Diff-in-diff

- Think of clinical trials to test the efficacy of a new drug
- i = individual, t = time
- Y_{it} = measure of health performance (e.g. temperature of the body)
- T_{it} = dummy for treatment (i.e. whether the individual is administered the drug)
- e_{it} = stochastic error term
- S7_Data_example.xlsx
- *(see last slide for xtreg command in Stata)*

Key ingredients

- In a previous period (t = 0) no one is administered the drug
- In a second period (t = 1) a random sample of the patients is administered the drug
- If so, estimate (by OLS)

$$Y_{it} = \text{alpha} + \text{beta}*T_{it} + e_{it}$$

- Beta is the diff-in-diff estimator
- Suppose individual i is treated while individual j is not:

$$Y_{i1} - Y_{i0} - (Y_{j1} - Y_{j0}) = \text{beta}$$

- Beta gives you an estimate of whether the treatment works

To be sure …

$$Y_{i1} - Y_{i0} - (Y_{j1} - Y_{j0}) = \text{beta} + e_{i1} - e_{i0} - (e_{j1} - e_{j0})$$

- Beta is the average treatment effect (ATE) if and only if $E[e_{i1} - e_{i0} - (e_{j1} - e_{j0})] = 0$
- Otherwise, the ATE depends on unobserved effects produced by the unobserved errors
- In practice, you must have randomized well between treatment and control
  - It does not matter that the two subsample are composed of different people, but they cannot be different people in a systematic way (e.g. by gender, age, else)

Challenges to randomization (I)

- The smaller the sample the less likely that you have randomized properly

Challenges to randomization (II)

- Use individual fixed effects if you believe that individuals are systematically different

$$Y_{it} = alpha + beta*T_{it} + FE_i + e_{it}$$

- What are fixed effects?
  - Dummies 0/1 for each individual
- What do they do?
  - They compute an estimator of the ATE that does not take into account potential systematic differences across individuals that do not vary over time (within estimator)
- You could also emply time fixed effects

$$Y_{it} = alpha + beta*T_{it} + FE_i + FE_t + e_{it}$$

- *If you really believe your randomization, you need no controls other than (possibly) FE*

Challenges to randomization (III)

- As opposed to a simple diff a diff-in-diff assumes that people could be different to start with
- Otherwise if you are sure that you randomized properly you can just go with the cross-section

$$Y_i = \text{alpha} + \text{beta}*T_i + e_i$$

- But a fully random sample to start with could be a big "if"

Generalize to panel data

- You can have more than just two periods

$$Y_{it} = \text{alpha} + \text{beta}*T_{it} + Fe_i + Fe_t + e_{it}$$

- In which case, T_{it} could be
  - a dummy variable taking value 1 for the treated group after some period
  - a continuous variable that undergoes some random shock for part of the sample after some period

Types of shocks/experiments

- Lab experiments: you generate the variation in a lab controlled session
  - Relatively easy to implement (if you have a lab and subjects)
  - Not the real world
- Natural experiment: unexpected shock that hits some part of your sample at some moment in time
  - Ideal, but may be hard to find for your question at hand
- Field experiment: you generate the variation yourself in the real world
  - Good, but (very) costly to make, and there can be ethical issues

- *Important*: shock has to be unanticipated by the agents in your sample (that's not a problem with clinical trials, it is a problem in social sciences)

Examples

- Field experiments:
  - Management matters (Bloom et al., 2013)
  - Social preferences and incentives (Bandiera et al., 2005)

- Natural experiment:
  - Chatterji & Fabrizio (2013)

