Honors Econometrics Exam
Swarthmore College, Spring 2017
Examiner: Paul Goldsmith-Pinkham, Federal Reserve Bank of New York
Time: 3 Hours

This exam has 5 questions. You may use a calculator. Statistical tables are **not** required to complete the exam. Make sure to answer all parts of each question! Partial credit will be awarded, so make sure to attempt each question.

Question 1 [35 minutes]

Choose four of these statements, and write (a) whether you believe this statement is True, False, or Uncertain and (b) why (in 2 to 4 sentences). Credit depends on the substance of your answer, not on whether you are correct in writing True, False or Uncertain.

1. A sample of 30 observations is sufficient for consistent estimates in regressions.
2. When modeling a binary outcome, it is fine to use a simple linear model.
3. When estimating models with missing data, it is fine to drop those observations.
4. When estimating a regression in Stata, you should always use the `robust` command to get heteroskedasticity-robust standard errors.
5. One should always include lagged values in a time series regression to avoid the problem of spurious regressions.
6. Fixed effects estimation assumes sequential exogeneity.
7. The OLS estimator is the best linear unbiased estimator.
8. If you omit a variable in a regression that is correlated with your outcome, your regression estimates will be biased.
9. If you run a regression of y on x and find a positive coefficient, this implies that increasing x will always lead to greater y.
10. Consistency is more important than efficiency.

Question 2 [40 minutes]

You are one of the chief risk modelers for Countrywide Financial in 2006, and studying the probability of mortgage default. In the interest of simplicity, you are studying the relationship between local annual house price appreciation (HPA) and share of local mortgages in default (FracDefault). You have N locations in your dataset, and know the following basic summary stats about your variables:

Variable	Mean	Minimum	Maximum
HPA	7.63	0.81	30.58
FracDefault	5.84	0	22.72

a

You first consider the simple linear regression

$$\text{FracDefault}_i = \alpha + \beta \text{HPA}_i + \epsilon_i.$$

Derive the formula for the OLS estimate of β.

b

You run the regression and get point estimates of $\hat{\alpha}$ =`8.11`, $\hat{\beta}$ =`-0.337`. Interpret, in words, these estimates. Be precise.

c

A time traveler from the future arrives briefly in front of you to warn that house prices will drop by `17` percent in the next year (i.e., HPA $= -17$), and then disappears in a flash. This number is unprecedented, and you want to know how this will affect your forecast of mortgage defaults, which was based on HPA continuing at average levels. Report the difference in predicted mortgage defaults, using the time traveler's value of HPA and your value of HPA (continuing at the average value).

d

Write down the formula that, under homoskedastictiy, would allow you to test for whether the difference reported in part c is statistically different from zero. Use algebraic placeholders where necessary.

e

You start to worry that your model of mortgage defaults is wrong as well. You estimate the following additional model:

$$\text{FracDefault}_i = \underbrace{\alpha}_{\texttt{10.58}} + \underbrace{\beta_1}_{\texttt{-0.91}} \text{HPA}_i + \underbrace{\beta_2}_{\texttt{0.02}} \text{HPA}_i^2 + \epsilon_i, \tag{1}$$

with point estimates of $\hat{\alpha}$ =`10.58`, $\hat{\beta}_1$ =`-0.91`, and $\hat{\beta}_2$ =`0.02`. What values would you predict under your original assumption of HPA continuing at average values? What values would you predict under the time traveller's values of HPA? Explain, graphically and in words, the difference between these predictions and the predictions in part c. What would you recommend to a researcher interested in predicting defaults under the time traveler's HPA scenario?

Question 3 [30 minutes]

Consider a panel regression, with outcome Y_{it}, and explanatory variable X_{it}, with i denoting individuals in a sample of N, and t denoting years, with T years.

a

Discuss the difference between using a random effects vs. fixed effects model in this context. Do these models necessarily estimate the same effect of X_{it} on Y_{it}? How can a researcher decide which to use?

b

Now consider the following regression:

$$Y_{it} = \alpha_i + X_{it}\beta + Y_{i,t-1}\gamma + \epsilon_{it}.$$

Explain in words why the OLS estimates of β and γ are inconsistent.

c

Describe a simple consistent estimator for β and γ in part b, assuming the α_i and ϵ_{it} are i.i.d. and independent of X_{it}.

