Political Science 150B/350B
Final Exam
Winter 2008

- This is a **closed book** examination.
- Answer all questions. Show work for partial credit where appropriate. You will be rewarded for answers that are clear, correct, insightful, and brief. You will not be rewarded for answers that are rambling, incorrect, or confused.
- Clearly indicate which question you are answering; note that Question 2 has many individual parts.
- Be particularly careful with labeling your answers if you answer the questions out of order.
- The maximum score is printed after the last question.
- A table of critical quantiles of the χ^2 distribution appears after the exam questions.

Question 1: (5 points) After estimating the least squares regression of **y** on **X**, a researcher finds that the correlation between the regression's predicted values $\hat{\mathbf{y}} = \mathbf{H}\mathbf{y}$ and the residuals $\mathbf{y} - \hat{\mathbf{y}}$ is approximately zero. What can the researcher conclude? [Hint: no proof necessary].

Question 2: For each of the following questions, simply select the correct response. No explanation is necessary.

(a): (4 points) A regression that has many statistically insignificant predictors, but a relatively high r^2 is most likely to be suffering from...

(a) group-wise heteroskedasticity
(b) data points with high leverages
(c) data points with high influence
(d) data that are skewed right
(e) multicollinearity

(b): (4 points) We would use a Chow test to test for...

(a) heteroskedastic disturbances
(b) exogeneity of instruments
(c) structural stability in regression coefficients
(d) endogeneity of regressors
(e) autocorrelated disturbances

(c): (4 points) We would use a Durbin-Wu-Hausmen test to test for...

(a) heteroskedastic disturbances
(b) exogeneity of instruments
(c) structural stability in regression coefficients
(d) endogeneity of regressors
(e) autocorrelated disturbances

(d): (4 points) We would use a Breusch-Pagan test to test for...

(a) heteroskedastic disturbances
(b) exogeneity of instruments
(c) omitted variable bias
(d) endogeneity of regressors
(e) autocorrelated disturbances

(e): (4 points) A researcher reporting the results of a regression analysis says that they have used a "HCCM". This is because the researcher is concerned about...

(a) regressors measured with error
(b) heteroskedastic disturbances
(c) exogeneity of instruments
(d) omitted variable bias
(e) endogeneity of regressors

(f): (4 points) A researcher conducts a regression analysis with all **X** variables centered around their respective averages. This implies that...

(a) all the estimated slope coefficients in the regression model will be equal to 1.0
(b) that the estimated intercept will be equal to the mean of y.
(c) the intercept in the regression equation will be zero, and so can be dropped from the model
(d) the estimated coefficients will have smaller standard errors than if the model was estimated with the uncentered variables
(e) none of the above

(g): (4 points) Complete the 2nd sentence. A researcher regresses y on a single predictor, X, and a constant. The r^2 from this regression is equal to...

(h): (4 points) Complete the sentence: "An estimator $\hat{\theta}$ whose sampling distribution (un-normalized by sample size) is asymptotically degenerate with point mass on θ is ..."

(i): (4 points) A researcher runs the regression of a survey-based measure of public opinion in region $i = 1, \ldots, n$, on various demographic characteristics of those regions. The surveys conducted in each region vary considerably in terms of the sample sizes that are used. The resulting regression is guaranteed to have...

(a) spatially correlated disturbances
(b) regressors measured with error
(c) highly influential observations
(d) heteroskedastic disturbances

(j): (4 points) In the presence of disturbances that are not "iid", the OLS estimator of $\boldsymbol{\beta}$ is generally...

(a) unbiased and consistent
(b) the best linear unbiased estimator
(c) biased but consistent
(d) inconsistent

Question 3: In the following question, consider data y_{ij} and predictors $\mathbf{x}_{ij}$ where $j = 1, \ldots, m_i$ indexes observations in geographic regions $i = 1, \ldots, n$. The researcher estimates a regression of y_{it} on predictors $\mathbf{x}_{it}$, including a vector of binary indicator (or "dummy") variables $\mathbf{d}_{ij}$, each one coded 1 if the particular observation is from region i and coded 0 otherwise.

(a): (2 points) How many dummy variables appear in the estimated regression? [Hint: there is more than one correct answer here; be clear].

(b): (3 points) What name is conventionally given to the *coefficients* that attach to the region-specific dummy variables?

(c): (3 points) When we estimate a model such as the one given above, what property do the residuals have, region-by-region?

(d): (4 points) The presence of these dummy variables in the model helps guard against what possible problem in the model specification? Hint: what does a regression look like without the region-by-region dummy variables, versus the regression that includes the **d** dummy variables?

