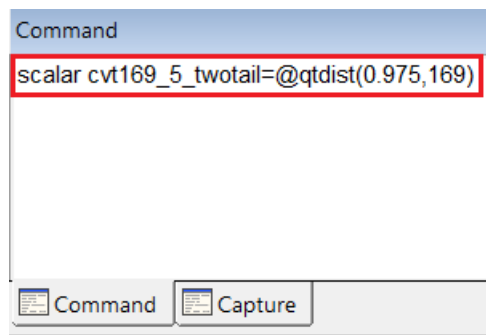
$$degrees\ of\ freedom = 169$$

Since 169 is not in the table, we take a conservative approach and choose the closest available degrees of freedom less than 169 i.e. $d.o.f = 120$.

TABLE G.2 Critical Values of the <i>t</i> Distribution						
		Significance Level				
1-Tailed:		.10	.05	.025	.01	.005
2-Tailed:		.20	.10	.05	.02	.01
D e g r e e s o f F r e e d o m	1	3.078	6.314	12.706	31.821	63.657
	2	1.886	2.920	4.303	6.965	9.925
	3	1.638	2.353	3.182	4.541	5.841
	4	1.533	2.132	2.776	3.747	4.604
	5	1.476	2.015	2.571	3.365	4.032
	6	1.440	1.943	2.447	3.143	3.707
	7	1.415	1.895	2.365	2.998	3.499
	8	1.397	1.860	2.306	2.896	3.355
	9	1.383	1.833	2.262	2.821	3.250
	10	1.372	1.812	2.228	2.764	3.169
	11	1.363	1.796	2.201	2.718	3.106
	12	1.356	1.782	2.179	2.681	3.055
	13	1.350	1.771	2.160	2.650	3.012
	14	1.345	1.761	2.145	2.624	2.977
	15	1.341	1.753	2.131	2.602	2.947
	16	1.337	1.746	2.120	2.583	2.921
	17	1.333	1.740	2.110	2.567	2.898
	18	1.330	1.734	2.101	2.552	2.878
	19	1.328	1.729	2.093	2.539	2.861
	20	1.325	1.725	2.086	2.528	2.845
	21	1.323	1.721	2.080	2.518	2.831
	22	1.321	1.717	2.074	2.508	2.819
	23	1.319	1.714	2.069	2.500	2.807
	24	1.318	1.711	2.064	2.492	2.797
	25	1.316	1.708	2.060	2.485	2.787
	26	1.315	1.706	2.056	2.479	2.779
	27	1.314	1.703	2.052	2.473	2.771
	28	1.313	1.701	2.048	2.467	2.763
	29	1.311	1.699	2.045	2.462	2.756
	30	1.310	1.697	2.042	2.457	2.750
	40	1.303	1.684	2.021	2.423	2.704
	60	1.296	1.671	2.000	2.390	2.660
	90	1.291	1.662	1.987	2.368	2.632
	120	1.289	1.658	1.980	2.358	2.617
	∞	1.282	1.645	1.960	2.326	2.576

To obtain the critical value using EViews,

Command window : `scalar cvt169_5_twotail = @qtdist(0.975,169)`



From Stat Table:

$$+t_{crit} = 1.980$$

$$-t_{crit} = -1.980$$

From EViews:

$$+t_{crit} = 1.9751$$

$$-t_{crit} = -1.9751$$

Rejection rule:

Comparing the calculated test statistic with the critical value, we reject H_0 if,

$$t_{calc} > +t_{crit}$$

or

$$t_{calc} < -t_{crit}$$

Comparing the p-value with the significance level, we reject H_0 if,

$$p - value < \alpha = 0.05$$

Conclusion

Since $p - value = 0.3317 > \alpha = 0.05$, we do not reject the null at the 5% significance level and conclude that there is insufficient evidence from our sample to suggest that Candidate A's political affiliation is statistically significant in explaining the share of votes received by Candidate A, holding campaign expenditure constant.

(d) Joint test of multiple linear restrictions.

