A. Torgovitsky

Lee, Miguel and Wolfram (2019)

Problem

This question is about Lee et al. (2019). The paper, as well as the data and code used in the paper, is available here: https://doi.org/10.1086/705417. A cleaner version of the data set is available on Canvas together with a data set containing descriptions of the variables.

(a) Replicate columns (1)–(3) rows C3 and D2 of Table 3. Note that the covariates the authors use are (in the cleaned data set): `busia`, `market transearly`, `connected_rate`, `population`, `female`, `age_resp`, `educ`, `bank`, `housing`, `asset_value`, `energyspending`, `round2`. Either use clustered standard errors like the authors do, or use unclustered but heteroskedasticity–robust standard errors.

(b) Discuss the difference between the ATE, ATT, ATU, and LATE(s) in the context of this empirical application. What types of counterfactual questions would each of these target parameters answer?

(c) Why do you think the authors label column (3) as "TOT?" Does this interpretation make sense? Discuss.

(d) Show that the results in (a) do not change much by not controlling for covariates.

(e) Estimate the ATE, ATU, and the ATT using the marginal treatment effect framework. Leave out covariates as in (d) and restrict your attention to parametric specifications of the MTR that are point identified. Use the bootstrap to produce either standard errors (clustered or unclustered) or confidence intervals. Examine the sensitivity of your results to the parameterization of the MTR.

(f) Repeat (e) while linearly controlling for the same set of covariates used in (a). Investigate the difference between specifications in which the MTR functions are nonseparable in covariates, and those in which it is separable in covariates. How does separability affect the range of specifications that are point identified?

(g) Consider a policy where the effective connection price is changed to \$200 for everyone. Construct an estimate of the (per-person) policy-relevant treatment effect of this new policy relative to the status quo, using the

two outcomes in (a) and parametric, point-identified specifications of the MTR, as in (e). (You may omit covariates if you like.) Use the bootstrap to produce either standard errors (clustered or unclustered) or confidence intervals. Examine the sensitivity of your results to the parameterization of the MTR.

(h) Pick a prior lower and upper bounds for the two outcomes in (a) and discuss your choices. Compute nonparametric bounds on the ATU that use all of the information in $\mathbb{E}[Y_i|D_i, Z_i]$, where D_i is the treatment (connection) and Z_i is the subsidy, and covariates are omitted.

(i) Parameterize the MTR function using the Bernstein polynomials, omiting covariates, and re-compute bounds on the ATU using the same prior outcome bounds as in (h). Examine how the bounds change with the flexibility of the parameterization of the MTR function.

References

Lee, K., E. Miguel, and C. Wolfram (2019): "Experimental Evidence on the Economics of Rural Electrification," *Journal of Political Economy*, 128, 1523–1565.