Assignment 6

Daniel Campos
Eric Gutierrez
Samuel Fraley

November 17, 2025

Contents

Question 1: Randomized Control Trials

This question is based on Duflo, Hanna, and Ryan (2012), who evaluate whether teacher monitoring combined with financial incentives can reduce teacher absenteeism and improve learning in primary schools.

The NGO Seva Mandir operates non-formal primary schools in rural villages of Rajasthan (India). Before the program, teacher absenteeism was high (around 35%). In 2003, Seva Mandir introduced a teacher incentive program in 57 randomly selected schools. A camera system was installed to monitor teacher attendance, and teachers were paid according to a nonlinear function of valid teaching days (at least 5 hours of teaching with at least 8 students).

The program generated an immediate and persistent improvement in attendance in treated schools.

Data

The dataset `ps1_q1.csv` is a simplified version of the original data collected for this RCT. Each observation corresponds to a visit to one of the study schools (identified by `schid`). The variable `time` equals 1 in the month before the program starts (baseline) and is greater than 1 in months after the program begins. Schools are randomly assigned to a treatment group (`treat`=1) or control group (`treat`=0). The main outcome variables are the number of students (`students`) and teacher attendance (`teacher_attendance`).

1.1 Baseline and Experiment Integrity

Under proper randomization, potential outcomes are independent of treatment status:

$$Y_{1i}, Y_{0i} \perp\!\!\!\perp D_i,$$

which implies that the average treatment effect on the treated (ATT) and the average treatment effect (ATE) coincide: $\alpha_{ATT} = \alpha_{ATE} = \beta$.

Before analyzing post-treatment outcomes, we should verify that treated and control schools are similar at baseline.

(a) Using only the observations from the month before the program starts (`time` = 1), compute the average teacher attendance and the average number of students per classroom separately for the treatment and control groups.

Replicate the information in Panels A and B of Table 1 in Duflo et al. (2012): produce baseline means by treatment status, report appropriate standard errors, and comment on whether the randomization appears to have produced comparable groups.

(b) Briefly discuss whether any observed baseline differences are economically and/or statistically significant. Explain why this check is important for interpreting the causal effect estimates later.

1.2 Results

When randomization is valid, the treatment effect can be estimated via the simple regression:

$$Y_i = \alpha + \beta D_i,$$

where Y_i is the outcome of interest (e.g. teacher attendance) and D_i is the treatment indicator.

(c) Using post-program data (`time` > 1), compute and compare average teacher attendance in treated and control schools. Replicate the first three columns of Panel A of Table 2 in Duflo et al. (2012), reporting the relevant means and differences.

(d) Based on your estimates, discuss whether the incentive program achieved its main goal of reducing teacher absenteeism. Comment on the magnitude and statistical significance of the estimated treatment effect.

Question 2: Matching

Jacobson, LaLonde, and Sullivan (1993, JLS) study earnings losses following job displacement. Using administrative data from Pennsylvania, they document that workers involved in mass employment reductions suffer long-term earnings losses of roughly 25% per year. They distinguish between separations due to mass layoffs and other separations, and use stayers as a control group.

Their model can be written as:

$$w_{it}^A = \mu_i + \sum_{k\geq -4}^{6} \phi_k L_{it}^k + \sum_{l\geq -4}^{6} \psi_l M_{it}^l + \beta' X_{it} + \rho_t + \varepsilon_{it},$$

where w_{it}^A is log annual earnings of worker i, L_{it}^k and M_{it}^l are sets of dummies indicating years relative to layoff and mass layoff, X_{it} is a vector of covariates, and ρ_t are time effects.

Couch and Placzek (2010, CP) revisit this question using matching estimators, arguing that displaced workers are systematically selected, so estimates based only on JLS-type comparisons may be biased upward.

