Applied Microeconometrics I
Discussion of Problem Set 3

Atte Pudas

Aalto University

04/10/2023

Summary

- Average grade 71

Exercise 1

Electronic monitoring (EM) for people committing crimes and waiting for their trial to be completed is a policy alternative to imprisonment. From a theoretical perspective, it is unclear whether EM reduces the chances of recidivism compared to imprisonment. Those in favor of EM stress the fact that prison can be criminogenic, whereas those in favor of incarcerating individuals while waiting for their trial to end stress that EM does not discourage future criminal acts. Di Tella and Schargrodsky (2013) study whether EM is a useful deterrence policy in Argentina by exploiting rich information on judges who decide whether to assign EM or not.

Exercise 1.1

Answer the following questions [Max 100 words in total]
a) What is the causal relationship of interest?
b) Why do we need an instrument for electronic monitoring?
c) Can you think of another credible empirical strategy to estimate the causal relationship of interest?

- a) The effect of EM (vs. imprisonment) on recidivism
- b) It's not random who gets EM and who is imprisoned. EM may be targeted to those with lower expected likelihood of recidivism.
- c) An alternative to the judge leniency design (IV) is to randomize EM and run RCT. We could even make sure of having perfect compliance, since detainees can be forced to prison/use EM against their will.

  Unethical? Legally impossible?

Exercise 1.2

Explain, in plain English, what is the exclusion restriction here. [Max 30 words]

- Judge leniency should affect recidivism only through the EM use.
- In general, we can think that it holds

Exercise 1.3

Check that there is a first stage:
a) Regress electronicMonitoring on percJudgeSentToEM for judges that see 10 offenders or more. Report the coefficient.
b) Regress electronicMonitoring on percJudgeSentToEM and all other covariates (as per column 5 in Table 2) for judges that see 10 offenders or more. Report the coefficient.
c) Where does the difference in coefficients come from? [Max 75 words]

Exercise 1.3 a) and b)

- Run:

```
use electronic_monitoring.dta, clear

* Strings to numerical (more convenient in regressions)
encode mostSeriousCrime, g(serious)
encode judicialDistrict, g(districtfe)

// 1. 3
*a)
reg electronicMonitoring percJudgeSentToEM if offendersPerJudge>9
*b)
reg electronicMonitoring percJudgeSentToEM  judgeAlreadyUsedEM i.serious ///
age ageSquared argentine numberPreviousImprisonments i.year i.districtfe if offendersPerJudge>9
```

- Coefficient for percJudgeSentToEM a):

electronicMo~g	Coefficient	Std. err.	t	P>|t|	[95% conf.	interval]
percJudgeSen~M	.8220764	.0257175	31.97	0.000	.7716685	.8724843

- Coefficient for percJudgeSentToEM b):

electronicMo~g	Coefficient	Std. err.	t	P>|t|	[95% conf.	interval]
percJudgeSen~M	.6417401	.0362389	17.71	0.000	.5707096	.7127705

Exercise 1.3 c)

- The difference in estimates for "percJudgeSentToEM" is given by the court fixed effects and not by the additional covariates (which, once court fixed are accounted for, should be independent of the instrument).
- Further explanation:

  The randomization happens within district (it's randomized by lottery which judge happens to be in charge). Between districts, the "type" of a judge is not random (districts are different etc.)

  When we estimate the FS as we do here (and also when doing RF and 2SLS) we have to condition on the variables in correspondence of which the randomization occurred (in this case, judge court district).

  Once randomization is captured via the inclusion of the district FE's, other covariates only improve precision but are not necessary, in general, to get an unbiased estimate.

Exercise 1.4

- *Replicate the results in the last three rows in Table 3. How do you interpret these results? [Max 50 words]*

```
* 3rd to last row (serious crimes):
   *Predicted difference between means of low and high EM rate judges and p-value for the difference
       reg seriousCrimes lowHighEMRateDummy i.districtfe if offendersPerJudge>9, cluster(districtfe)
       local est_diff_in_means = _b[lowHighEMRateDummy]            //save coef. to a local
   *column 1 (mean for judges with low EM rate)
       sum seriousCrimes if e(sample)==1 & lowHighEMRateDummy == 0 //e(sample) gives the sample in regression
       local low_EM_mean = `r(mean)'                               //save mean to a local
   *column 2 (predicted mean for judges with high EM rate)
       local high_EM_mean = `low_EM_mean' + `est_diff_in_means'    //save the predict. diff to a local
       di `high_EM_mean'                                           //display
* 2nd to last row (middle crimes):
   *predicted difference between means of low and high EM rate judges and p-value for the difference
       reg middleCrimes lowHighEMRateDummy i.districtfe if offendersPerJudge>9, cluster(districtfe)
       local est_diff_in_means = _b[lowHighEMRateDummy]
   *column 1 (mean for judges with low EM rate)
       sum middleCrimes if e(sample)==1 & lowHighEMRateDummy == 0
       local low_EM_mean = `r(mean)'
   *column 2 (predicted mean for judges with high EM rate)
       local high_EM_mean = `low_EM_mean' + `est_diff_in_means'
       di `high_EM_mean'
* The last row (minor crimes):
   *predicted difference between means of low and high EM rate judges and p-value for the difference
       reg minorCrimes lowHighEMRateDummy i.districtfe if offendersPerJudge>9, cluster(districtfe)
       local est_diff_in_means = _b[lowHighEMRateDummy]
   *column 1 (mean for judges with low EM rate)
       sum minorCrimes if e(sample)==1 & lowHighEMRateDummy == 0
       local low_EM_mean = `r(mean)'
   *column 2 (predicted mean for judges with high EM rate)
       local high_EM_mean = `low_EM_mean' + `est_diff_in_means'
       di `high_EM_mean'
```