Test the joint hypothesis that controlling for campaign expenditure, being a democratic candidate does not contribute to the % vote received and that the effect (on *votea*) of every percentage increase in campaign expenditure by Candidate A can be offset exactly by the same percentage increase in the opponent's campaign expenditure. Perform this test at the 5% level of significance.

From the statement,

“..controlling for campaign expenditure, being a democratic candidate does not contribute to the % of vote received..”

we have our first linear restriction,

$$\beta_3 = 0$$

From the statement,

“..the effect (on *votea*) of every percentage increase in campaign expenditure by Candidate A can be offset exactly by the same percentage increase in the opponent's campaign expenditure..”

we have our second linear restriction,

$$\beta_1 + \beta_2 = 0$$

$$\beta_2 = -\beta_1$$

It tells us that the effect on *votea* when both Candidate's campaign expenditure increases by the same percentage equals to 0.

Unrestricted model (the model before imposing restrictions):

$$votea = \beta_0 + \beta_1 \log(expenda) + \beta_2 \log(expendb) + \beta_3 democa + u$$

Restricted model (the model after imposing restrictions):

$$votea = \beta_0 + \beta_1 \log(expenda) - \beta_1 \log(expendb) + 0 \times democa + u$$
$$votea = \beta_0 + \beta_1 (\log(expenda) - \log(expendb)) + u$$

State the null and alternative hypothesis

$$H_0 : \beta_3 = 0, \beta_2 = -\beta_1$$

$$H_1 : \beta_3 \neq 0 \text{ and/or } \beta_2 \neq -\beta_1$$

The test statistic and its distribution under H_0

$$F = \frac{(SSR_r - SSR_{ur})/q}{SSR_{ur}/(n - k - 1)} = \frac{(SSR_r - SSR_{ur})/2}{SSR_{ur}/(173 - 3 - 1)} \sim F_{q, n-k-1} \quad \text{under } H_0$$

$$n = \text{sample size} = 173$$

$$k = \text{number of regressors in the unrestricted model} = 3$$

$$q = \text{number of restrictions} = 2$$

$$SSR_r = \text{sum of squared residuals from estimated restricted model}$$

$$SSR_{ur} = \text{sum of squared residuals from estimated unrestricted model}$$

Calculate the test statistic

From the regression output of the estimated unrestricted model,

$$SSR_{ur} = 10351.17$$

Dependent Variable: VOTEA
Method: Least Squares
Date: 08/19/17 Time: 09:07
Sample: 1 173
Included observations: 173

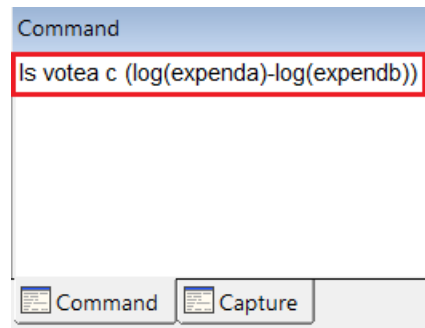
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	51.13410	2.903327	17.61224	0.0000
LOG(EXPENDA)	6.299279	0.375274	16.78582	0.0000
LOG(EXPENDB)	-6.666045	0.391187	-17.04054	0.0000
DEMOCA	1.208824	1.241612	0.973593	0.3317
R-squared	0.786385	Mean dependent var	50.50289	
Adjusted R-squared	0.782593	S.D. dependent var	16.78476	
S.E. of regression	7.826209	Akaike info criterion	6.975683	
Sum squared resid	10351.17	Schwarz criterion	7.048592	
Log likelihood	-599.3966	Hannan-Quinn criter.	7.005262	
F-statistic	207.3816	Durbin-Watson stat	1.652138	
Prob(F-statistic)	0.000000			

To obtain SSR_r , we need to estimate the restricted model,

$$votea = \beta_0 + \beta_1(\log(expenda) - \log(expendb)) + u$$

To estimate this model from the Command window,

ls votea c (log(expenda)-log(expendb))