Let $D_i = 1$ if worker i is displaced (due to a mass layoff or other separation) and $D_i = 0$ otherwise, and let $p(X_i)$ denote the propensity score. The average treatment effect on the treated (ATT) is:

$$\alpha_{TT} = \mathbb{E}\Big[\mathbb{E}[w_{1i}^A \mid D_i = 1, p(X_i)] - \mathbb{E}[w_{0i}^A \mid D_i = 0, p(X_i)] \;\Big|\; D_i = 1\Big].$$

To compare outcomes relative to a reference year t_0, CP consider a differenced ATT:

$$\alpha_{ATT}^D = \mathbb{E}\Big[\big(\mathbb{E}[w_{1it}^A \mid D_i = 1, p(X_i)] - \mathbb{E}[w_{1it_0}^A \mid D_i = 1, p(X_i)]\big) - \big(\mathbb{E}[w_{0it}^A \mid D_i = 0, p(X_i)] - \mathbb{E}[w_{0it_0}^A \mid D_i = 0, p(X_i)]\big) \;\Big|\; D_i = 1$$

Your task is to revisit CP's findings using a different dataset.

Data

The dataset `ps1_q2.dta` is built from the Veneto Workers Histories (VWH), an administrative panel including all individuals working in the Italian region of Veneto from 1975–2001. The file `ps1_q2.dta` contains a subsample of workers who, in 1999, either:

- experienced a mass employment reduction,
- separated from the firm without being part of a mass layoff, or
- stayed with the same employer.

The panel covers the years 1995–2001. Mass layoffs are defined using the endogenous separation rate, following JLS and von Wachter, Song, and Manchester (2009). Displaced workers satisfy the standard requirements in this literature.

2.1

By computing the propensity score using gender, 1995 earnings decile, and decade of birth, the sample is divided into 9 blocks, ensuring that within each block the mean propensity score is not different between the control and treatment groups, that is, between non-displaced and displaced workers. However, and as the output indicates, the balancing property is not satisfied, which means that within the blocks created, the covariates are not balanced. Several variables differ significantly between the treatment and control groups, and consequently, the model is not properly accounting for systematic, idiosyncratic, differences between both groups.

2.2

The Nearest-Neighbor (NN) method matches individuals in the treatment group with the closest individual in the control group, using the observed values of the latter as a counterfactual to compute the αTT of the former. The Kernel estimator, instead, takes into account all the individuals in the control group, downweighting those that are farther from any given individual in the treatment group. Since both methods operate differently, we expect them to yield different results, as it is confirmed by the results obtained, which are detailed in the following table:

Year	**NN Estimate**	**NN SE**	**Kernel Estimate**	**Kernel SE**
1996	-421.06	2098.80	-554.54	150.52
1997	352.93	2226.50	-648.94	232.19
1998	-293.59	2423.26	-949.65	295.29
1999	-1700.42	2680.80	-1884.44	312.05
2000	-2440.29	2713.74	-2596.17	330.81
2001	1088.11	2714.94	-1046.38	395.99

Overall, the estimates obtained from both methods differ significantly, as demonstrated by the wide gap between the estimated αTT for the year 2001. The estimates from the Kernel method are significantly more precise than those from the NN method, as can be observed by their corresponding standard error measures. Consequently, we consider the estimated obtained through the Kernel method to be more reliable.

2.3

Compared to Figure 1, the estimates obtained through NN and Kernel matching follow similar trends, with both figures showing declining earnings after displacement. This confirms that the results from the matching estimators replicate the patterns observed in Figure 1. Additionally, our graph also remarks the reliability of the Kernel method, and highlights the volatility of NN estimates.

