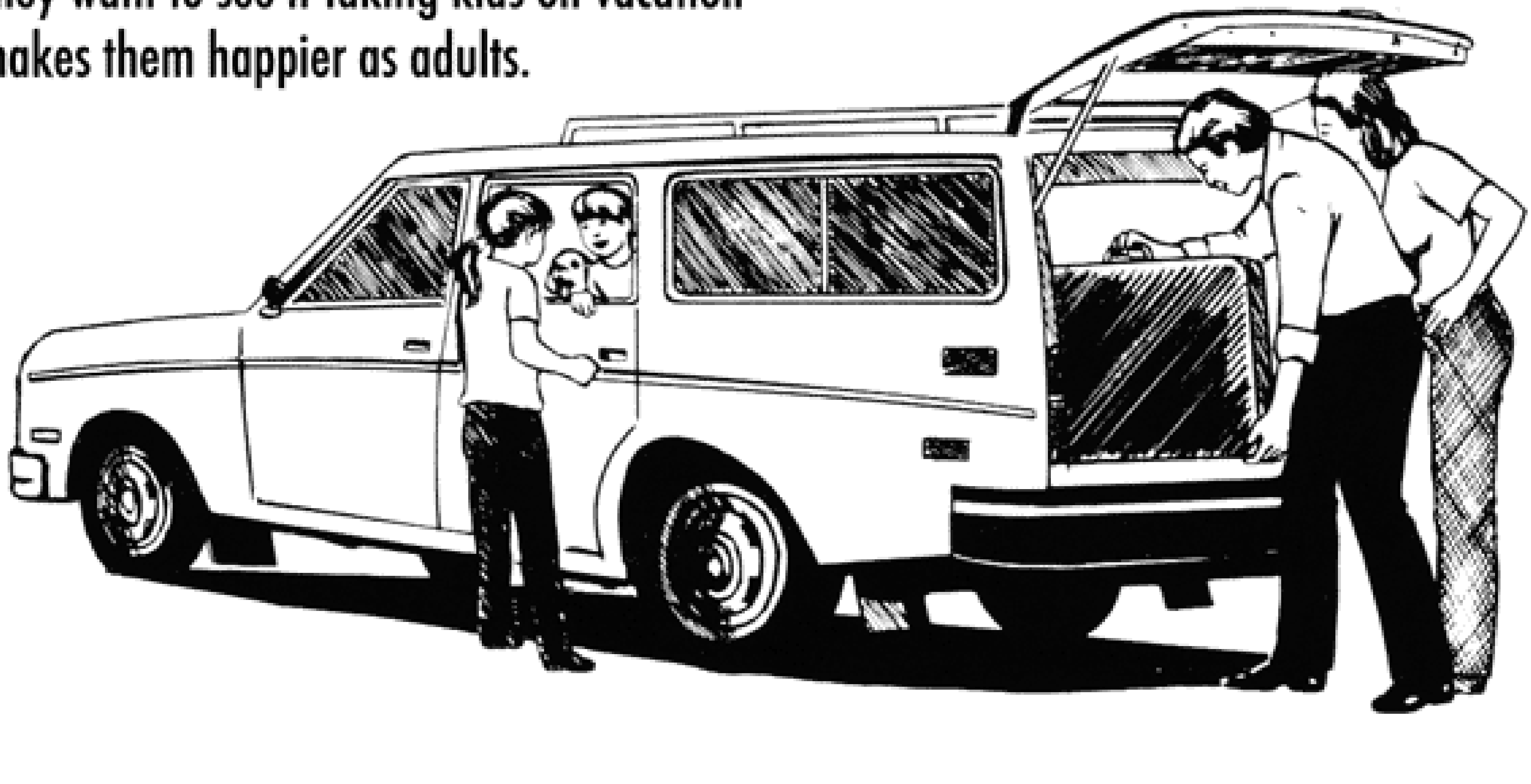
CAUSALITY

Again

I told you that you weren't coming with us, Nicole!
Mom and Dad are using you as the control group.
They want to see if taking kids on vacation
makes them happier as adults.



Conditional Average Treatment Effects

Average treatment effect (ATE) $E[\delta] = E[Y^1 - Y^0]$

Average treatment effect for the treated (ATT) $E[\delta|D = 1] = E[Y^1 - Y^0|D = 1]$

Average treatment effect for the controls (ATC) $E[\delta|D = 0] = E[Y^1 - Y^0|D = 0]$

ATT = average treatment effect for those that typically are (choose to be) treated based on counterfactual comparison.