Exercise 1.4

- Putting all the parts together, we get:

Pretreatment Characteristic	Cases before Judges with Low EM Rate (1)	Cases before Judges with High EM Rate (2)	p-Value (3)
Serious crimes	.600	.623	.230
Middle crimes	.311	.297	.310
Minor crimes	.090	.081	.294

- Table 3 presents the observable characteristics of alleged offenders across liberal and conservative judges. The last three rows show no evidence of significant differences in how severe crimes the two groups of judges face (p-values for the differences in means are high).

Exercise 1.5

Check whether the reduced form is significant (use the specification in column 2 of Table 5). [Max 10 words]

- Run commands:

```
reg recidivism percJudgeSentToEM i.serious age ageSquared argentine numberPreviousImprisonments ///
    i.year i.districtfe if offendersPerJudge>9, cluster(districtfe)

global reduced_form = _b[percJudgeSentToEM] //save the RF estimate in a global to be used in part 1.6
```

- The reduced form estimate is significant at 5% level.

recidivism	Coefficient	Robust std. err.	t	P>|t|	[95% conf.	interval]
percJudgeSentToEM	-.4156035	.1832045	-2.27	0.037	-.8021312	-.0290758

Exercise 1.6

Replicate column 2 in Table 5 (second and first stage regressions). What is LATE according to your results? Verify that you can obtain the LATE by dividing the reduced-form coefficient by the first-stage coefficient. [Max 10 words]

- Run commands:

```
ivregress 2sls recidivism (electronicMonitoring=percJudgeSentToEM) i.serious age ///
        ageSquared argentine numberPreviousImprisonments i.year i.districtfe ///
        if offendersPerJudge>9, first cluster(districtfe)  //choose option "first" to show FS

di ${reduced_form}/3.08845 // 2sls estimate = reduced form/first stage
```

- LATE:

recidivism	Coefficient	Robust std. err.	z	P>|z|	[95% conf.	interval]
electronicMonitoring	-.134567	.0553502	-2.43	0.015	-.2430514	-.0260827

- FS:

electronicMonitoring	Coefficient	Robust std. err.	t	P>|t|	[95% conf.	interval]
percJudgeSentToEM	3.08845	.3331948	9.27	0.000	2.434859	3.742041

- RF/FS: $-0.41563/3.08845 \approx -0.1346$

Exercise 2

Gerber et al. (2003) study whether voting in one election increases one's propensity to vote in the future. The authors document that the persistence in voting behavior ranks among the most robust empirical generalizations in political science. The formation and the relevance of habits in voters is however still largely unknown, and the authors seek to fill this gap in the literature.

Exercise 2.1

Suppose that we run a regression of current turnout on past voting controlling for age, race, education, and sex. Can we credibly identify the causal effect of voting in the latest election on the probability of voting in the current election? Discuss briefly why/why not. [Max 50 words]

- No. Endogeneity issue: for example, unobserved psychological traits such as sense of responsibility may explain both the past and present voting.

Exercise 2.2

Describe the experimental design in the paper and how the treatment effect of interest is identified. How is randomization used to solve the selection bias problem in this setting? [Max 75 words]

- Empirical strategy is RCT with imperfect compliance.
- Due to imperfect compliance (not everyone who is randomly encouraged to vote goes to vote), we can cast the analysis in an IV setting where we instrument "voting in 98" with the randomized treatment (being encouraged to vote). The voting behavior in 1999 (the outcome of interest) is then regressed on the instrumented voting behavior in 1998.
- **Dependent variable:** voting in 1999
  Endogenous variable: voting in 1998
  Instrument: randomized encouragement to vote in 1998

Exercise 2.3

Answer the following questions [Max 100 words in total]
a) A valid instrument should fulfill the exclusion restriction. What does this mean here?
b) Do you think that it is plausible that the instrument in the paper fulfills this criterion? Provide an example of one possible violation in the paper setting.