Question 4: A researcher conducting a regression analysis believes that at least one of her regressors is endogenous, and seeks to remedy the situation using instrumental variables. She seeks your advice regarding the following issues. Answer merely by selecting the correct response.

(a): (4 points) In the presence of an endogenous regressor, the least squares estimates of the parameters of the structural equation of substantive interest are...

(a) biased, but consistent
(b) unbiased and consistent
(c) biased and inconsistent
(d) unbiased, but inconsistent

(b): (4 points) Complete the sentence. The instrumental variables estimator of the parameters in the structural equation of substantive interest is just identified when...

(c): (4 points) Let the structural model of substantive interest be $\mathbf{y} = \mathbf{X}_1\boldsymbol{\beta}_1 + \mathbf{X}_2\boldsymbol{\beta}_2 + \boldsymbol{\varepsilon}$, where $\mathbf{X}_1$ are endogneous regressors and $\mathbf{X}_2$ are exogenous regressors. Let $\mathbf{Z}$ be a matrix of exogenous regressors. Then the "first stage" of the two-stage least squares estimator consists of regressing each endogenous regressor $\mathbf{x}_j$, where j indexes the columns of $\mathbf{X}_1$, on...

(a) $\mathbf{Z}$
(b) $\mathbf{X}_2$ and $\mathbf{Z}$
(c) $\mathbf{X}_2$, $\mathbf{Z}$ and $\mathbf{x}_k$, $\forall\ k \neq j$.
(d) any of the above.

(d): (4 points) Using the notation in the previous question, the guiding principles for choosing $\mathbf{Z}$ variables are...

(a) $\text{plim}\frac{1}{n}(\mathbf{Z}'\boldsymbol{\varepsilon}) = \mathbf{0}$.
(b) $\text{plim}\frac{1}{n}(\mathbf{Z}'\mathbf{X}_1)$ be large.
(c) in order of importance, (a), then (b)
(d) in order of importance, (b), then (a)

(e): (4 points) Again using the notation from above, the weakness of an instrument (or collection of instruments) is best assessed by examining...

(a) the correlations between $\mathbf{y}$ and the columns of $\mathbf{Z}$
(b) the correlations between the columns of $\mathbf{X}_1$ and the columns of $\mathbf{Z}$
(c) the correlations between the columns of $\mathbf{X}_2$ and the columns of $\mathbf{Z}$
(d) the partial correlations between the columns of $\mathbf{X}_1$ and the columns of $\mathbf{Z}$ controlling for $\mathbf{X}_2$
(e) any of (b), (c), or (d)
(f) none of the above

Question 5: Professor Shanto Iyengar and I (Jackman) have data assessing reactions to political advertising on television. Each subject ($i = 1, \ldots, n$) is shown a 30 second ad, and is given a "dial", a device that the respondent turns to the left or the right to indicate "disliking" or "liking" what they are viewing as the ad progresses. The dial generates readings Y_{it} on a 0 - 100 scale every

second ($t = 0, \ldots, 30$), which are captured and stored on a computer. The dial is initialized at $Y_{i0} = 50 \, \forall \, i$. Iyengar and Jackman considered fitting the model

$$E(Y_{it}|t) = \alpha \exp\left[-\gamma \exp(-\beta t)\right]$$

to the data, where α, γ and β are unknown parameters to be estimated.

(a): (5 points) The dials data are constrained to start at 50 at $t = 0, \forall \, i$ and to lie in the 0-100 interval. What might this imply about the disturbances Y_{it} - $E(Y_{it}|t)$ of the model given above?

(b): (5 points) Since $Y_{it}|(t = 0) = 50 \, \forall \, i$, what constraint is implied on one or more of the parameters in the model?

(c): (5 points) Irrespective of the answers to the previous questions, can ordinary least squares be used to estimate the parameters of the proposed model? Explain your answer.

Question 6: Suppose we have data $(y_i, \mathbf{x}_i)$, $i = 1, \ldots, n$, for which we posit the probability model $y_i|\mathbf{x}_i \overset{\text{iid}}{\sim} N(\mathbf{x}_i\boldsymbol{\beta}, \sigma^2)$, where $\boldsymbol{\beta}$ is a vector of unknown parameters and σ^2 is an unknown scalar.

(a): (3 points) What is the maximum likelihood estimate of $\boldsymbol{\beta}$?

(b): (3 points) What is the maximum likelihood estimate of σ^2?