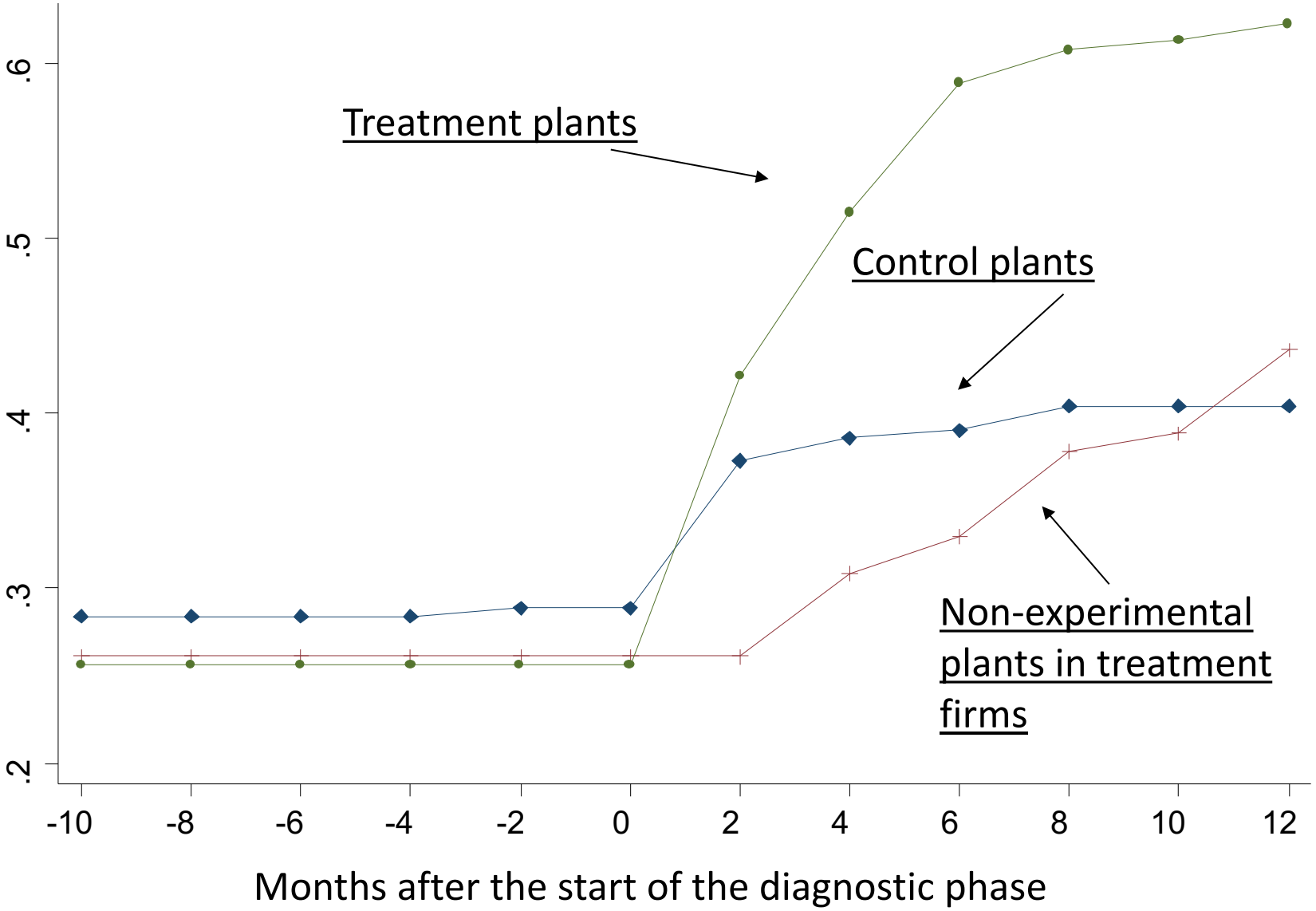
Management matters

- Bloom, N. et al. (2013) Does Management matter? *Quarterly Journal of Economics*
- Experiment on 20 plants in large multi-plant firms (average 300 employees and $7m sales) near Mumbai making cotton fabric
- Randomized **treatment** plants get 5 months of management consulting intervention, **controls** get 1 month
- Consulting is on 38 specific practices tied to factory operations, quality and inventory control
- Collect weekly data on all plants from 2008 to 2010

Examples of targeted practices

Area	Specific practice
Factory Operations	Preventive maintenance is carried out for the machines
	Preventive maintenance is carried out per manufacturer's recommendations
	The shop floor is marked clearly for where each machine should be
	The shop floor is clear of waste and obstacles
	Machine downtime is recorded
	Machine downtime reasons are monitored daily
	Machine downtime is analyzed at least fortnightly & action plans created and implemented to try to reduce this
	Daily meetings take place that discuss efficiency with the production team
	Written procedures for warping, drawing, weaving & beam gaiting are displayed
	Visual aids display daily efficiency loomwise and weaverwise
	These visual aids are updated on a daily basis
	Spares stored in a systematic basis (labeling and demarked locations)
	Spares purchases and consumption are recorded and monitored
	Scientific methods are used to define inventory norms for spares
Quality Control	Quality defects are recorded
	Quality defects are recorded defect wise
	Quality defects are monitored on a daily basis
	There is an analysis and action plan based on defects data
	There is a fabric gradation system
	The gradation system is well defined
	Daily meetings take place that discuss defects and gradation
	Standard operating procedures are displayed for quality supervisors & checkers