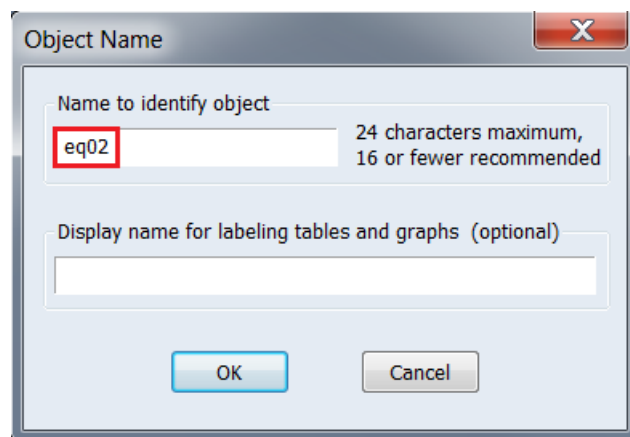
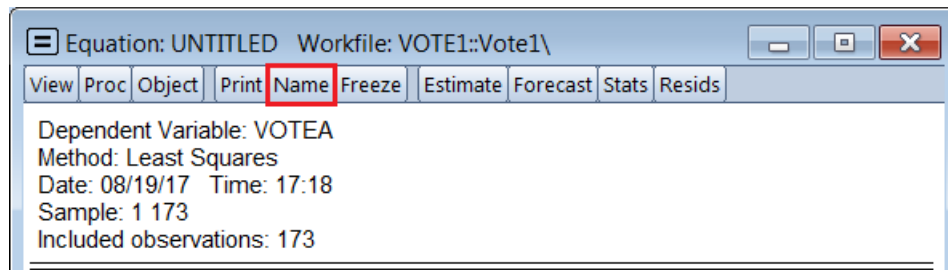


(Press Enter to execute code)

To name (save) the estimated equation,

Name → *Name to identify object* : eq02

*(This names the equation **eq02**)*



Dependent Variable: VOTEA
Method: Least Squares
Date: 08/19/17 Time: 17:18
Sample: 1 173
Included observations: 173

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	49.97148	0.594598	84.04253	0.0000
LOG(EXPENDA)-LOG(EXPENDB)	6.545465	0.262391	24.94545	0.0000
R-squared	0.784438	Mean dependent var	50.50289	
Adjusted R-squared	0.783178	S.D. dependent var	16.78476	
S.E. of regression	7.815690	Akaike info criterion	6.961637	
Sum squared resid	10445.54	Schwarz criterion	6.998091	
Log likelihood	-600.1816	Hannan-Quinn criter.	6.976426	
F-statistic	622.2757	Durbin-Watson stat	1.657779	
Prob(F-statistic)	0.000000			

Table 3: Regression output of *votea* on a constant and $\log(\text{expenda}) - \log(\text{expendb})$.

$$SSR_r = 10445.54$$

$$F_{calc} = \frac{(SSR_r - SSR_{ur})/2}{SSR_{ur}/(173 - 3 - 1)} = \frac{(10445.54 - 10351.17)/2}{10351.17/(169)} = 0.77$$

Critical value and rejection region

5% *significance level* $\rightarrow \alpha = 0.05$

To obtain the critical value using the Stats Table, locate the F distribution table at the 5% significance level,

$$\text{Numerator } d.o.f = 2$$

$$\text{Denominator } d.o.f = 169$$

Since 169 is not in the table, we take a conservative approach and choose the closest available degrees of freedom less than 169 i.e. $d.o.f = 120$.

TABLE G.3b 5% Critical Values of the F Distribution

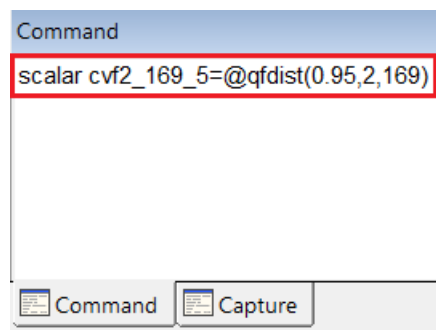
		Numerator Degrees of Freedom									
		1	2	3	4	5	6	7	8	9	10
D e n o m i n a t o r D e g r e e s o f F r e e d o m	10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98
	11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85
	12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75
	13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67
	14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60
	15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54
	16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49
	17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45
	18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41
	19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38
	20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35
	21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32
	22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30
	23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32	2.27
	24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25
	25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28	2.24
	26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27	2.22
	27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25	2.20
	28	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24	2.19
	29	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22	2.18
	30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16
	40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08
	60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04	1.99
	90	3.95	3.10	2.71	2.47	2.32	2.20	2.11	2.04	1.99	1.94
	120	3.92	3.07	2.68	2.45	2.29	2.17	2.09	2.02	1.96	1.91
	∞	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88	1.83

