ECON3150/4150: Introductory Econometrics – Postponed Exam Spring 2023

1. (60%) You are interested in estimating the effect of income on maternal smoking during pregnancy. In your dataset `smoke` equals 1 if smoked per day while pregnant (0 otherwise), `faminc` is 1988 family income (1000 USD), `motheduc` is mother's years of education:

```
##              mean     SD min max    N
## faminc   29.0422 18.737 0.5  65 1387
## motheduc 12.9358  2.377 2.0  18 1387
## smoke     0.1528  0.360 0.0   1 1387
```

You estimate the following models:

```
##                             reg1           reg2           reg3         logit1
## Dependent Var.:            smoke          smoke          smoke          smoke
##
## Constant          0.256 (0.019)  0.357 (0.040)  0.643 (0.057) -0.882 (0.133)
## faminc          -0.004 (0.0005)                                -0.033 (0.005)
## log(faminc)                     -0.067 (0.011) -0.035 (0.012)
## motheduc                                       -0.030 (0.004)
## _______________ _______________ ______________ ______________ ______________
## Family                      OLS            OLS            OLS          Logit
## S.E. type       Heteroske.-rob. Heterosk.-rob. Heterosk.-rob. Heterosk.-rob.
## SSR                      173.50         174.43         168.74             --
## Observations              1,387          1,387          1,387          1,387
```

a. Interpret the estimates and their standard errors in the 1st column (`reg1`).

b. Compute the predicted value for `faminc`=65 using the estimates in column 1 (`reg1`) and interpret the result.

c. Interpret the coefficient on `log(faminc)` in the 2nd column (`reg2`).

d. What is the significance level of the confidence interval (-.07184, -.06216) for the coefficient on `log(faminc)` in the 2nd column (`reg2`)?

e. Compute the R-squared of the regression in the 3rd column (`reg3`).

f. The third column (`reg3`) adds mother's education to the specification. Compute the correlation between `log(faminc)` and `motheduc`.

g. Perform a joint test of the null-hypothesis that the coefficients on `log(faminc)` and `motheduc` are zero in the 3rd column (`reg3`).

h. Use the logit estimates in column 4 (`logit`) to predict the outcome at the average family income and interpret the result.

i. Compute the marginal effect of income at the sample average using the logit estimates in column 4 (`logit`), and compare them to the OLS results in column 1 (`reg1`).

j. Someone tells you that if one is interested in estimating an average marginal effect of income on smoking that is causal, then the logit model is to be preferred over the linear probability model. What do you respond?

2. (40%) Gibson & Shrader (Review of Economics & Statistics, 2018) use American data to estimate the effect of sleep on productivity. Information on the average night-time sleep (hours per week) comes from a time use survey and productivity is measured by earnings and wages. To account for potential omitted variable bias Gibson & Shrader implement an instrumental variable approach. Their instrumental variable is annual average sunset time (hours), and is expected to affect the time when people go to bed and thereby the amount they sleep. The following figure shows the variation – within US time zones as well as sharp changes in sunset time around the boundaries of time zones – that is used in the estimation.

In their estimation Gibson & Shrader adjust for both geographic characteristics (coastal distance and latitude) and demographics (gender, age, race and occupation shares, plus population density). The following table summarizes some of their main estimation results for log(earnings):

	First-stage Sleep	Reduced-form log(earnings)
Sunset time	-0.93	-0.045
	(0.28)	(0.017)
Mean dep. var.	57.9	6.67

a. Compute the IV estimate of the effect of sleep on log(earnings) and interpret the result.

b. If the IV estimate of sleep on log(wage) is 0.083, what is then the effect of sunset time on log(wage)?

c. What could violate the instrument's exclusion restriction? Do the controls (hint) help in this respect? Explain.

d. Is the instrument relevant? Motivate your answer.

e. Explain how you could use the variation highlighted in the figure above to estimate the effect of sleep on earnings using a regression discontinuity design. Make clear whether you would use a sharp or a fuzzy design.