In a well designed experiment ATE should (over many replications) be the same for those that just happen to be randomized to the treatment and those that just happen to be randomized to the controls.

Q. Is there any reason to expect this in observational data?

Sources of bias in the estimation of the ATE

Details of estimation from sample data skipped here (see Morgan & Winship , 2007: 44-46)

Turns out that:

$$E[Y^1|D=1] - E[Y^0|D=0] =$$

$$E[\delta]$$ The true average treatment effect

$$+\{E[Y^0|D=1] - E[Y^0|D=0]\}$$ A “baseline” bias

$$+ (1-\pi)\{E[\delta|D=1] - E[\delta|D=0]\}$$ A differential treatment effect bias

(Where π is the proportion receiving the treatment)

So, much depends on whether the second and third component can be either shown or assumed to = 0.

Numerical example – discussed in Morgan & Winship pp 47

Table 2.3: An Example of the Bias of the the Naïve Estimator of the ATE

Group	$E[Y^1 \mid .]$	$E[Y^0 \mid .]$
Treatment group (D=1)	**10**	**6**
Control group (D=0)	**8**	**5**

Effect of college degree on a labour market outcome. Assume $\pi = 0.3$

Average PO under treatment for treated = 10 and average PO under control for controls = 5. This is what is observed. Q. Is the ATE = 5?

Consider what would have happened in the control state. Those in D=1 would have done better, 6 versus 5. Baseline difference.
Consider what would have happened in the treatment state. Those in D=0 would have done worse, 8 versus 10.

ATT = 4; ATC = 3 therefore
ATE = 0.3(10-6) + 0.7(8-5) = 3.3 NOT 5!

Example: Church Schools

We observe (hypothetically) that children attending religious schools do better in exams

Why?

1) Religious schools do a better job of teaching kids (there is a causal impact).

2) Kids that entered religious schools were different (smarter, came from more advantaged homes) right from the start.

3) Kids that actually entered religious schools flourish more in religious schools than would the kids whose parents (actually) chose a secular school for them.

What we want to know about is 1). But 2) & 3) get in the way. 2) is a problem of heterogeneity; 3) could be a problem of self-selection on the basis of the anticipated outcome (parents select schools they think will suit their kids).

Example 2

Random assignment to treatment and control solves (**on average**) 2) and 3)

But for problems like this randomization is, except under special circumstances, likely to be difficult (or impossible). Many of the problems sociologists are interested in share these characteristics:

Treatment and control groups are heterogeneous. Normal move is to try to deal with heterogeneity by :

1) Matching on observables
2) Conditioning on observables (ie introducing relevant control variables)
3) Or both.

No guarantee this will work!

Units self-select themselves into or are selected into treatment and control on the basis of anticipated outcomes. If we notice that kids who read for an hour a day for pleasure do better in school, would we expect the same result if we forced reading on kids that wouldn't normally choose it as a leisure pursuit?

These thoughts also have implications for what it makes sense to target

Causal effect of treatment on the treated?

Effect of Catholic school on those kids that would choose a Catholic school

Effect of a training programme on those that would choose to take it

Average causal effect?

Effect of Catholic school on Catholics and others?

Causal effect of treatment on those that would choose the control?

Effect of Catholic school on Northern Ireland Protestants

Effect of marriage on those who prefer cohabitation

Internal Validity

- Internally valid designs are resistant to rival explanations in terms of factor(s) other than the applied treatment is (are) the **cause(s)** of any observed effect
- Good experimental design seeks to maximize internal validity against a number of particular threats, in other words weaknesses of design that call into doubt the unique attribution of changes in values of the response to the experimental manipulation.

Threats to Validity

- Internal Validity
  - Maturation
  - Selection
  - History
  - Testing
  - Instrumentation
  - Mortality
  - Regression to the mean

External Validity

- Generalisability beyond the experimental setting
  - The Hawthorne Effect
- The extent to which the effects discovered in the experiment can be extended to the target population.

