計量經濟學導論暨實習

期末考

2021.6.23.

請依序作答。

一、單選題，回答正確選項即可，不需說明。每題3分，共45分。

1. In the linear probability model, the interpretation of the slope coefficient is
   (a) the change in odds associated with a unit change in X, holding other regressors constant.
   (b) not all that meaningful since the dependent variable is either 0 or 1.
   (c) the change in probability that $Y = 1$ associated with a unit change in X, holding others regressors constant.
   (d) the response in the dependent variable to a percentage change in the regressor.

2. The maximum likelihood estimation method produces, in general, all of the following desirable properties with the **exception** of
   (a) efficiency.
   (b) consistency.
   (c) normally distributed estimators in large samples.
   (d) unbiasedness in small samples.

3. Having more relevant instruments
   (a) is a problem because instead of being just identified, the regression now becomes over-identified.
   (b) is like having a larger sample size in that the more information is available for use in the IV regressions.
   (c) typically results in larger standard errors for the TSLS estimator.
   (d) is not as important for inference as having the same number of endogenous variables as instruments.

4. Weak instruments are a problem because
   (a) the TSLS estimator may not be normally distributed, even in large samples.
   (b) they result in the instruments not being exogenous.
   (c) the TSLS estimator cannot be computed.
   (d) you cannot predict the endogenous variables any longer in the first stage.

5. Failure to follow the treatment protocol means that
   (a) the OLS estimator cannot be computed.
   (b) instrumental variables estimation of the treatment effect should be used where the initial random assignment is the instrument for the treatment actually received.
   (c) you should use the TSLS estimator and regress the outcome variable Y on the initial random assignment in the first stage to get predicted values of the outcome variable.
   (d) the Hawthorne effect plays a crucial role.

6. In a sharp regression discontinuity design:
   (a) crossing the threshold influences receipt of the treatment but is not the sole determinant.
   (b) the population regression line must be linear above and below the threshold.
   (c) X_i will in general be correlated with u_i.
   (d) receipt of treatment is entirely determined by whether W exceeds the threshold.

7. In the Project STAR experiment, the observations are not plausibly i.i.d. As a result,
   (a) we should use homoskedasticity-only standard errors.
   (b) we should use clustered standard errors at the school level.
   (c) we should use clustered standard errors across schools.
   (d) the OLS estimator will be biased.

8. The MSPE by m-fold cross validation is given by the following formula:
   (a) $\widehat{MSPE}_{m-fold\ cross\ validation} = \frac{1}{m}\sum_{i=1}^{m}\widehat{MSPE}_i$.
   (b) $\widehat{MSPE}_{m-fold\ cross\ validation} = \sum_{i=1}^{m}\left(\frac{n_i}{n/m}\right)\widehat{MSPE}_i$.
   (c) $\widehat{MSPE}_{m-fold\ cross\ validation} = \frac{1}{m}\sum_{i=1}^{m}\left(\frac{n_i}{n/m}\right)^2\widehat{MSPE}_i$.
   (d) $\widehat{MSPE}_{m-fold\ cross\ validation} = \frac{1}{m}\sum_{i=1}^{m}\left(\frac{n_i}{n/m}\right)\widehat{MSPE}_i$.

9. In order for the ridge regression estimator to work well for the data at hand, you choose λ_{Ridge}
   (a) arbitrarily.
   (b) by minimizing $S^{Ridge}(b, \lambda_{Ridge})$.
   (c) by setting it to 1.
   (d) to minimize the estimated $MSPE$.

10. When $k = 1$, the Lasso estimator is given by the following formula when $\hat{\beta} \geq 0$:
   (a) $\hat{\beta}^{Lasso} = \min\left(\hat{\beta} - \frac{1}{2}\frac{\lambda^{Lasso}}{\sum_{i=1}^{n}X_i^2}, 0\right)$
   (b) $\hat{\beta}^{Lasso} = \max\left(\hat{\beta} - \frac{1}{2}\frac{\lambda^{Lasso}}{\sum_{i=1}^{n}X_i^2}, 0\right)$
   (c) $\hat{\beta}^{Lasso} = \min\left(\hat{\beta} + \frac{1}{2}\frac{\lambda^{Lasso}}{\sum_{i=1}^{n}X_i^2}, 0\right)$
   (d) $\hat{\beta}^{Lasso} = \max\left(\hat{\beta} + \frac{1}{2}\frac{\lambda^{Lasso}}{\sum_{i=1}^{n}X_i^2}, 0\right)$

11. The BIC is a statistic
   (a) used to help the researcher choose the number of lags in an autoregression.
   (b) commonly used to test for serial correlation.
   (c) only used in cross-sectional analysis.
   (d) developed by the Bank of England in its river of blood analysis.