- a) Encouragement to vote should affect the voting behavior in 1999 only through its effect on the voting behavior in 1998.
- b) Maybe the encouragement in 1998 also affected directly in 1999 elections if people remembered the encouragement.

  Or maybe voting encouragement made their own past voting behavior salient to people, reinforcing the past habit of going/not going to vote, and therefore the encouragement would have an effect on the 1999 voting also via pre-1998 voting behavior.

Exercise 2.4

Can you prove empirically whether an instrument is valid? Can you prove empirically whether an instrument is not valid? [Max 50 words]

- You can prove that instrument is not valid, but you can't prove that it's valid.
- Strength of the first stage can be tested (F-test for the instruments).
- You may be able to show that exogeneity is violated (test whether instrument correlates with predetermined variables).
- Exclusion restriction cannot be shown to hold, but you may be able to show that it's violated.

Exercise 2.5

Explain the following terms [Max 100 words in total]
a) first stage
b) reduced form
c) 2SLS.
When you explain the terms, use the setting provided in the paper (explain the relevant variables and how they enter in the equations of interest. You don't need to list all the control variables).

- a) First stage: regress voter turnout in 1998 on the voting encouragement treatments.
- b) Reduced form: regress voter turnout in 1999 on the voting encouragement treatments.
- c) 2SLS: regresses voter turnout in 1999 on the predicted voter turnout in 1998 from the first stage.

Exercise 2.6

Answer the following questions [Max 75 words in total]
a) Who are the compliers in the setting?
b) Do you think that the causal estimates provided in the paper are policy relevant?

- a) The compliers are those who voted in 1998 because encouraged to do so, and who would not have voted otherwise.
- b) The causal estimate is LATE and it's policy relevant if we think that those at the margin of going/not going to vote are an important group to encourage.

Exercise 3

Edin et al. (2003) are interested in studying the economic consequences of living in ethnic "enclaves" in the Swedish context. Whether the geographical concentration of immigrants helps them in terms of socioeconomic outcomes is highly debated and is fundamentally an empirical question: theory can easily rationalize both why ethic concentration might be good and why it might be detrimental for newly arrived immigrants.

Read the paper by Edin et al. (2003) and answer to the following questions.

Exercise 3.1

Answer the following [Max 100 words in total]
a) What do the authors mean with "sorting"?
b) Discuss what the endogeneity problem is that the authors try to solve and why an IV strategy helps to solve it.

- a) Self-locating in different areas is sorting.
- b) Individuals endogenously locate in different areas based on many reasons. IV uses conditionally random (conditional on observed characteristics, independent of unobserved immigrant characteristics) initial location of immigrants as a source of variation to contemporary location. Immigrants can't be forced to stay in one place, and therefore, initial location is used as an instrument.

Exercise 3.2

Answer the following [Max 100 words in total]
a) How would you formulate (in words) the exogeneity assumption in this setting?
b) Propose a potential violation of this assumption.

- a) Initial location is independent of all unobservable characteristics (conditional on observed characteristics) that may also affect the outcome. In other words, initial location is as good as random.
- b) Maybe the most determined or high-ability immigrants could affect their initial location in some way (persuasion?), and these unobserved factors also correlate with labor market outcomes.

Exercise 3.3

Who are the compliers in this setting? [Max 30 words]

- Those who stay on a certain location only if induced to do so by the assignment policy, otherwise not.

Exercise 3.4

What does post-assignment mobility mean for the interpretation of results? [Max 50 words]

- If many individuals move away from the place where they were initially assigned, the first stage gets weaker. Weak first stage biases the IV estimates.

Exercise 3.5

Answer the following questions [Max 100 words in total]
a) What is the exclusion restriction in this setting?
b) Is it likely to hold? Why/why not?

- a) Conditionally random initial location affects labor market outcomes only trough contemporary location.
- b) Maybe not. Initial location may have an impact to labor market outcomes that is not channeled through contemporary locations. This could occur due to the shock of poor initial conditions of the place (e.g., high local unemployment rate in the area), which might negatively affect people's early employment; this might have lasting effects that show up in the outcomes measured 8 years later.

Exercise 3.6

Assume that the randomization holds so that the instrument is exogenous. Discuss what the
a) reduced form
b) 2SLS equations
are in this setting and whether they have a causal interpretation (without further assumptions). [Max 100 words in total]

- a) Regression of dependent variable (labor market outcomes) on the instrument (ethnic concentration in the initial location)
- b) Regression of dependent variable on predicted ethnic concentration in contemporary location, predicted with the instrument.
- Under exogeneity, the reduced form has a causal interpretation, whereas the 2SLS in general does not (we additionally need the exclusion restriction to hold).