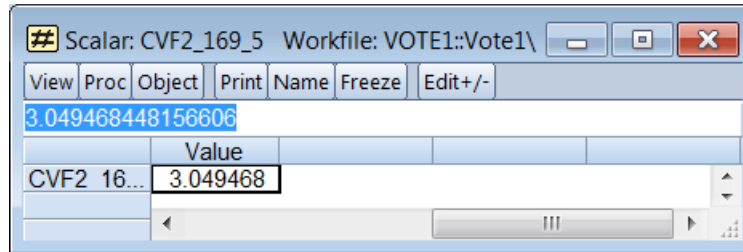
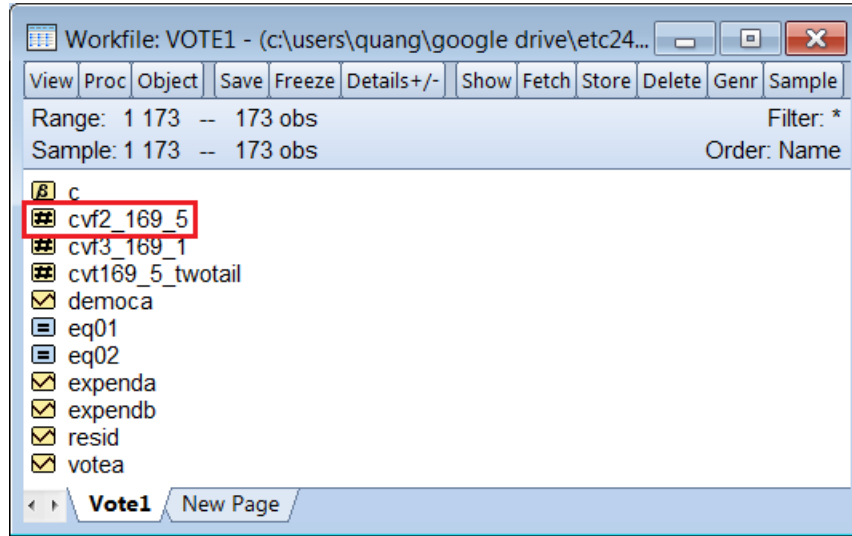
Example: The 5% critical value for numerator $df = 4$ and large denominator $df(\infty)$ is 2.37.

Source: This table was generated using the Stata® function invFtail.

To obtain the critical value using EViews,

Command window : `scalar cvf2_169_5 = @qfdist(0.95,2,169)`





$$F_{crit} \text{ (from Stat Table)} = 3.07$$

$$F_{crit} \text{ (from EViews)} = 3.049$$

Rejection rule:

Comparing the calculated test statistic with the critical value, we reject H_0 if,

$$F_{calc} > F_{crit}$$

Conclusion

Since $F_{calc} = 0.77 < F_{crit} = 3.95$, we do not reject the null at the 5% significance level and conclude that there is insufficient evidence to reject the assumption that being a democratic candidate has no impact on share of votes holding campaign expenditure constant **and** that the effect on share of votes received by Candidate A from a percentage increase in Candidate A's campaign expenditure can be offset exactly by the same percentage increase in Candidate B's campaign expenditure.

(e) Testing a single hypothesis about a linear combination of parameters.