12. You should use the QLR test for breaks in the regression coefficients, when
   (a) the Chow F-test has a p value of between 0.05 and 0.10.
   (b) the suspected break data is not known.
   (c) there are breaks in only some, but not all, of the regression coefficients.
   (d) the suspected break data is known.

13. Problems caused by stochastic trends include all of the following with the **exception** of
   (a) the estimator of an AR(1) is biased towards zero if its true value is one.
   (b) the model can no longer be estimated by OLS.
   (c) t-statistics on regression coefficients can have a non-normal distribution, even in large samples.
   (d) the presence of spurious regression.

14. HAC standard errors should be used because
    (a) they are convenient simplifications of the heteroskedasticity-robust standard errors.
    (b) conventional standard errors may result in misleading inference.
    (c) they are easier to calculate than the heteroskedasticity-robust standard errors and yet still allow you to perform inference correctly.
    (d) when there is a structural break, then conventional standard errors result in misleading

15. Consider the distributed lag model $Y_t = \beta_0 + \beta_1 X_t + \beta_2 X_{t-1} + \beta_3 X_{t-2} + \cdots + \beta_{r+1} X_{t-r} + u_t$. The dynamic causal effect is
    (a) $\beta_0 + \beta_1$.
    (b) $\beta_1 + \beta_2 + \cdots + \beta_{r+1}$.
    (c) $\beta_0 + \beta_1 + \cdots + \beta_{r+1}$.
    (d) β_1.

二、非選擇題,共55分。答題時請務必簡要清楚,並適當說明你的想法,答案以能讓閱卷者瞭解爲原則。

1. (20%) Consider a model where random variables X and Y are jointly determined by the following equations:

$$Y = \beta_1 + \beta_2 X + \beta_3 Z + u, \quad (1)$$
$$X = \gamma_1 + \gamma_2 Y + v, \quad (2)$$

where u and v are error terms, mutually uncorrelated with mean zero and variance σ_u^2 and σ_v^2, and Z is an exogenous variable.

    (a) (2%) Are β_2 and γ_2 identified? Explain.
    (b) (2%) Derive the covariance between Y and v, $\text{Cov}(Y, v)$.
    (c) (4%) Is the OLS estimate of γ_2 in equation (2) consistent in general? Is it consistent if $\beta_2 = 0$?
    (d) (4%) Is there any valid instrument to estimate γ_2. Why? If yes, show that the corresponding instrument variable estimate γ_2^{IV} is consistent.
    (e) (4%) Show that the two stage least square (2SLS) estimate of γ_2 is identical to the IV estimate in (d). Note that that 2SLS estimate is the OLS estimate of γ_2 in (2) when Y is replace by its fitted values by OLS in the equation $Y = \delta_1 + \delta_2 Z + w$, where w is an error term.
    (f) (4%) Suppose that $\gamma_1 = 0$. Is $\frac{\bar{X}}{\bar{Y}}$ a consistent estimator of γ_2? Why?

2. (10%) Suppose Y_t follows a random walk, $Y_t = Y_{t-1} + u_t$, for $t = 1, \cdots, T$, where $Y_0 = 0$ and u_i is i.i.d. with mean 0 and variance σ_u^2.

    (a) (4%) Compute the mean and variance of Y_t.
    (b) (3%) Compute the covariance of Y_t and Y_{t-k}.
    (c) (3%) Use the results in (a) and (b) to determine whether Y_t is stationary or nonstationary.