Adoption of 38 management practices over time



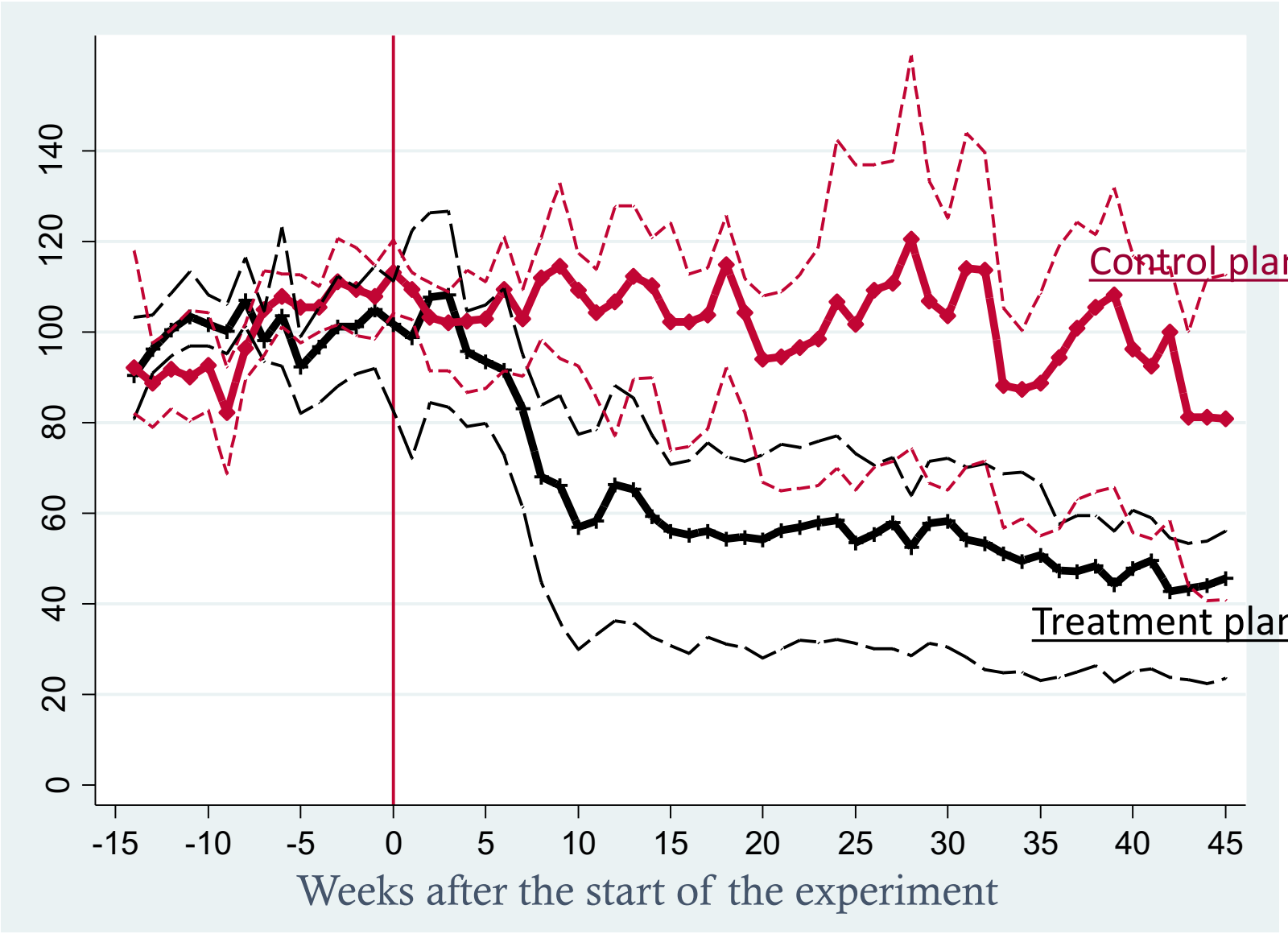
Measures of performance

1. **Quality:** Measured by Quality Defects Index (QDI) – a weighted average of quality defects (higher=worse quality)
2. **Inventory**: Measured in log tons
3. **Output**: Production picks (one pick=one run of the shuttle)
4. **Productivity:** Log(VA) – 0.42*log(K) – 0.58*log(L)

Estimate Intention to Treat (ITT) regression:

$$OUTCOME_{it} = alpha + beta*INTERVENTION_{it} + e_{it}$$

Quality improved after treatment



Note: solid lines are point estimates, dashed lines are 95% confidence intervals

Inventory fell after treatment

Note: solid lines are point estimates, dashed lines are 95% confidence intervals

TFP increased after treatment

Note: solid lines are point estimates, dashed lines are 95% confidence intervals

Intention to treat (ITT) Regression

Dep. Var.	Quality Defects$_{i,t}$	Inventory$_{i,t}$	Output$_{i,t}$	TFP$_{i,t}$
Intervention$_{i,t}$	-0.565*** (0.231)	-0.273** (0.116)	0.098*** (0.036)	0.169** (0.067)
Small sample robustness				
Ibragimov-Mueller (95% Conf. Intervals)	[-0.782, -0.441]	[-0.219, 0.001]	[0.218, 0.470]	[0.183, 0.511]
Permutation Test (p-values)	0.04	0.13	0.04	0.05
Time FEs	125	122	125	122
Observations	1396	1627	1966	1447

References

- Jeffrey Wooldridge, Introductory Econometrics: A Modern Approach, 4th ed. Thomson, chapter 13.2
- Angrist, J.and J. Pischke. *Mostly Harmless Econometrics*: An Empiricist's Companion. Princeton, chapters 1 and 2