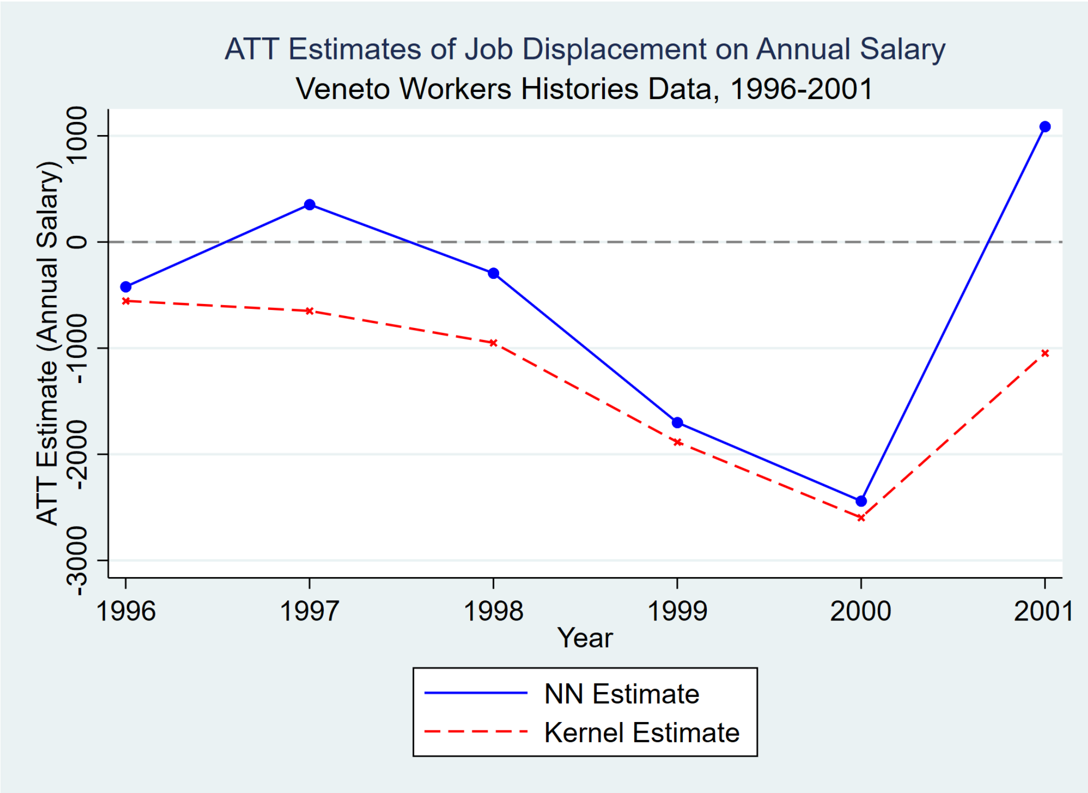

Figure 1: Matching estimates using NN and Kernel methods.

Question 3: Instrumental Variables

Angrist and Evans (1998) use an IV strategy to analyze how the number of children affects parents' labor supply. They find a sizable negative effect for mothers and essentially no effect for fathers. Here we focus on mothers and on employment (rather than hours worked).

Data

The dataset `ps1_q3.dta` is a subset of the data used by Angrist and Evans (1998) and contains only mothers. The key variables are:

- `sexk`: sex of the first child,
- `kidcount`: total number of children,
- `agem`: age of the mother,
- `twin_latest`: indicator equal to 1 if the last birth was a twin birth,
- `blackm`, `hispm`, `othracem`: race dummies,
- `workedm`: indicator equal to 1 if the mother is employed.

3.1 Baseline Models

Consider the model:

$$y_i = \beta_0 + \beta_1 \text{kidcount}_i + X_i'\beta + \varepsilon_i,$$

where y_i is the mother's employment status and X_i is a vector of controls.

(a) Estimate this equation using OLS.

Table 1: OLS regression

	Coefficient	Std. err.	z-stat	p-value	95 conf. low	95 conf. high
KIDCOUNT	-0.0910744	0.0009621	-94.67	0.000	-0.0929600	-0.0891888
AGEM	0.0146906	0.0002214	66.36	0.000	0.0142567	0.0151244
blackm	0.1506704	0.0023869	63.12	0.000	0.1459921	0.1553487
hispm	-0.0083050	0.0045233	-1.84	0.066	-0.0171705	0.0005605
othracem	0.0275062	0.0046310	5.94	0.000	0.0184296	0.0365828
Constant	0.3376580	0.0068619	49.21	0.000	0.3242089	0.3511071
Number of obs	4.00e+05					
R-squared	0.0343846					
Adj R-squared	0.0343725					
F-statistic	2.85e+03					
Root MSE	0.4870991					

Figure 1: OLS regression output from Stata

(b) Estimate the same specification using a probit model.