Drop *democa* from the model.

$$votea = \beta_0 + \beta_1 \log(expenda) + \beta_2 \log(expendb) + u$$

In close races each candidate believes that he or she needs to increase their campaign expenditure by more than 1% to offset the effect of a 1% increase in their opponent's expenditure. The null hypothesis is,

$$H_0 : \beta_1 + \beta_2 = 0$$

and although it involves two parameters, it tests only one restriction. The alternative is,

$$H_1 : \beta_1 + \beta_2 < 0$$

That is, the impact on the share of votes received by Candidate A for a 1% increase in Candidate A's campaign expenditure is not enough to offset the impact from a 1% increase in Candidate B's campaign expenditure on the share of votes received by Candidate A. Candidate A needs to increase their campaign expenditure by more than 1% increase to offset the effect of a 1% increase in Candidate B's campaign expenditure.

so we cannot use the F test because F test provides inference against $\beta_1 + \beta_2 \neq 0$. In such cases that we have only one restriction about a linear combination, we use a reparameterisation trick:

$$\text{Define } \delta = \beta_1 + \beta_2 \implies \beta_2 = \delta - \beta_1$$

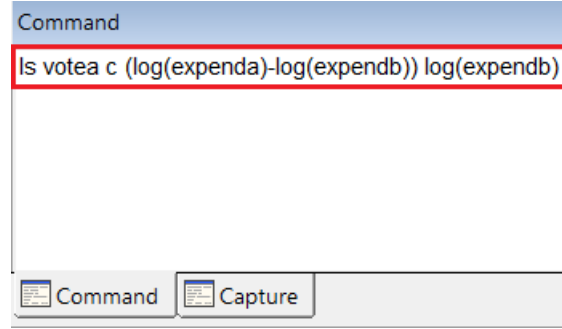
Substitute for β_2 in the population model and re-arrange, you will see that δ becomes the coefficient of one of the explanatory variables in the reparameterised model. You can see that testing $\delta = 0$ against $\delta < 0$ can be performed with a simple t test in this reparameterised model.

Reparameterised model

$$\begin{aligned} votea &= \beta_0 + \beta_1 \log(expenda) + \beta_2 \log(expendb) + u \\ &= \beta_0 + \beta_1 \log(expenda) + (\delta - \beta_1) \log(expendb) + u \\ &= \beta_0 + \beta_1 \log(expenda) + \delta \log(expendb) - \beta_1 \log(expendb) + u \\ &= \beta_0 + \beta_1 (\log(expenda) - \log(expendb)) + \delta \log(expendb) + u \end{aligned}$$

To estimate the reparameterised model from the Command window in EViews,

$$ls \text{ votea } c (\log(\text{expenda}) - \log(\text{expendb})) \log(\text{expendb})$$



(Press Enter to execute code)

Dependent Variable: VOTEA
Method: Least Squares
Date: 08/19/17 Time: 18:42
Sample: 1 173
Included observations: 173

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	52.03893	2.750137	18.92231	0.0000
LOG(EXPENDA)-LOG(EXPENDB)	6.341950	0.372649	17.01858	0.0000
LOG(EXPENDB)	-0.414801	0.538688	-0.770019	0.4424
R-squared	0.785187	Mean dependent var	50.50289	
Adjusted R-squared	0.782660	S.D. dependent var	16.78476	
S.E. of regression	7.825009	Akaike info criterion	6.969716	
Sum squared resid	10409.23	Schwarz criterion	7.024397	
Log likelihood	-599.8804	Hannan-Quinn criter.	6.991900	
F-statistic	310.6936	Durbin-Watson stat	1.662940	
Prob(F-statistic)	0.000000			

Table 4: Regression output of *votea* on a constant, $\log(\text{expenda}) - \log(\text{expendb})$ and $\log(\text{expendb})$.

$$\widehat{votea} = \underset{(se(\hat{\beta}_0))}{\hat{\beta}_0} + \underset{(se(\hat{\beta}_1))}{\hat{\beta}_1} (\log(expenda) - \log(expendb)) + \underset{(se(\hat{\delta}))}{\hat{\delta}} \log(expendb)$$

$$\widehat{votea} = 52.0389 + 6.3420(\log(expenda) - \log(expendb)) - 0.4148\log(expendb)$$

(2.7501)
 (0.3726)
 (0.5387)

State the null and alternative hypothesis

$$H_0 : \delta = 0 \quad (\text{same as } \beta_1 + \beta_2 = 0)$$

$$H_1 : \delta < 0 \quad (\text{same as } \beta_1 + \beta_2 < 0)$$

(one – sided t test)