The Cumulative Standard Normal Distribution Function, $\Pr(Z \le z)$

```
## Rows denote 1st decimal value of z, and columns 2nd decimal value of z

## So for example, P(Z <= 0.22) = 0.5871

##            0      1      2      3      4      5      6      7      8      9
## 0.0 : 0.5000 0.5040 0.5080 0.5120 0.5160 0.5199 0.5239 0.5279 0.5319 0.5359
## 0.1 : 0.5398 0.5438 0.5478 0.5517 0.5557 0.5596 0.5636 0.5675 0.5714 0.5753
## 0.2 : 0.5793 0.5832 0.5871 0.5910 0.5948 0.5987 0.6026 0.6064 0.6103 0.6141
## 0.3 : 0.6179 0.6217 0.6255 0.6293 0.6331 0.6368 0.6406 0.6443 0.6480 0.6517
## 0.4 : 0.6554 0.6591 0.6628 0.6664 0.6700 0.6736 0.6772 0.6808 0.6844 0.6879
## 0.5 : 0.6915 0.6950 0.6985 0.7019 0.7054 0.7088 0.7123 0.7157 0.7190 0.7224
## 0.6 : 0.7257 0.7291 0.7324 0.7357 0.7389 0.7422 0.7454 0.7486 0.7517 0.7549
## 0.7 : 0.7580 0.7611 0.7642 0.7673 0.7704 0.7734 0.7764 0.7794 0.7823 0.7852
## 0.8 : 0.7881 0.7910 0.7939 0.7967 0.7995 0.8023 0.8051 0.8078 0.8106 0.8133
## 0.9 : 0.8159 0.8186 0.8212 0.8238 0.8264 0.8289 0.8315 0.8340 0.8365 0.8389
## 1.0 : 0.8413 0.8438 0.8461 0.8485 0.8508 0.8531 0.8554 0.8577 0.8599 0.8621
## 1.1 : 0.8643 0.8665 0.8686 0.8708 0.8729 0.8749 0.8770 0.8790 0.8810 0.8830
## 1.2 : 0.8849 0.8869 0.8888 0.8907 0.8925 0.8944 0.8962 0.8980 0.8997 0.9015
## 1.3 : 0.9032 0.9049 0.9066 0.9082 0.9099 0.9115 0.9131 0.9147 0.9162 0.9177
## 1.4 : 0.9192 0.9207 0.9222 0.9236 0.9251 0.9265 0.9279 0.9292 0.9306 0.9319
## 1.5 : 0.9332 0.9345 0.9357 0.9370 0.9382 0.9394 0.9406 0.9418 0.9429 0.9441
## 1.6 : 0.9452 0.9463 0.9474 0.9484 0.9495 0.9505 0.9515 0.9525 0.9535 0.9545
## 1.7 : 0.9554 0.9564 0.9573 0.9582 0.9591 0.9599 0.9608 0.9616 0.9625 0.9633
## 1.8 : 0.9641 0.9649 0.9656 0.9664 0.9671 0.9678 0.9686 0.9693 0.9699 0.9706
## 1.9 : 0.9713 0.9719 0.9726 0.9732 0.9738 0.9744 0.9750 0.9756 0.9761 0.9767
## 2.0 : 0.9772 0.9778 0.9783 0.9788 0.9793 0.9798 0.9803 0.9808 0.9812 0.9817
## 2.1 : 0.9821 0.9826 0.9830 0.9834 0.9838 0.9842 0.9846 0.9850 0.9854 0.9857
## 2.2 : 0.9861 0.9864 0.9868 0.9871 0.9875 0.9878 0.9881 0.9884 0.9887 0.9890
## 2.3 : 0.9893 0.9896 0.9898 0.9901 0.9904 0.9906 0.9909 0.9911 0.9913 0.9916
## 2.4 : 0.9918 0.9920 0.9922 0.9925 0.9927 0.9929 0.9931 0.9932 0.9934 0.9936
## 2.5 : 0.9938 0.9940 0.9941 0.9943 0.9945 0.9946 0.9948 0.9949 0.9951 0.9952
## 2.6 : 0.9953 0.9955 0.9956 0.9957 0.9959 0.9960 0.9961 0.9962 0.9963 0.9964
## 2.7 : 0.9965 0.9966 0.9967 0.9968 0.9969 0.9970 0.9971 0.9972 0.9973 0.9974
## 2.8 : 0.9974 0.9975 0.9976 0.9977 0.9977 0.9978 0.9979 0.9979 0.9980 0.9981
## 2.9 : 0.9981 0.9982 0.9982 0.9983 0.9984 0.9984 0.9985 0.9985 0.9986 0.9986
```

Critical Values for the $F_{m,\infty}$ Distribution

```
## Rows denote degrees of freedom (m), and columns significance level (%)

##          10%     5%     1%

##
##   1 :  2.7055 3.8415 6.6349
##   2 :  2.3026 2.9957 4.6052
##   3 :  2.0838 2.6049 3.7816
##   4 :  1.9449 2.3719 3.3192
##   5 :  1.8473 2.2141 3.0173
##   6 :  1.7741 2.0986 2.8020
##   7 :  1.7167 2.0096 2.6393
##   8 :  1.6702 1.9384 2.5113
##   9 :  1.6315 1.8799 2.4073
##   10 :  1.5987 1.8307 2.3209
##   11 :  1.5705 1.7886 2.2477
##   12 :  1.5458 1.7522 2.1847
##   13 :  1.5240 1.7202 2.1299
##   14 :  1.5046 1.6918 2.0815
##   15 :  1.4871 1.6664 2.0385
##   16 :  1.4714 1.6435 2.0000
##   17 :  1.4570 1.6228 1.9652
##   18 :  1.4439 1.6038 1.9336
##   19 :  1.4318 1.5865 1.9048
##   20 :  1.4206 1.5705 1.8783
##   21 :  1.4102 1.5557 1.8539
##   22 :  1.4006 1.5420 1.8313
##   23 :  1.3916 1.5292 1.8104
##   24 :  1.3832 1.5173 1.7908
##   25 :  1.3753 1.5061 1.7726
##   26 :  1.3678 1.4956 1.7554
##   27 :  1.3608 1.4857 1.7394
##   28 :  1.3541 1.4763 1.7242
##   29 :  1.3478 1.4675 1.7099
##   30 :  1.3419 1.4591 1.6964
```