Question 4 [45 minutes]

You are a researcher studying the effects of a training program. You have randomly assigned a binary treatment, $T \in \{0, 1\}$, to a sample of N individuals. You are estimating the effect of the treatment on employment, Y:

$$Y_i = \alpha + \beta T_i + \epsilon_i.$$

a

Derive the OLS estimator of the effect of T_i on Y_i. Assuming that $Cov(T_i, \epsilon_i) = 0$ and that the law of large numbers holds, write the probability limit of this estimator. Show that this estimate can be written as the difference in two means.

Hint: remember that T_i is binary.

b

Now, assume that $Cov(T_i, \epsilon_i) \neq 0$, and instead you have a binary instrument $Z_i \in \{0, 1\}$ such that $Cov(Z_i, \epsilon_i) = 0$ and $Cov(Z_i, T_i) \neq 0$. Derive the instrumental variables estimator for β. Again, assume that the law of large numbers holds and write the probability limit of this estimator. Show that this estimate can be written as the ratio of two differences in means.

Hint: remember that Z_i is binary.

c

Now assume that $Cov(Z_i, \epsilon_i) = \mu_1$, and $Cov(T_i, \epsilon_i) = \mu_2$. Compare the biases of the IV estimator and the OLS estimator. If $|\mu_2| > |\mu_1|$, does that mean that the bias of the IV estimator is smaller than the OLS estimator?

d

Now assume that there are two (unobservable) types g: economists ($g = e$) and humans ($g = h$). We assume that

$$Y_i = \alpha + \beta_g T_i + \epsilon_i,$$

where $\beta_g \in \{\beta_e, \beta_h\}$ depending on whether person i is type e or type h. Assume that $\Pr(g = e|T_i = 1) = \Pr(g = e|T_i = 0) = \Pr(g = e) = \pi$. Under the assumptions from part a, write down the probablity limit of the OLS estimator of T_i on Y_i.

Hint: consider rewriting $\beta_g = \beta_h + 1(g = e) \times (\beta_e - \beta_h)$.

e

Finally, consider the same heterogeneous relationship between both T_i and Z_i, as well as T_i and Y_i:

$$Y_i = \alpha + \beta_g T_i + \epsilon_i$$
$$T_i = \tau + \gamma_g Z_i + u_i.$$

Under the assumptions of part b, derive the IV estimate of using Z_i as an instrument for T_i.

f

Discuss, in words, how the heterogeneity across groups implies that even with a valid instrument, you could find an arbitrarily large or small effect as an estimate of β, the effect of T_i on Y_i.

Question 5 [40 minutes]

Read the Economist article on the following page, and in about 3-5 paragraphs, discuss the following:

- Compare and contrast the methodological concerns that Cartwright and Deaton raise about RCTs with the concerns raised about machine learning by Cathy O'Neil.
- Do you agree with the following statement: "by offering alluringly simple ways of evaluating certain policies, economists lose sight of policy questions that are not easily testable using RCTs, such as the effects of institutions, monetary policy or social norms."

Economists are prone to fads, and the latest is machine learning

Big data have led to the latest craze in economic research What is the collective noun for a group of economists? Options include a gloom, a regression or even an assumption. In January, when PhD students jostle for jobs at the annual meeting of the American Economic Association, a "market" might seem the mot juste. Or perhaps, judging by the tendency of those writing economic papers to follow the latest fashion, a "herd" would be best. This year the hot technique is machine learning, using big data; Imran Rasul, an economics professor at University College, London, is expecting to read a pile of papers using this voguish technique.

