ECON 484
Spring 2011
Prof. Simon Woodcock

FINAL EXAM

There are 2 questions, each with multiple parts. All parts are given equal weight. Don't Panic!

1. Be sure to read all parts of this question carefully before you begin (your answer to part d will depend on how you choose to answer part c, so choose carefully).
    a. What is selection bias? Explain how and why selection bias causes problems for estimating the causal effect of a treatment D on an outcome of interest, Y.
    b. Suppose the university decides to offer free weekly tutoring sessions to any SFU student who wants tutoring. Let Y_i denote the GPA of student i at graduation. Let $D_i = 1$ if student i participates in the free tutoring sessions, and $D_i = 0$ if student i does not. Let X_i denote a set of other observable characteristics of student i that are thought to affect GPA (e.g., major, number of hours worked for pay, etc.). Suppose we estimate the regression, $Y_i = X_i\beta + D_i\gamma + \varepsilon_i$. Does γ estimate the causal effect of the free tutoring sessions on GPA at graduation? Is there a selection bias problem here? If so, what is the nature of the selection bias, and what is its consequence? Explain.
    c. Describe a better econometric methodology for estimating the causal effect of the tutoring program on GPA. Feel free to change the tutoring program (e.g., who is eligible for tutoring, what kind tutoring is offered, how often, etc.) if doing so is helpful. Be sure to explain what kind of data your method requires, what model you would estimate, how your method deals with the selection bias problem, and what assumptions are required for your method to give an unbiased or consistent estimate of the causal effect of interest. If you choose to change the tutoring policy as part of your answer, be sure to clearly explain the change(s).
    d. Choose one paper that we discussed this semester, and briefly describe how it applies your method from part c to estimate a causal effect. Be sure to explain what causal effect the authors are trying to estimate, what data they use, what model they estimate, how their method deals with the selection bias problem, and how it identifies the causal effect of interest.

2. As you probably know, there is a Federal election underway in Canada. The Liberal Party of Canada recently announced that if they are elected, they will introduce a new policy called the Canadian Learning Passport (CLP). This policy will provide $4000 to each student who graduates from high school and enrolls in a university or college. Suppose that the Liberal Party is elected, that they implement the CLP, and that high school students graduating in 2012 are the first cohort to receive the $4000 CLP grant. Suppose that sometime in the future, you are hired by the government of Canada to measure the effect of the CLP on **university** graduation. You have data on a random sample of 4000 young Canadians: 1000 who graduated from high school in 2010, 1000 who graduated from high school in 2011, 1000 who graduated from high school in 2012, and 1000 who graduated from high school in 2013. Your sample is randomly selected from all ten provinces. For each member *i* of the sample, you observe whether they eventually graduate from university (Y_i), the year in which they graduated from high school (C_i), whether they received the CLP grant (G_i), the province in which they went to high school (P_i), their birth day (B_i), their score (out of 100) on a high school exit exam (S_i), and a set of background characteristics (X_i). The high school exit exam is designed to test graduates' mastery of high school-level academic skills (mathematics, language, etc.).
    a. Suppose that the CLP is implemented in all provinces except Quebec. Explain how you can use this information to measure the causal effect of the CLP on university graduation. What assumptions are required for your method to estimate the causal effect?
    b. Suppose instead that the CLP is implemented in all 10 provinces. However, high school graduates are only eligible to receive the CLP if they were born on an odd-numbered day (i.e., graduates born on the 1^{st}, 3^{rd}, 5^{th}, (etc.), day of the month receive the CLP grant; but graduates born on the 2^{nd}, 4^{th}, 6^{th}, (etc.), day of the month do not; yes, I know this is a strange policy). Explain how you can use this information to measure the causal effect of the CLP on university graduation. What assumptions are required for your method to estimate the causal effect?
    c. Suppose instead that the CLP is implemented in all 10 provinces, and eligibility for CLP no longer depends on birth day. However, only students who score 80 or better on the high school exit exam receive the CLP grant. Explain how you can use this information to measure the causal effect of the CLP on university graduation. What assumptions are required for your method to estimate the causal effect?
    d. Suppose instead that the CLP is implemented in all 10 provinces, but now eligibility for CLP depends on birth day **and** the exit exam score. Specifically, the **probability** that student *i* receives the CLP grant is $S_i/200 + B_i/50$. Explain how you can use this information to measure the causal effect of the CLP on university graduation. What assumptions are required for your method to estimate the causal effect?
    e. Which of your four estimators from parts a-d do you prefer, and why? Explain.