3. (10%) Suppose there are panel data for $T = 2$ time periods for a randomized controlled experiment, where the first observation ($T = 1$) is taken before the experiment and the second observation ($t = 2$) is for the posttreatment period. Suppose the treatment is binary; that is, suppose $X_{it} = 1$ if the i^{th} individual is in the treatment group and $t = 2$, and $X_{it} = 0$ otherwise. Further suppose the treatment effect can be modeled using the specification

$$Y_{it} = \alpha_i + \beta_1 X_{it} + u_{it},$$

where α_i are individual-specific effects with a mean of 0 and a variance of σ_α^2 and u_{it} is an error term, where u_{it} is **homoskedastic**, $\text{Cov}(u_{i1}, u_{i2}) = 0$, and $\text{Cov}(u_{it}, \alpha_i) = 0$ for all i. Let $\hat{\beta}_1^{differences}$ denote the differences estimator— that is, the OLS estimator in a regression of Y_{i2} on X_{i2} with an intercept— and let $\hat{\beta}_1^{diffs\text{-}in\text{-}diffs}$ denote the differences-in-differences estimator— that is, the estimator of β_1 based on the OLS regression of $\Delta Y_i = Y_{i2} - Y_{i1}$ against $\Delta X_i = X_{i2} - X_{i1}$ and an intercept. (**Note**: The homoskedasticity-only formula for the variance of the OLS estimator of β_1 for $Y_i = \beta_0 + \beta_1 X_i + \epsilon$ is $\text{Var}(\hat{\beta}_1) = \frac{\sigma_\epsilon^2}{n\text{Var}(X_i)}$.)

(a) (4%) What is the variance of the differences estimator, $\text{Var}(\hat{\beta}_1^{differences})$?

(b) (4%) What is the variance of the differences-in-differences estimator, $\text{Var}(\hat{\beta}_1^{diffs\text{-}in\text{-}diffs})$?

(c) (2%) Based on your answers to (a) and (b), when would you prefer the differences-in-differences estimator over the differences estimator, based purely on efficiency consideration.

4. (15%) A prospective employer receives two resumes: a resume from a **white** job applicant and similar resume from an **African American** applicant. Is the employer more likely to call back the white applicant to arrange an interview? Bertrand and Mullainathan (2004) carried out a randomized controlled experiment to answer this question. Because race is not typically included on a resume, they differentiated resumes on the basis of "white-**sounding** names" and "African American-**sounding** names". A large collection of fictitious resumes was created, and the presupposed "race" (base on the "sound" of the name) was randomly assigned to each resume. These resumes were sent to prospective employers to see which resumes generated a phone call (a **call back**) from prospective employer. Refer to the attached pages for "**Document for Names Data**" and various regression results from the **STATA log file**, answer the following questions.

(a) (4%) Define the *callback rate* as the fraction of resumes that generate a phone call from the prospective employer. What was the call back rate for whites? For African Americans? Is the difference statistically significant? Is it large in a real-world sense?

(b) (3%) Is the African American/white callback rate differential different for men than for women?

(c) (4%) What was the callback rate for high-quality resumes? For low-quality resumes? Is there a significant difference in callback rates for **high-quality** versus **low-quality** resumes?

(d) (4%) What is the high-quality/low-quality difference for white applicants? For African American applicants? Is there a significant difference in this high-quality/low-quality differences for whites versus African Americans?

Documentation for Names Data

Names contains resume, call-back and employer information for 4,870 fictitious resumes sent in response to employment advertisements in Chicago and Boston in 2001, in a randomized controlled experiment conducted by Marianne Bertrand and Sendhil Mullainathan. The resumes contained information concerning the race of the applicant. Because race is not typically included on a resume, resumes were differentiated on the basis of so-called "white sounding names" (such as Emily Walsh or Gregory Baker) and "African American sounding names" (such as Lakisha Washington or Jamal Jones). A large collection of fictitious resumes were created and the presupposed "race" (based on the "sound" of the name) was randomly assigned to each resume. These resumes were sent to prospective employers to see which resumes generated a phone call (a "call back") from the prospective employer. These data were provided by Professor Marianne Bertrand of the University of Chicago, and were used in her paper with Sendhil Mullainathan "Are Emily and Greg More Employable that Lakisha and Jamal? A Field Experiment on Labor Market Discrimination," *American Economic Review*" 2004, Vol. 94, no. 4.

Variable Descriptions are provided on the next page.