The test statistic and its distribution under H_0

$$t = \frac{\hat{\delta} - \delta}{se(\hat{\delta})} = \frac{\hat{\delta}}{se(\hat{\delta})} \sim t_{n-k-1} \quad \text{under } H_0$$

$$n = \text{sample size} = 173$$

$$k = \text{number of regressors in the model} = 2$$

Calculate the test statistic

$$t_{calc} = \frac{-0.414801}{0.538688} = -0.77$$

p-value

$$p - \text{value} = \frac{0.4424}{2} = 0.2212$$

Note: The *Prob.* values from the EViews regression output are p-values for a two-sided t-tests of individual significance so to calculate the p-value for our one-sided t-test, we divided the two-sided p-value by 2.

Dependent Variable: VOTEA
Method: Least Squares
Date: 08/19/17 Time: 18:42
Sample: 1 173
Included observations: 173

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	52.03893	2.750137	18.92231	0.0000
LOG(EXPENDA)-LOG(EXPENDB)	6.341950	0.372649	17.01858	0.0000
LOG(EXPENDB)	-0.414801	0.538688	-0.770019	0.4424
R-squared	0.785187	Mean dependent var	50.50289	
Adjusted R-squared	0.782660	S.D. dependent var	16.78476	
S.E. of regression	7.825009	Akaike info criterion	6.969716	
Sum squared resid	10409.23	Schwarz criterion	7.024397	
Log likelihood	-599.8804	Hannan-Quinn criter.	6.991900	
F-statistic	310.6936	Durbin-Watson stat	1.662940	
Prob(F-statistic)	0.000000			

Critical value and rejection region

5% *significance level* $\rightarrow \alpha = 0.05$

To obtain the critical value using the Stats Table, locate the t distribution table,

$$\text{degrees of freedom} = 170$$

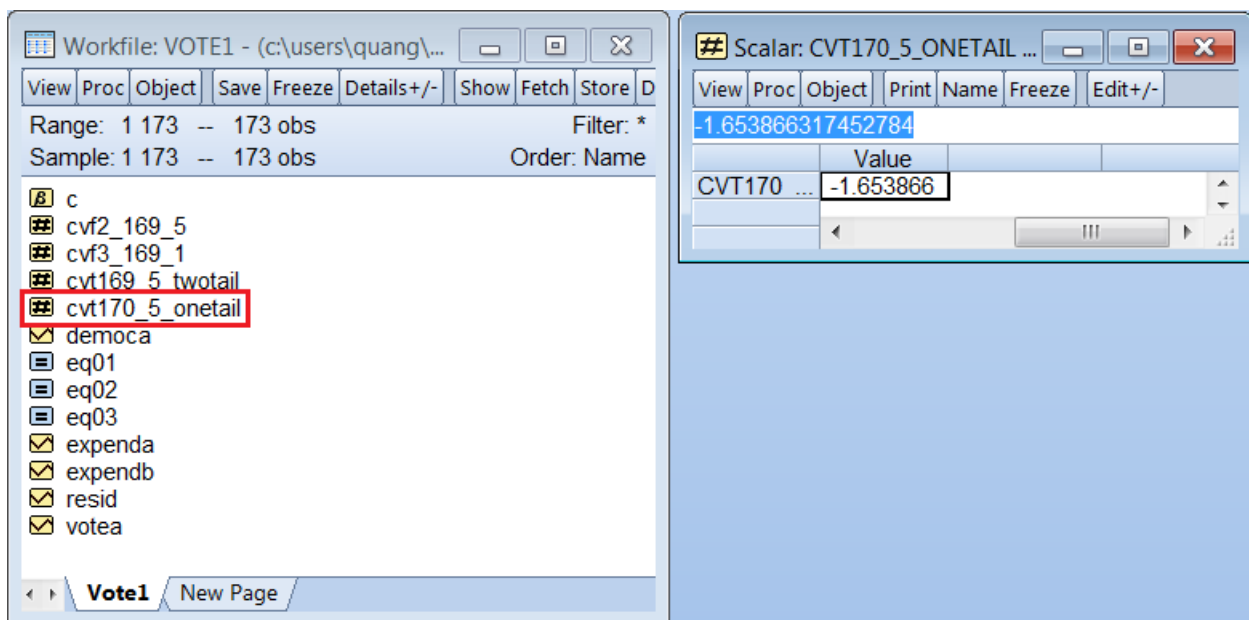
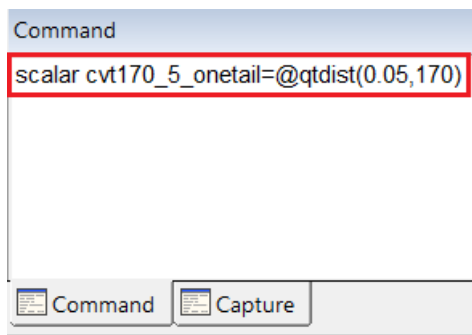
Since 170 is not in the table, we take a conservative approach and choose the closest available degrees of freedom less than 170 i.e. $d.o.f = 120$.

