

The University of Tennessee, Knoxville
Econ 381: Introduction to Econometrics
Midterm 2
October 13, 2016

Do not open this booklet until instructed to do so.

Name:

Instructions and advice:

1. There is plenty of time. Take it, do not rush.
2. There are blank pages at the back for rough work, which you should use to sketch out answers. Write only your final answers on the answers pages.
3. The back page of this exam includes the a table of critical values for the t -distribution.
4. Including the bonus question, there are 110 points available on this exam. The maximum score is capped at 100 points. Note the points allocated to each question.
5. There is a big difference between 3 and -3, and there is a big difference between Y and \bar{Y} . Be careful with your answers.
6. There is enough room in the spaces provided for a perfect answer. Every question can be answered with a well-structured sentence or two; and you will not receive extra points for long-winded answers.
7. This is a closed book exam, and calculators capable of anything complicated are not allowed.
8. If you cheat, you will fail the course. Avoid the temptation to look into the work of the person next to you.
9. Best of luck.

Question 1 – House Prices

Some Stata output is shown below. It is a regression of houses prices (in thousands of US Dollars), on the number of bathrooms, the number of bedrooms, the size of the house (in square feet), and whether the house has central air-conditioning. Some information is suppressed.

```
. reg price baths beds house_size aircon
```

Source	SS	df	MS	Number of obs	=	43
Model	209701.46	4	52425.3649	F(4, 38)	=	36.87
Residual	54025.6101	38	1421.72658	Prob > F	=	0.0000
Total	263727.07	42	6279.21595	R-squared	=	0.7951
				Adj R-squared	=	0.7736
				Root MSE	=	37.706

price	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
baths	7.858424	20.25533			
beds	1.824019	13.42338			
house_size	.1272404	.0277299			
aircon	3.790653	17.35817			
_cons	36.1058	28.39124			

- a. After running this regression, you conduct a Breusch-Pagan test. Stata produces the following output.

```
Breusch-Pagan / Cook-Weisberg test for heteroskedasticity
Ho: Constant variance
Variables: baths beds house_size aircon

chi2(4)      =    4.52
Prob > chi2  =    0.3403
```

What command do you input into Stata to produce this test? Include any relevant options/sub-commands. [5 points]

- b. What is your response to this test, and what adjustments (if any) would you make as a result? Take your time to consider this, and explain your reasoning concisely. [10 points]

- c. Towards the top right-hand corner, Stata reports $F(4, 38) = 36.87$. What do the numbers 4 and 38 mean in this context? Be precise. [5 points]
- d. What is your conclusion on the null hypothesis from that F -test? Briefly explain. [5 points]
- e. After running this regression, you run the `vif` command. Stata produces the following output.

```
. vif
```

Variable	VIF	1/VIF
house_size	5.98	0.167223
baths	3.42	0.292170
beds	2.26	0.441773
aircon	1.63	0.614841
Mean VIF	3.32	

Let us use `baths` as an example. Write out the regression equation (either in math or in “Stata code”) needed to calculate the VIF for `baths`. [5 points]

- f. This regression will give us an R^2 . Write out the formula relating the VIF for `baths` to this R^2 . [5 points]
- g. In light of this output, how would you proceed in terms of model selection? [5 points]

Question 2 – Auditing Politicians

In class we discussed the Bobonis et al (2016) paper on corruption, published in a recent edition of the *American Economic Review*. Table 3 from that paper is reproduced below.