Variable Descriptions

Variable Name	Description
Key Variables	
firstname	applicant's first name
female	1 = female
black	1 = black
high	1= high quality resume
call_back	1= applicant was called back
chicago	1 = data from Chicago
Detailed Information on Resume	
ofjobs	number of jobs listed on resume
yearsexp	number of years of work experience on the resume
honors	1=resume mentions some honors
volunteer	1=resume mentions some volunteering experience
military	1=applicant has some military experience
empholes	1=resume has some employment holes
workinschool	1=resume mentions some work experience while at school
email	1=email address on applicant's resume
computerskills	1=resume mentions some computer skills
specialskills	1=resume mentions some special skills
college	applicant has college degree or more
Detailed Information Concerning Employer	
expminreq	min experience required, if any
eoe	1=ad mentions employer is EOE
manager	1=manager wanted
supervisor	1=supervisor wanted
secretary	1=secretary wanted
offsupport	1=office support
salesrep	1=sales representative wanted
retailsales	1=retail sales worker wanted
req	1=ad mentions any requirement for job
expreq	1=ad mentions some experience requirement
comreq	1=ad mentions some communication skills requirement
educreq	1=ad mentions some educational requirement
compreq	1=ad mentions some computer skill requirement
orgreq	1=ad mentions some organizational skills requirement
manuf	1=employer industry is manufacturing
transcom	1=employer industry is transport/communication
bankreal	1=employer industry is finance, insurance, real estate
trade	1=employer industry is wholesale or retail trade
busservice	1=employer industry is business and personal services
othservice	1=employer industry is health, educ. and social services
missind	1=employer industry is other/unknown

```
------------------------------------------------------------------------------------------------------
      name:  <unnamed>
       log:  final_names.log
  log type:  text
 opened on:  17 Jun 2021, 11:34:01

.
. use names.dta

. desc

Contains data from names.dta
  obs:         4,870
 vars:            38                          16 Jan 2006 11:05
-------------------------------------------------------------------------------------------------------
variable name   type    format     label      variable label
-------------------------------------------------------------------------------------------------------
ofjobs          byte    %8.0g                 number of jobs listed on resume
yearsexp        byte    %8.0g                 number of years of work experience on the resume
honors          byte    %8.0g                 1=resume mentions some honors
volunteer       byte    %8.0g                 1=resume mentions some volunteering experience
military        byte    %8.0g                 1=applicant has some military experience
empholes        byte    %8.0g                 1=resume has some employment holes
workinschool    byte    %8.0g                 1=resume mentions some work experience while at school
email           byte    %8.0g                 1=email address on applicant's resume
computerskills  byte    %8.0g                 1=resume mentions some computer skills
specialskills   byte    %8.0g                 1=resume mentions some special skills
firstname       str8    %9s                   applicant's first name
expminreq       str4    %9s                   min experience required, if any
eoe             float   %8.0g                 1=ad mentions employer is EOE
manager         float   %9.0g                 manager wanted
supervisor      float   %9.0g                 supervisor wanted
secretary       float   %9.0g                 secretary wanted
offsupport      float   %9.0g
salesrep        float   %9.0g                 sales representative wanted
retailsales     float   %9.0g                 retail sales worker wanted
req             float   %9.0g                 1=ad mentions any requirement for job
expreq          float   %9.0g                 1=ad mentions some experience requirement
comreq          float   %9.0g                 1=ad mentions some communication skills requirement
educreq         float   %9.0g                 1=ad mentions some educational requirement
compreq         float   %9.0g                 1=ad mentions some computer skill requirement
orgreq          float   %9.0g                 1=ad mentions some organizational skills requirement
manuf           float   %9.0g                 employer industry is manufacturing
transcom        float   %9.0g                 employer industry is transport/communication
bankreal        float   %9.0g                 employer industry is finance, insurance, real estate
trade           float   %9.0g                 employer industry is wholesale or retail trade
busservice      float   %9.0g                 employer industry is business and personal services
othservice      float   %9.0g                 employer industry is health, educ. and social services
missind         float   %9.0g                 employer industry is other/unknown
black           float   %9.0g
chicago         float   %9.0g
high            float   %9.0g
female          float   %9.0g
college         float   %9.0g
call_back       float   %9.0g
--------------------------------------------------------------------------------------------------------
Sorted by:

. summarize

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
      ofjobs |      4,870    3.661396    1.219126          1          7
    yearsexp |      4,870    7.842916    5.044612          1         44
      honors |      4,870    .0527721    .2236011          0          1
   volunteer |      4,870     .411499    .4921558          0          1
    military |      4,870    .0971253    .2961587          0          1
-------------+---------------------------------------------------------
    empholes |      4,870    .4480493    .4973449          0          1
workinschool |      4,870    .5595483    .4964923          0          1
       email |      4,870    .4792608     .499621          0          1
computersk~s |      4,870    .8205339    .3837815          0          1
specialski~s |      4,870    .3287474    .4698062          0          1
-------------+---------------------------------------------------------