Pretest-Postest Nonequivalent Control Group Design

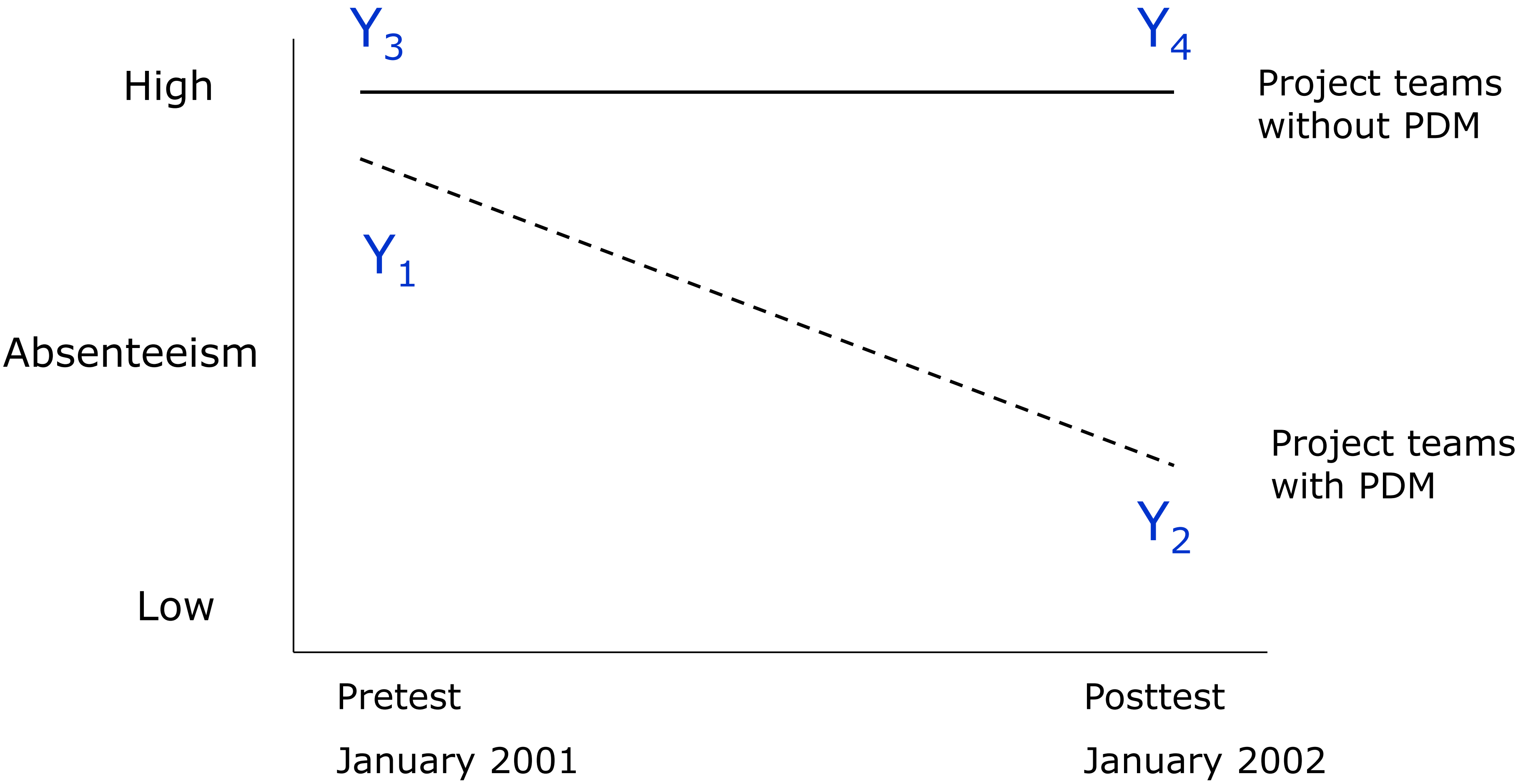
GROUP 1	Y_1	X	Y_2
- - - - - - - - - - - - - - - - - - - -			
GROUP 2	Y_3		Y_4

-------- = non random allocation to groups

Participative decision making

- A large management consultancy organizes work on the basis of “project teams” with a team leader
- Senior partners are worried about morale level (as reflected in absenteeism rate)
- Team leaders are allowed to adopt (if they wish) more participatory ways of making decisions
- Some do and others don’t
- Do participatory teams do better than others?

