UNIVERSITY OF CAMBRIDGE
Judge Business School

Updated: February 19, 2012

Lecturer: Dr. Paul Kattuman
Contact: pak13@cam.ac.uk

Contest Quiz 6

Question Sheet

In this quiz we will review concepts of binary choice models covered in lecture 7.

Question 1

Civil wars can cause widespread civilian disruption and suffering. Various explanations have been proposed for these conflicts, including religious tensions, undemocratic governments, and heightened economic hardship resulting from declining economic conditions.

Here, we examine the relation between civil conflicts and several of their proposed causes using annual data on 41 sub-Saharan African countries for the years 1981 to 1999. Here, a civil conflict is defined to be an armed fight between the government and a domestic insurgency resulting in at least 25 battle deaths per year. During the period covered by our data, 27 of the 41 countries experienced a civil conflict.

The variables are described in Table 1. The unit of observation is a country i in year t, where $i = 1, ..., 41$ and $t = 1981, ..., 1999$ (so the number of time observations is $T = 19$).

Table 1: Variable description

variable	description
Civil Conflict$_{it}$	= 1 if country i has an internal conflict with at least 25 deaths in year t; = 0 otherwise
2-year GDP growth$_{it}$	= decimal average annual growth rate of real GDP in country i from year $t-2$ to year t; $= \frac{1}{2} \times [ln(GDP_{it}) - ln(GDP_{it-2})]$

Consider the following (1) OLS and (2) Probit Models of Civil Conflict based on 731 observations each. Standard errors in parentheses.

$$\begin{aligned} \text{E(Civil Conflict}_{it} \mid \text{2-year GDP growth}_{it}) &= 0.27 - 0.38 \times \text{2-year GDP growth}_{it} \\ &\quad (.05) \quad (.49) \end{aligned} \tag{1}$$

$$\begin{aligned} \text{P(Civil Conflict}_{it} = 1 \mid \text{2-year GDP growth}_{it}) &= -0.63 - 1.03 \times \text{2-year GDP growth}_{it} \\ &\quad (.16) \quad (1.38) \end{aligned} \tag{2}$$

Using the OLS regression (1):

(i) Compute the *absolute value* of the difference in the probabilities of countries A and B experiencing a civil conflict, where country A has 2-year GDP growth of $+0.10$ while country B has 2-year GDP growth of -0.10.

(ii) Compute a 95% confidence interval for the difference in probabilities in Question 1-i and give the upper bound. [use 1.96 as the value of the critical quantile]

Using the Probit regression (2):

(iii) Compute difference in the probabilities of countries A and B experiencing a civil conflict, for countries A and B as given in Question 1-i. [give your result in percentage points]

(iv) Is the difference in predicted probabilities computed in Question 1-iii statistically significantly different from zero at the 5% significance level?

a) Yes
b) No

(v) In real-world terms, are the numerical answers to Question 1-i and 1-iii similar or different?

a) similar
b) different

Question 2

Let y be a binary dependent variable and $d_1; d_2; ...; d_M$ be dummy regressors for exhaustive and mutually exclusive categories, i.e., each unit falls into only one category.

(i) Assume we run the linear regression

$$y_i = \beta_1 d_{1i} + \beta_2 d_{2i} + ... + \beta_M d_{Mi}; \quad i = 1, 2, ..., N$$

Which of the following are true?

a) The fitted values, $\hat{y}_i$ always fall in the interval $[0, 1]$
b) The estimated coefficients $\hat{\beta}$ always fall in the interval $[0, 1]$
c) The estimated coefficients $\hat{\beta}$ may fall outside the interval $[0, 1]$
d) The fitted values, $\hat{y}_i$ may fall outside the interval $[0, 1]$

(ii) Assume we regress instead

$$y_i = \beta_1 + \beta_2 d_{2i} + ... + \beta_M d_{Mi}; \quad i = 1, 2, ..., N$$

where the regressors are M linearly independent, linear combinations of $d_{1i}, ..., d_{Mi}$. Which of the following are true?

a) The fitted values, $\hat{y}_i$ for all M categories will be the same
b) The fitted values, $\hat{y}_i$ for the M categories may be different
c) The estimated coefficients $\hat{\beta}$ will be the same
d) The estimated coefficients $\hat{\beta}$ may be different

Question 3

A biologist is trying to kill mosquitos with a poisonous drug. The dataset *data_x.csv*[1] contains the following information: The id of the mosquito is coded in the first column, the amount of poison in the second column, the status of the mosquito is shown in the third column. The fourth column, finally, contains the id of the research assistant who administered the poison.

(i) Estimate a logit model that explains survival of the mosquito as a function of the amount of poison. What is the critical amount of poison such that the survival probability is just 1/2?

Our biologist remembers that two of her assistants joined the team later. As a result, their animals were treated for a shorter time. Unfortunately, our biologist forgot who these assistants were.

(ii) To help her, plot for each of the assistants a diagram with the marginal effect on the vertical axis and the amount of poison on the horizontal axis. Are they all similar? Can you identify the two who are different? Note the two assistant ids and enter their sum in the entry sheet.

[1] To access the data go to http://www.camcon.eu/data/series6/ and download *data_x.csv* where x is your group id, e.g. 1 for Group1, displayed in the upper right corner of the website after login.