Name: _____ Spring 2018

EC200 Econometrics and Applications

Unit 1 Quiz

You have 75 minutes to complete this quiz. There are 36 total points. Please round to three decimal places when necessary. Note that some sub-questions build on previous sub-questions; if you get stuck on one part, at least make a guess so have you have values to apply to the following parts - you won't be penalized a second time if the numbers you carry over are incorrect.

1. *(8 points)* We've spent a decent amount of time talking about the zero conditional mean assumption behind our ordinary least squares estimates, that $E[u_i|X_i] = 0$. Consider one of two research questions:

   - A: What is the relationship between city-level crime rates and city-level median household income in 2015?
   - B: What is the relationship between quarterly CO_2 emissions and quarterly GDP in the United States, from 1980-2014.

   (a) For *each* research question, identify the dependent variable and independent variable. *[2 points]*

A: B:

   (b) Select question A *or* B and provide an example of a variable that would be captured in the error term but *would not* violate the zero conditional mean assumption. Explain your reasoning in one sentence. *[3 points]*

   (c) Select question A *or* B (same or different from your choice in part b) and provide an example of a variable that would be captured in the error term but *would* violate the zero conditional mean assumption. Explain your reasoning in one sentence. *[3 points]*

2. *(14 points)* Consider the relationship between annual per-capita cheese consumption (*cheese*, dependent variable) and unemployment rates (*unemp*, independent variable). You collect the following data:

Year	2005	2006	2007	2008
Per-capita cheese consumption (pounds):	31.7	32.6	33.1	32.7
Unemployment rate (percent):	4.9	4.4	5.0	7.3

(a) Write a population model for the relationship between per-capita cheese consumption and unemployment rates. *[2 points]*

(b) Estimate $\widehat{\beta_0}$ and $\widehat{\beta_1}$. *[4 points]*

(c) Interpret $\widehat{\beta_0}$ and $\widehat{\beta_1}$, making sure to include appropriate units. *[2 points]*

(d) Calculate the residual for 2008. What does it mean? *[2 points]*

(e) Calculate R^2. How much does unemployment explain per-capita cheese consumption? *[4 points]*

3. *(14 points)* Consider the following summary statistics and regression results from a nationally representative random sample of 1,172 mothers surveyed in the 2016 General Social Survey.

```
.
. sum _hrs1  _doesnthurt childs  educ age
    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
       _hrs1 |      1,172    19.40614    21.28042         0         89
 _doesnthurt |      1,172    .4974403    .5002069         0          1
      childs |      1,172    2.547782    1.376366         1          8
        educ |      1,172    13.43686    2.996486         0         20
         age |      1,172    52.71672    16.84845        19         89
.
```

where variables are defined as follows:

- `hrs1` = hours worked last week
- `_doesnthurt` = 1 if agrees that it doesn't hurt children for their mothers to work outside of home, 0 otherwise
- `childs` = number of children
- `educ` = years of completed education
- `age` = age in years

```
.
. regress _hrs1  _doesnthurt childs  educ age if _female == 1 & childs > 0
      Source |       SS           df       MS      Number of obs   =     1,172
-------------+----------------------------------   F(4, 1167)      =     70.12
       Model |  102760.201         4  25690.0502   Prob > F        =    0.0000
    Residual |  427534.475     1,167  366.353449   R-squared       =    0.1938
-------------+----------------------------------   Adj R-squared   =    0.1910
       Total |  530294.676     1,171  452.856256   Root MSE        =     19.14

------------------------------------------------------------------------------
       _hrs1 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
 _doesnthurt |   .5572165   1.120283     0.50   0.619    -1.640778     2.75521
      childs |  -.8703755    .426837    -2.04   0.042    -1.707829    -.0329217
        educ |   1.146882   .1924429     5.96   0.000     .7693097     1.524455
         age |  -.4831164   .0339084   -14.25   0.000    -.5496446    -.4165883
       _cons |    31.4043   3.456793     9.08   0.000     24.62208     38.18653
------------------------------------------------------------------------------
```

(a) What share of respondents believe that it doesn't hurt children for their mothers to work? *[2 points]*

answer in percents (0%-100%)

(b) Interpret $\widehat{\beta}_{childs}$, the coefficient on the number of children. Exactly what does it mean, in words? Is it statistically significant? Is it large? *[2 points]*

(c) Fill out the following table based on the regression results above. *[2 points]*

R^2		TSS	
ESS		SSR	
d.f.		SER	

(d) Under what assumptions would it be the case that $\widehat{\beta}_{childs}$ is BLUE? Which assumptions are likely to hold, which are not, and for which would you need more information? Give your reasoning for *each* assumption you list. *[4 points]*

(e) Consider two friends, Drithi and Ayako. They are both 33-year-old mothers who believe that when mothers work outside of the home, it could hurt their children. However, Drithi is a high school graduate (12 years completed) with 3 children, and Ayako is a college graduate (16 years completed) with 2 children. What is the *difference* in predicted number of work hours between them? *[2 points]*

(f) Marius says that since older people likely have had more children, $\widehat{\beta}_{childs}$, the impact of number of children on hours of work, is biased. Do you agree? Explain. *[2 points]*