Question 7: Let $\theta \in [0, 1]$ be the probability that a coin comes up "heads" when flipped. Suppose the coin is flipped five times and comes up heads each time.

(a): (3 points) What is the maximum likelihood estimate of θ?

(b): (5 points) What is the value of the likelihood function for these data evaluated at the maximum likelihood estimate of θ? In answering this part of the question and the parts that follow, you can assume the outcomes of the coin flips are independent events.

(c): (5 points) What is the value of the likelihood function for these data evaluated at $\theta = .5$?

(d): (7 points) Given the data, assess the plausibility of the hypothesis $H_0 : \theta = .5$ (the coin is "fair") against the one-sided alternative $H_A : \theta > .5$. Be explicit as to how you conduct this test (there is more than one way to do it). What do you conclude?

Question 8: (6 points) Logistic regression models (i.e., logit models for binary data) are sometimes said to be non-linear models. What is linear about a logistic regression model, and what is non-linear? Be brief, but precise and explicit in your answer.

Question 9: On April 15th, 1912, the *Titanic* collided with an iceberg and sank with much loss of life. Logistic regression models were used to analyze data available for 2,201 passengers and crew: a binary dependent variable (coded 1 for survival, 0 otherwise), with predictors measuring class of travel (0 for crew, 1 for first class, 2 for second class, 3 for third class), adult/child, and gender. Table 1 summarizes this analysis, presenting maximum likelihood estimates of coefficients (with standard errors in parentheses) and empty table entries indicating whether that the corresponding variable was omitted from the respective logit model.

	Model 1	Model 2	Model 3	Model 4	Model 5
Intercept	-.74	2.61			1.27
	(.05)	(.29)			(.34)
Class		-.33			
(0-3)		(.05)			
1st class			3.11	2.07	.80
(0-1)			(.30)	(.35)	(.16)
2nd class			2.09	1.04	-.23
(0-1)			(.28)	(.34)	(.18)
3rd class			1.33	.26	-1.01
(0-1)			(.25)	(.32)	(.15)
Crew			2.25	1.27	
(0-1)			(.30)	(.34)	
Adult		-1.01	-1.06	.11	.11
(0-1)		(.25)	(.24)	(.34)	(.34)
Male		-2.61	-2.42	-.72	-.72
(0-1)		(.13)	(.14)	(.41)	(.41)
Adult × Male				-1.90	-1.90
				(.43)	(.43)
Log-Likelihood	-1384.73	-1137.45	-1105.03	-1096.02	-1096.02

Table 1: Logit Estimates, Analysis of *Titanic* Disaster. Maximum likelihood estimates and standard errors in parentheses.

(a): (5 points) What proportion of passengers survived the disaster?

(b): (5 points) What is the substantive content of the null hypothesis tested by a comparison of Models 2 and 3?

(c): (5 points) Use a likelihood ratio test to compare Models 2 and 3.

(d): (5 points) How do Models 4 and Model 5 differ? That is, what is being tested in Model 5, versus Model 4 (or vice-versa)?

(e): (5 points) You need not compute a likelihood ratio test to compare Model 3 and 4. Why?

(f): (5 points) A time-honored policy in disasters at sea is that "women and children" are the first to be rescued. How might we augment or change the analysis presented above to test the extent to which different classes adhered to the "women and children first" policy?

END OF EXAM

Total Number of Points: 154

	Upper Tail Area				
df	.25	.10	.05	.01	.001
2	2.77	4.61	5.99	9.21	13.8
3	4.11	6.25	7.81	11.3	16.3
4	5.39	7.78	9.49	13.3	18.5
5	6.63	9.24	11.1	15.1	20.5
6	7.84	10.6	12.6	16.8	22.5
7	9.04	12.0	14.1	18.5	24.3
8	10.2	13.4	15.5	20.1	26.1
9	11.4	14.7	16.9	21.7	27.9
10	12.5	16	18.3	23.2	29.6
11	13.7	17.3	19.7	24.7	31.3
12	14.8	18.5	21.0	26.2	32.9
13	16.0	19.8	22.4	27.7	34.5
14	17.1	21.1	23.7	29.1	36.1
15	18.2	22.3	25	30.6	37.7
20	23.8	28.4	31.4	37.6	45.3
30	34.8	40.3	43.8	50.9	59.7
50	56.3	63.2	67.5	76.2	86.7
100	109	118	124	136	149
200	213	226	234	249	268
300	316	332	341	360	381
500	521	541	553	576	603
1000	1030	1058	1075	1107	1144
3000	3052	3100	3129	3183	3245

Table 2: Critical values of the χ^2 distribution.