Economists are prone to methodological crazes. Mr Rasul recalls past paper-piles using the regression-discontinuity technique, which compared similar people either side of a sharp cut-off to gauge a policy's effect. An analysis by The Economist of the key words in working-paper abstracts published by the National Bureau of Economic Research, a think-tank (see chart), shows tides of enthusiasm for laboratory experiments, randomised control trials (RCTs) and the difference-in-differences approach (ie, comparing trends over time between different groups).

When a hot new tool arrives on the scene, it should extend the frontiers of economics and pull previously unanswerable questions within reach. What might seem faddish could in fact be economists piling in to help shed light on the discipline's darkest corners. Some economists, however, argue that new methods also bring new dangers; rather than pushing economics forward, crazes can lead it astray, especially in their infancy.

In 1976 James Heckman developed a simple way of correcting for the problem of a specific type of sample selection. For example, economists had difficulty estimating the effect of education on women's wages, because the ones who chose to work (for whom pay could be measured) were particularly likely to enjoy high returns. When Mr Heckman offered economists a simple way of correcting this bias, which involved accounting for the choice to enter work, it took the social sciences by storm. But its seductive simplicity led to its misuse.

A paper by Angus Deaton, a Nobel laureate and expert data digger, and Nancy Cartwright, an economist at Durham University, argues that randomised control trials, a current darling of the discipline, enjoy misplaced enthusiasm. RCTs involve randomly assigning a policy to some people and not to others, so that researchers can be sure that differences are caused by the policy. Analysis is a simple comparison of averages between the two. Mr Deaton and Ms Cartwright have a statistical gripe; they complain that researchers are not careful enough when calculating whether two results are significantly different from one another. As a consequence, they suspect that a sizeable portion of published results in development and health economics using RCTs are "unreliable".

With time, economists should learn when to use their shiny new tools. But there is a deeper concern: that fashions and fads are distorting economics, by nudging the profession towards asking particular questions, and hiding bigger ones from view. Mr Deaton's and Ms Cartwright's fear is that RCTs yield results while appearing to sidestep theory, and that "without knowing why things happen and why people do things, we run the risk of worthless causal ('fairy story') theorising, and we have given up on one of the central tasks of economics." Another fundamental worry is that by offering alluringly simple ways of evaluating certain policies, economists lose sight of policy questions that are not easily testable using RCTs, such as the effects of institutions, monetary policy or social norms.

Elsewhere in economics one methodology has on occasion crowded others out. An excess of consensus among macroeconomists in the run-up to the financial crisis has haunted them. In August, Olivier Blanchard, a heavyweight macroeconomist, wrote a plea to colleagues to be less "imperialistic" about their use of dynamic stochastic general equilibrium models, adding that, for forecasting, their theoretical purity might be "more of a hindrance than a strength". He issued a reminder that "different model types are needed for different tasks."

Machine learning is still new enough for the backlash to be largely restricted to academic eye-rolling. But some familiar themes are emerging in this latest craze. In principle, these new techniques should protect economists from their own sloppy theorising. Before, economists would try to predict things using only a few inputs. With machine learning, the data speak for themselves; the machine learns which inputs generate the most accurate predictions.

This powerful method appears to have improved the accuracy of economists' predictions. For example, researchers have started to use big data to predict whether a criminal suspect is likely to come back to court for a trial, influencing bail decisions. But, as with RCTs, a powerful algorithm might seduce its users into ignoring underlying causal factors. In her new book,"Weapons of Math Destruction", Cathy O'Neil, a data scientist, points out that some factors, such as race or coming from a high-crime neighbourhood, might be excellent predictors of recidivism. But they could reflect racism in law enforcement or zero-tolerance "broken windows" policies that lead to high recorded crime rates in poor or minority neighbourhoods. If so, those predictions risk punishing people for factors beyond their control.

Mr Rasul is not very worried by the "little bit of overshooting" that excitement at new methods engenders. Over time, their merits and limitations are better appreciated and they join the toolkit alongside older methods. But the critics of faddishness have one thing right. Good economics is about asking the right questions. Of all the tools at the discipline's disposal, its practitioners' scepticism is the most timeless.