Table 2: Probit regression

	Coefficient	Std. err.	z-stat	p-value	95 conf. low	95 conf. high
workedm						
KIDCOUNT	-0.2370727	0.0025496	-92.99	0.000	-0.2420698	-0.2320757
AGEM	0.0382601	0.0005794	66.03	0.000	0.0371244	0.0393957
blackm	0.4039446	0.0064219	62.90	0.000	0.3913579	0.4165314
hispm	-0.0203433	0.0117697	-1.73	0.084	-0.0434116	0.0027249
othracem	0.0717805	0.0120918	5.94	0.000	0.0480810	0.0954800
Constant	-0.4263508	0.0178661	-23.86	0.000	-0.4613677	-0.3913339
Number of obs	4.00e+05					
Log likelihood	-2.67e+05					
LR chi2(5)	1.40e+04					
Prob > chi2	0.0000000					
Pseudo R2	0.0255562					

Figure 2: Probit regression output from Stata

(c) Discuss whether these approaches (OLS and probit) are appropriate for identifying the causal effect of the number of children on mothers' labor supply.

Just using OLS or probit alone is probably not the best approach to identifying the casual effect of number of children on labor supply due to endogeneity problems. For example, women with richer partners may not be required to work, and can decide to have more kids, thus showing that family income could influence both number of children and labor force participation. Various other endogeneity problems may arise as we have a relatively simple model (we already give household/family income, but what about access to childcare such as grandparents present, access to contraceptives, etc).

3.2 IV Probit

(d) Re-estimate the model using an IV probit specification, instrumenting `kidcount` with `twin_latest`.

Using twin_latest as an instrument for kidcount is a valid choice as having twins is arguably exogenous to the mother's labor supply decision. The occurrence of twins is largely random and not influenced by the mother's employment status or other socio-economic factors. Therefore, it satisfies the relevance condition (it affects the number of children) and the exclusion restriction (it does not directly affect the mother's employment status except through its effect on the number of children).

To test, we run two stage regression. The first stage regresses kidcount on twin_latest and other controls, and the second stage regresses workedm on the predicted values of kidcount from the first stage and other controls.

Table 3: First-stage regression

	Coefficient	Std. err.	z-stat	p-value	95 conf. low	95 conf. high
twin_latest	0.3850044	0.0099461	38.71	0.000	0.3655104	0.4044983
SEXK	0.0138442	0.0025263	5.48	0.000	0.0088927	0.0187958
AGEM	0.0309710	0.0003598	86.08	0.000	0.0302658	0.0316761
blackm	0.3235942	0.0038811	83.38	0.000	0.3159874	0.3312011
hispm	0.4370486	0.0073863	59.17	0.000	0.4225716	0.4515256
othracem	0.1210053	0.0075927	15.94	0.000	0.1061238	0.1358868
Constant	1.5577578	0.0110493	140.98	0.000	1.5361014	1.5794142
Number of obs	4.00e+05					
R-squared	0.0404864					
Adj R-squared	0.0404720					
F-statistic	2.81e+03					
Root MSE	0.7988634					

Here we see that twin_latest satisfies the relevance condition as it is statistically significant in predicting kidcount.

Table 4: IV Probit regression

	Coefficient	Std. err.	z-stat	p-value	95 conf. low	95 conf. high
workedm						
KIDCOUNT	-0.0715877	0.0413285	-1.73	0.083	-0.1525901	0.0094146
SEXK	0.0002138	0.0040526	0.05	0.958	-0.0077291	0.0081568
AGEM	0.0328903	0.0015186	21.66	0.000	0.0299138	0.0358667
blackm	0.3473656	0.0161035	21.57	0.000	0.3158033	0.3789279
hispm	-0.0915425	0.0210846	-4.34	0.000	-0.1328676	-0.0502174
othracem	0.0515402	0.0131292	3.93	0.000	0.0258073	0.0772730
Constant	-0.6801591	0.0638258	-10.66	0.000	-0.8052555	-0.5550628
KIDCOUNT						
SEXK	0.0138442	0.0025270	5.48	0.000	0.0088914	0.0187970
AGEM	0.0309710	0.0003467	89.33	0.000	0.0302914	0.0316505
blackm	0.3235942	0.0046597	69.45	0.000	0.3144614	0.3327271
hispm	0.4370486	0.0097855	44.66	0.000	0.4178694	0.4562278
othracem	0.1210053	0.0084547	14.31	0.000	0.1044344	0.1375763
twin_latest	0.3850044	0.0108025	35.64	0.000	0.3638318	0.4061769
Constant	1.5577578	0.0103840	150.01	0.000	1.5374055	1.5781101
/						
athrho2_1	-0.1318425	0.0328383	-4.01	0.000	-0.1962043	-0.0674807
lnsigma2	-0.2245740	0.0020035	-112.09	0.000	-0.2285007	-0.2206473
Number of obs	4.00e+05					
Log likelihood	-7.45e+05					
Wald chi2	5.17e+03					
Prob > chi2	0.0000000					
Pseudo R2						

