Note that you are not allowed to collaborate. Please provide a self-contained report on the problems with maximum 10 pages (not a hard constraint), and provide further documentation in an appendix (without page limit). You can hand in the file as a single pdf-document.

Problem 1 (20 pct - about 2 pages of text):

Explain with as much depth as possible how we can construct standard errors and perform statistical inference when we use instrumental variable estimation techniques.

Hints: We want to know about 2sls asymptotics, the difference between 2sls standard errors and OLS standard errors from the second stage, the problem of weak instruments and robust inference based on the reduced form.

Problem 2 (10 pct - about 1 page of text):

In a recent machine learning textbook, Matt Taddy (2019) writes (in the context of diff-in-diff analyses and fixed effects models): "If the clustered standard errors are unintuitive, note that you can get a similar result by controlling for the DMA-specific revenue levels in the regression mean equation itself. That is, you replace equation (*) with a model including DMA-specific intercepts (i.e. what economists would call "DMA fixed effects"). Since we're now assuming independence conditional upon the DMA-specific levels, the usual regression standard errors apply." He adds in a footnote: "Adding fixed effects is not the same as clustering standard errors, however both are approaches to accounting for dependence within a DMA. One approach is not obviously superior to the other, and each has different advantages. For example, the approach of adding fixed effects is easily extensible to more complex dependence structures." That is, he is claiming that including fixed effects is an alternative to clustered standard errors and that when we use fixed effects we do not need clustered standard errors.

Please clarify!

Problem 3 (30 pct - about 3 pages of text, including tables/graphics)

Use data from Favara and Imbs (2015) (the same data that we worked with in the problems after Lecture 8). In section 3 of this paper, the authors study the relationship between financial deregulation and credit supply. In their models, the data for most of the variables appear in differenced form, that is, as year-on-year changes. Here, you will do a more standard difference-in-difference analysis of the data.

1. Describe briefly what happens if you pre-difference your data and then enter the data into a fixed effects model.

2. Please generate the level form of the variables in question and estimate a two-way fixed effects model relating financial deregulation and credit supply based on the same data as Table 2 in the paper. You can do this analysis without controls or lagged dependent variables and please use the "number of loans"-variable as the dependent variable. Note that it makes sense to use the variable "inter_bra" in not-lagged version. Do the analysis also for the placebo sample. Comment on the results.

3. The motivation for specifying the model with variables in differenced form in the original paper is to get rid of state-specific linear trends. Please estimate models with linear state-specific trends as a robustness check to the above analysis. Do the same analysis also for the placebo sample and comment on the results. In the following problems, use the models without linear state-specific trends.

4. Please use event-study techniques to study if there is a (one-period) anticipation effect to deregulation and to study the dynamic effects of deregulation (does all the effect of deregulation kick in immediately?) Specifically, estimate a model with three lags and one lead and comment on the results.

5. Provide a consistent estimate of the simple model from 2. above, but with a lagged dependent variable added to the model. Explain what the challenge is, and how you deal with this challenge.

Problem 4 (20 pct - about 2 pages of text)

Read Cengiz et al. (2019) and explain the basic empirical strategy in this paper. Discuss in particular what techniques the authors use to study how minimum wage changes not only affects the mean wages but also where in the wage distribution changes occur. Compare their approach to estimation of distributional effects to estimation using quantile regression techniques.

Problem 5 (20 pct - about 2 pages of text, including tables/graphs)

Use the data from Angrist and Evans (1998) on family size and mothers' labor supply (the same data that we worked with in the problems after Lecture 12). Apply the same-sex instrument to measure the effect of family size on the number of weeks worked. Explain the basic intuition behind this instrument. Estimate how the decision to have one more child beyond the second affects the distribution of weeks worked (beyond the mean) using several different techniques. Explain carefully how you do this: it may be possible to do well on this problem even if it is hard to get sensible numerical estimates. Compare the results to the standard instrumental variable estimates.

References

Angrist, Joshua and William Evans (1998) "Children and Their Parents' Labor Supply: Evidence from Exogenous Variation in Family Size," *American Economic Review*, Vol. 88, No. 3, pp. 450–77.

Cengiz, Doruk, Arindrajit Dube, Attila Lindner, and Ben Zipperer (2019) "The effect of minimum wages on low-wage jobs," *The Quarterly Journal of Economics*, Vol. 134, No. 3, pp. 1405–1454.

Favara, Giovanni and Jean Imbs (2015) "Credit supply and the price of housing," *American Economic Review*, Vol. 105, No. 3, pp. 958–92.

Taddy, Matt (2019) *Business data science: Combining machine learning and economics to optimize, automate, and accelerate business decisions*: McGraw Hill Professional.