   firstname |          0
   expminreq |          0
         eoe |      4,870    .2911704    .4543485          0          1
     manager |      4,870    .1521561     .359209          0          1
  supervisor |      4,870    .0772074    .2669476          0          1
-------------+---------------------------------------------------------
   secretary |      4,870    .3328542    .4712832          0          1
  offsupport |      4,870    .1186858    .3234517          0          1
    salesrep |      4,870    .1511294    .3582117          0          1
 retailsales |      4,870    .1679671    .3738755          0          1
```

```
         req |      4,870     .787269     .409281          0          1
-------------+---------------------------------------------------------
     expreq |      4,870    .4353183    .4958495          0          1
     comreq |      4,870     .124846    .3305781          0          1
    educreq |      4,870    .1067762    .3088602          0          1
    compreq |      4,870    .4371663    .4960872          0          1
     orgreq |      4,870    .0726899    .2596535          0          1
-------------+---------------------------------------------------------
       manuf |      4,870    .0829569    .2758454          0          1
   transcom |      4,870    .0303901     .171676          0          1
   bankreal |      4,870    .0850103    .2789256          0          1
      trade |      4,870     .213963    .4101431          0          1
 busservice |      4,870    .2677618    .4428382          0          1
-------------+---------------------------------------------------------
 othservice |      4,870    .1548255    .3617754          0          1
    missind |      4,870    .1650924     .371302          0          1
      black |      4,870          .5    .5000513          0          1
    chicago |      4,870    .5552361    .4969906          0          1
       high |      4,870    .5022587    .5000462          0          1
-------------+---------------------------------------------------------
     female |      4,870    .7691992    .4213885          0          1
    college |      4,870    .7195072    .4492861          0          1
  call_back |      4,870    .0804928    .2720826          0          1

. gen f_b = female*black

. gen h_b = high*black

.
. /* model (1) */
. reg call_back black, r

Linear regression                               Number of obs     =      4,870
                                                F(1, 4868)        =      16.93
                                                Prob > F          =     0.0000
                                                R-squared         =     0.0035
                                                Root MSE          =     .27164

------------------------------------------------------------------------------
             |               Robust
   call_back |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       black |  -.0320329     .007785    -4.11   0.000    -.0472949   -.0167708
       _cons |   .0965092    .0059853    16.12   0.000     .0847753    .1082431
------------------------------------------------------------------------------

.
. /* model (2) */
. reg call_back black f_b, r

Linear regression                               Number of obs     =      4,870
                                                F(2, 4867)        =       8.80
                                                Prob > F          =     0.0002
                                                R-squared         =     0.0035
                                                Root MSE          =     .27166

------------------------------------------------------------------------------
             |               Robust
   call_back |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       black |  -.0382214    .0116566    -3.28   0.001    -.0610735   -.0153693
         f_b |     .00799    .0115272     0.69   0.488    -.0146085    .0305886
       _cons |   .0965092    .0059859    16.12   0.000     .0847741    .1082443
------------------------------------------------------------------------------

.
. /* model (3) */
. reg call_back high, r

Linear regression                               Number of obs     =      4,870
                                                F(1, 4868)        =       3.25
                                                Prob > F          =     0.0713
                                                R-squared         =     0.0007
                                                Root MSE          =     .27202

------------------------------------------------------------------------------
             |               Robust
   call_back |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        high |   .0140574    .0077932     1.80   0.071    -.0012207    .0293356
       _cons |   .0734323    .0052991    13.86   0.000     .0630436     .083821
------------------------------------------------------------------------------

.
```

```
. /* model (4) */
. reg call_back black high h_b, r

Linear regression                               Number of obs     =      4,870
                                                F(3, 4866)        =       6.61
                                                Prob > F          =     0.0002
                                                R-squared         =     0.0044
                                                Root MSE          =     .27157

------------------------------------------------------------------------------
             |               Robust
   call_back |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       black |  -.0231023   .0105901    -2.18   0.029    -.0438636    -.002341
        high |   .0229478   .0119584     1.92   0.055     -.000496    .0463917
         h_b |  -.0177808   .0155605    -1.14   0.253    -.0482864    .0127248
       _cons |   .0849835   .0080133    10.61   0.000     .0692739    .1006931
------------------------------------------------------------------------------

.
. log close
      name:  <unnamed>
       log:  final_names.log
  log type:  text
 closed on:  17 Jun 2021, 11:34:02
-------------------------------------------------------------------------------------------------
```