Midterm 1

Eco 231 - Undergraduate Econometrics

10/04/2010 (Prof. Carolina Caetano)

1 Instructions

Reading and understanding the instructions is your responsibility. Failure to comply may result in loss of points, and there will be no leniency on that respect.

1. This exam has 11 questions and it's worth 100 points. Questions are in the front and back of the exam pages. You have until 5 minutes before the end of the regular class time to answer it.
2. You must answer each question in the space provided. If you need more space, you must ask for extra paper from the TA. It is your responsibility at the end of the exam to staple the extra page exactly in the right place in your exam. A page that is out of order may not be graded. You may ask for draft paper if you like.
3. You are not allowed the use of notes, cheat sheets, calculators, or electronic devices of any kind. If your answers are unclear or illegible you may lose points. You may answer in pencil.
4. You may return the exam at any time until 10 minutes before the end of class time. If you finished within 10 minutes of the end of class time, you must remain seated. Do not get up when the TA announces the time is up. Remain seated and follow the TA's instructions.
5. Fill out the information on this page now.

2 Student information

1. **Name:** _____
2. **Student ID:** _____
3. **Class time:** _____

Good wishes!

Question 1 (50 points): Suppose that you have a data set which contains information on births to women in the United States. Two variables of interest are the dependent variable, infant birth weight in ounces ($bwght$), and an explanatory variable, average number of cigarettes the mother smoked per day during pregnancy ($cigs$). The following simple regression was estimated using data on $n = 2000$ births:

$$\widehat{bwght} = 120 - 0.5\, cigs$$

1.1 (10 points): What is the predicted birth weight when $cigs = 0$? What about when $cigs = 20$ (one pack per day)? Comment on the difference.

1.2 (10 points): Does this simple regression necessarily capture a causal relationship between the child's birth weight and the mothers smoking habits? Explain (more space on the next page).

1.3 (10 points): To predict a birth weight of 125 ounces, what would *cigs* have to be? Comment.

1.4 (10 points, hard): The proportion of women in the sample who do not smoke while pregnant is about 0.85. Does this help reconcile your finding from question (1.3)? Do you believe that a linear model is ideal for this problem?

1.5 (10 points): 1 ounce is equivalent to approximately 30 grams. What would be the values of $\hat{\beta}_0$ and $\hat{\beta}_1$ if birth weight was measured in grams instead of ounces? You don't need to do the calculations. It is enough to leave the multiplications and divisions explicitly stated (more space on the next page).

Question 2 (30 points): Suppose that your data set from the last question also contains information about the women's education in years ($yeduc$), yearly income ($yinc$), and number of alcoholic drinks she consumes per week ($wdrink$). You write the following econometric model:

$$bwght = \beta_0 + \beta_1\, cigs + \beta_2\, yeduc + \beta_3\, yinc + \beta_4\, wdrink + u$$

2.1 (10 points): Incidence of diabetes is more common among smokers than among non smokers. Also, diabetic mothers, or mothers that will be diabetic in the future, have a higher likelihood of giving birth to babies weighting over 140 ounces. Should information about the mother's diabetic status be included in the model? Explain.

2.2 (10 points): Are the variables in the model enough to explain the relationship between the mother's smoking behavior and her child's weight at birth? Are there any important variables that you think are missing? If your answer to the last question is yes, discuss which variables should be included, why, and whether you think they are observable or not (more space on the next page).

2.3 (10 points): The econometric model estimated on question 1 is the following:

$$bwght = \beta_0 + \beta_1\ cigs + u, \tag{1}$$

while the model proposed in question 2 is

$$bwght = \beta_0 + \beta_1\ cigs + \beta_2\ yeduc + \beta_3\ yinc + \beta_4\ wdrink + u. \tag{2}$$

The estimation performed in model (1) yielded $\hat{\beta}_0 = 120$ and $\hat{\beta}_1 = -0.5$. $\tilde{\beta}_0$ and $\tilde{\beta}_1$ are the OLS estimators of β_0 and β_1 using the econometric model (2).

2.3.1 (5 points): Do you believe that $\tilde{\beta}_1$ will be higher or smaller than $\hat{\beta}_1$? Explain.

2.3.2 (5 points, hard): Do you believe that $\tilde{\beta}_0$ will be higher or smaller than $\hat{\beta}_0$? Explain.

Question 3 (20 points):

3.1 (10 points): What are the assumptions required for the unbiasedness of the OLS estimator?

3.2 (10 points): Suppose that you already proved that $E(\hat{\beta}_1 \mid x_1, \ldots, x_n) = \beta_1$. Using this information, prove that $E(\hat{\beta}_0) = \beta_0$.

3 FOR NEXT TIME

Question 2.3 has different tilde and hat definitions than the book. For the book, $\tilde{\beta}$ is the univariate one. Maybe it's good to be careful, change them or something.

This exam format is silly. In the end, since I am returning the midterm to the students, they will keep it, won't they? I should think better about what is the right way to do this.

Question for the final:

$$y = \beta_0 + \beta_1 x_i + u$$

satisfies the Gauss-Markov assumptions. Let $\tilde{\beta}_0$ and $\tilde{\beta}_1$ be obtained by solving the equations

$$\sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i) w_i = 0$$
$$\sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_1) x_i w_i = 0$$

for $\beta_0, \beta_1, \ldots, \beta_k$, where $\sum_{i=1}^{n} w_i = 1$. Derive