In the IV probit, we see that the coefficient on kidcount is -0.0715877, which is larger in magnitude compared to the OLS and probit estimates. This suggests that the negative effect of having more children on mothers' employment is more pronounced when accounting for endogeneity using the IV approach.

3.3 Marginal Effects

(e) Using your preferred IV probit specification, estimate the marginal effect of an additional child on the probability that a mother is employed.

Table X reports the predicted probability of employment for mothers with 0–5 children, based on the IV probit estimates. The probability falls from about 0.63 for one child to 0.43 for three children, implying that an additional child reduces employment probability by roughly zz percentage points around this range (see also Figure Y).

	Pred. prob.	Std. err.	95 conf. low	95 conf. high
0 children	0.6355	0.0393	0.5584	0.7125
1 child	0.6086	0.0245	0.5605	0.6566
2 children	0.5812	0.0090	0.5635	0.5989
3 children	0.5534	0.0071	0.5395	0.5672
4 children	0.5253	0.0233	0.4796	0.5710
5 children	0.4971	0.0396	0.4194	0.5747
Observations	4.00e+05			

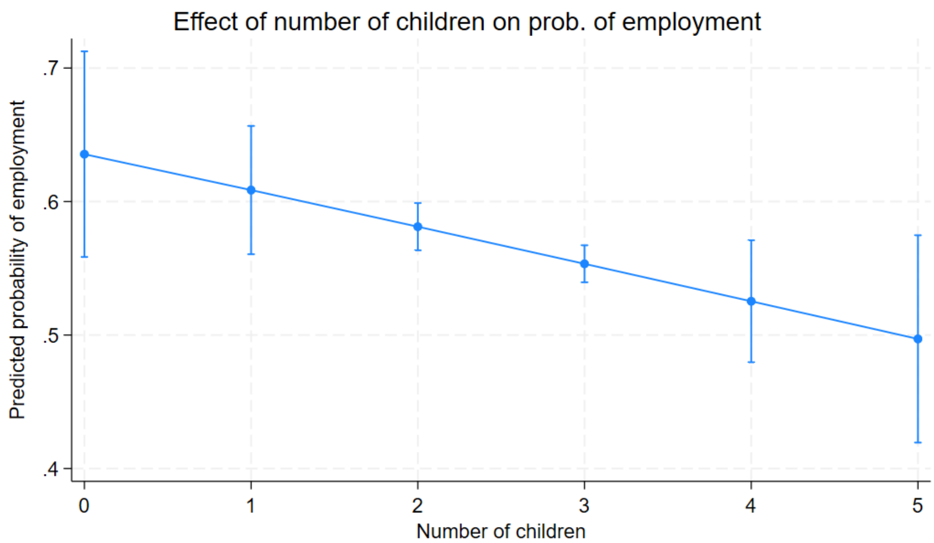

Figure 2: Predicted probability of maternal employment by number of children

References

- Angrist, J.D. and Evans, W.N. (1998). "Children and Their Parents' Labor Supply: Evidence from Exogenous Variation in Family Size." *American Economic Review*, 88(3): 450–477.
- Couch, K.A. and Placzek, D.W. (2010). "Earnings Losses of Displaced Workers Revisited." *American Economic Review*, 100(1): 572–589.
- Duflo, E., Hanna, R., and Ryan, S.P. (2012). "Incentives Work: Getting Teachers to Come to School." *American Economic Review*, 102(4): 1241–1278.
- Jacobson, L.S., LaLonde, R.J., and Sullivan, D.G. (1993). "Earnings Losses of Displaced Workers." *American Economic Review*, 83(4): 685–709.