TABLE 3—EFFECTS OF THE (*Timing of the*) AUDITS
ON THE NUMBER OF CORRUPT VIOLATIONS IN THE CURRENT AUDIT

	Number of corrupt violations, per report			Share of findings classified as corrupt violations, findings	
	All OLS (1)	By mayor/ vice-mayor OLS (2)	Referred to DOJ OLS (3)	By mayor/ vice-mayor OLS (4)	Referred to DOJ OLS (5)
<i>Panel A. Average effects</i>					
Timely audit	-1.43 (0.22)	-0.63 (0.14)	-0.65 (0.16)	-0.082 (0.040)	-0.160 (0.055)
Municipality controls	Yes	Yes	Yes	Yes	Yes
Election year and municipality FEs	Yes	Yes	Yes	Yes	Yes
<i>Panel B. Effects by party advantage</i>					
Timely audit	-1.90 (0.32)	-0.92 (0.17)	-1.12 (0.25)	-0.174 (0.045)	-0.229 (0.069)
Timely audit × incumbent's party has won in previous 3+ elections	0.96 (0.41)	0.61 (0.24)	0.97 (0.32)	0.191 (0.061)	0.153 (0.106)
Municipality controls	Yes	Yes	Yes	Yes	Yes
Election year and municipality FEs	Yes	Yes	Yes	Yes	Yes
Observations	326	326	326	326	326
Mean of dep. variable (untimely audits)	2.17	0.95	0.99	0.22	0.34

- In the column labelled “All OLS”, the coefficient on **Timely Audit** is -1.90 with a standard error of 0.32. What is the t -statistic on a hypothesis that the true coefficient equals zero? [5 points]
- Do you reject or fail to reject that hypothesis? A good answer will make reference to the p -value implied by that t -statistic. [5 points]
- Construct a 95% confidence interval for the true value of that coefficient. (If you wish, you may make invoke a common approximation.) [10 points]

d. The coefficient for the interaction effect of `Timely audit × incumbent's party has won in previous 3+ elections` is 0.96. With appropriate reference to the base category, precisely explain how to interpret this interaction effect. [5 points]

e. Bonus question In one sentence, what is a “Municipality Fixed Effect”? (Hint: we discussed fixed effects in relation to the effect of education on wages.) [10 points]

Question 3 – German GDP

In class we discussed a German quarterly GDP dataset, some results of which are reproduced below. The output shows a regression of consumption based on income.

```
. reg consump inc
```

Source	SS	df	MS
Model	31752505.3	1	31752505.3
Residual	23846.8925	90	264.965472
Total	31776352.2	91	349190.684

consump	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
inc	.8451531	.0024414			
_cons	21.41795	3.718209			

- Suppose we have a problem with autocorrelation here. What effect (if any) will this have on our estimated β ? [5 points]
- A Durbin-Watson test indicates $DW = 0.5719$. Does this indicate positive autocorrelation, negative autocorrelation, or no autocorrelation? Explain. [5 points]
- Given that DW statistic, what effect (if any) will this have on our SE? Be precise. [5 points]
- Name one major theoretical drawback with using Durbin-Watson to test for autocorrelation. [5 points]

- e. After running this regression, what command would you use to run to conduct a Breusch-Godfrey test? [5 points]
- f. Suppose a Breusch-Godfrey test rejects a null of no serial correlation. What adjustment would you make to the regression shown above, and how would you implement this adjustment in Stata? [10 points]

Inference for two-sided tests

Critical Values for t distribution

df	<i>Significance Level</i>		
	0.9	0.95	0.99
1.	6.314	12.71	63.66
2.	2.920	4.303	9.925
3.	2.353	3.182	5.841
4.	2.132	2.776	4.604
5.	2.015	2.571	4.032
6.	1.943	2.447	3.707
7.	1.895	2.365	3.499
8.	1.860	2.306	3.355
9.	1.833	2.262	3.250
10.	1.812	2.228	3.169
11.	1.796	2.201	3.106
12.	1.782	2.179	3.055
13.	1.771	2.160	3.012
14.	1.761	2.145	2.977
15.	1.753	2.131	2.947
16.	1.746	2.120	2.921
17.	1.740	2.110	2.898
18.	1.734	2.101	2.878
19.	1.729	2.093	2.861
20.	1.725	2.086	2.845
21.	1.721	2.080	2.831
22.	1.717	2.074	2.819
23.	1.714	2.069	2.807
24.	1.711	2.064	2.797
25.	1.708	2.060	2.787
26.	1.706	2.056	2.779
27.	1.703	2.052	2.771
28.	1.701	2.048	2.763
29.	1.699	2.045	2.756
30.	1.697	2.042	2.750
40.	1.684	2.021	2.704
60.	1.671	2.000	2.660
100.	1.660	1.984	2.626
120.	1.658	1.980	2.617
∞	1.645	1.960	2.576

Source: nist.gov