TABLE G.2 Critical Values of the <i>t</i> Distribution						
		Significance Level				
1-Tailed:	.10	.05	.025	.01	.005	
2-Tailed:	.20	.10	.05	.02	.01	
1	3.078	6.314	12.706	31.821	63.657	
2	1.886	2.920	4.303	6.965	9.925	
3	1.638	2.353	3.182	4.541	5.841	
4	1.533	2.132	2.776	3.747	4.604	
5	1.476	2.015	2.571	3.365	4.032	
6	1.440	1.943	2.447	3.143	3.707	
7	1.415	1.895	2.365	2.998	3.499	
8	1.397	1.860	2.306	2.896	3.355	
9	1.383	1.833	2.262	2.821	3.250	
10	1.372	1.812	2.228	2.764	3.169	
11	1.363	1.796	2.201	2.718	3.106	
12	1.356	1.782	2.179	2.681	3.055	
13	1.350	1.771	2.160	2.650	3.012	
14	1.345	1.761	2.145	2.624	2.977	
15	1.341	1.753	2.131	2.602	2.947	
16	1.337	1.746	2.120	2.583	2.921	
17	1.333	1.740	2.110	2.567	2.898	
18	1.330	1.734	2.101	2.552	2.878	
19	1.328	1.729	2.093	2.539	2.861	
20	1.325	1.725	2.086	2.528	2.845	
21	1.323	1.721	2.080	2.518	2.831	
22	1.321	1.717	2.074	2.508	2.819	
23	1.319	1.714	2.069	2.500	2.807	
24	1.318	1.711	2.064	2.492	2.797	
25	1.316	1.708	2.060	2.485	2.787	
26	1.315	1.706	2.056	2.479	2.779	
27	1.314	1.703	2.052	2.473	2.771	
28	1.313	1.701	2.048	2.467	2.763	
29	1.311	1.699	2.045	2.462	2.756	
30	1.310	1.697	2.042	2.457	2.750	
40	1.303	1.684	2.021	2.423	2.704	
60	1.296	1.671	2.000	2.390	2.660	
90	1.291	1.662	1.987	2.368	2.632	
120	1.289	1.658	1.980	2.358	2.617	
∞	1.282	1.645	1.960	2.326	2.576	

Note : This gives us $+t_{crit}$ but we need $-t_{crit}$ for this test

To obtain the critical value using EViews,

Command window : `scalar cvt170_5_onetail = @qtdist(0.05, 170)`



From Stat Table:

$$-t_{crit} = -1.658$$

From EViews:

$$-t_{crit} = -1.6539$$

Rejection rule:

Comparing the calculated test statistic with the critical value, we reject H_0 if,

$$t_{calc} < -t_{crit}$$

Comparing the p-value with the significance level, we reject H_0 if,

$$p - value < \alpha = 0.05$$

Conclusion

Since $t_{calc} = -0.77 > -t_{crit} = -1.658$, we do not reject the null at the 5% significance level and conclude that there is insufficient evidence from our sample to suggest that Candidate A needs a larger than 1% increase in campaign expenditure to offset the effect of a 1% increase in Candidate B's campaign expenditure on share of votes received by Candidate A.