PDM Outcome 1



PDM Outcome 1 continued

- PDM adopters already have lower absenteeism rates

- Difference becomes bigger after adoption of PDM

  - But PDM adopters might already be on a downward trend (perhaps team leaders are more easy going)

  - aka **selection by maturation** threat
    - Can’t be ruled out without more pretest observations

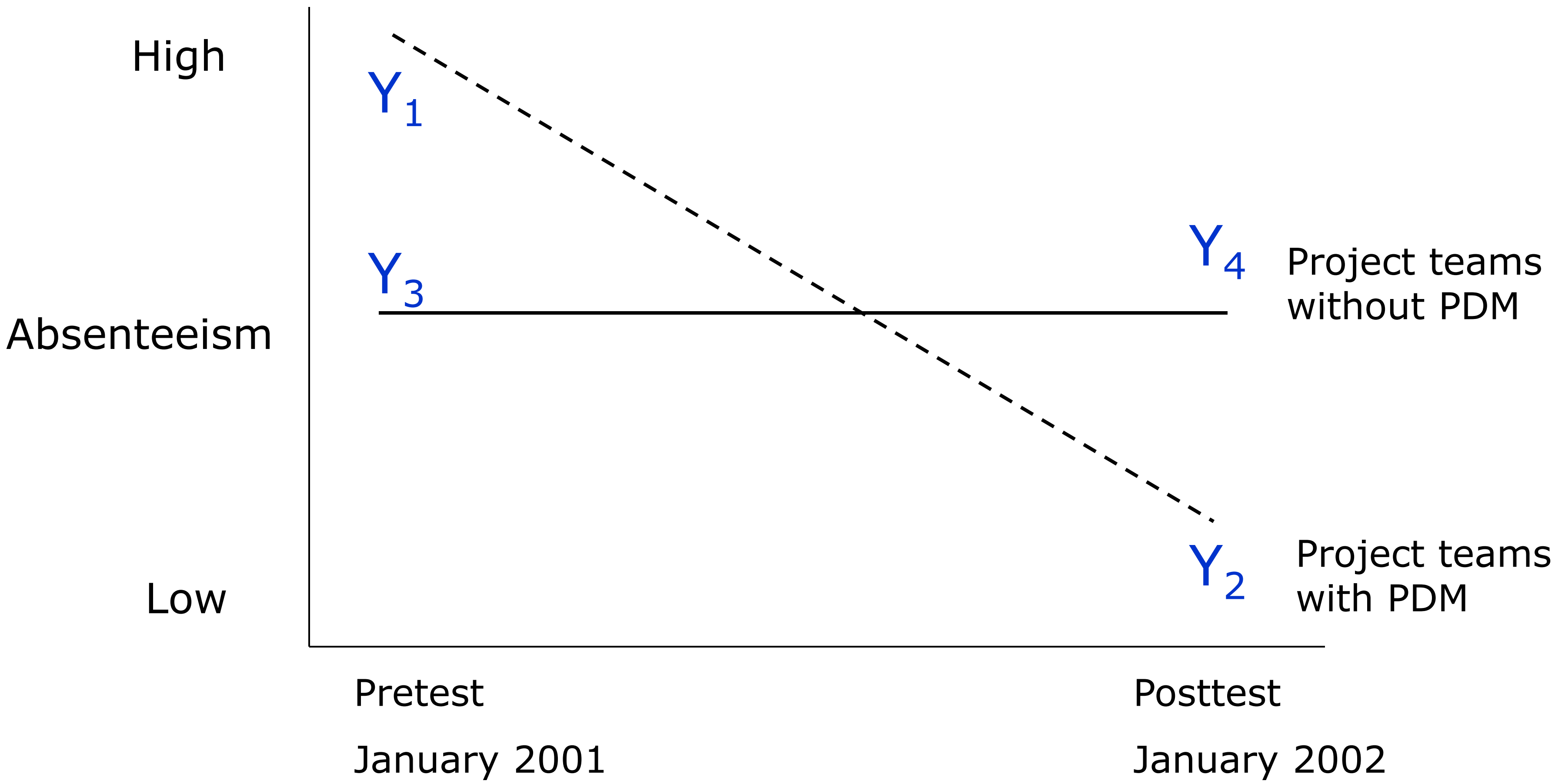
- Conclusion depends on the plausibility of the selection by maturation threat **in this particular case**

PDM Outcome 2

- Imagine a different set up

- PDM is imposed by the senior partners on the project teams with the highest absenteeism rates

PDM Outcome 2 continued



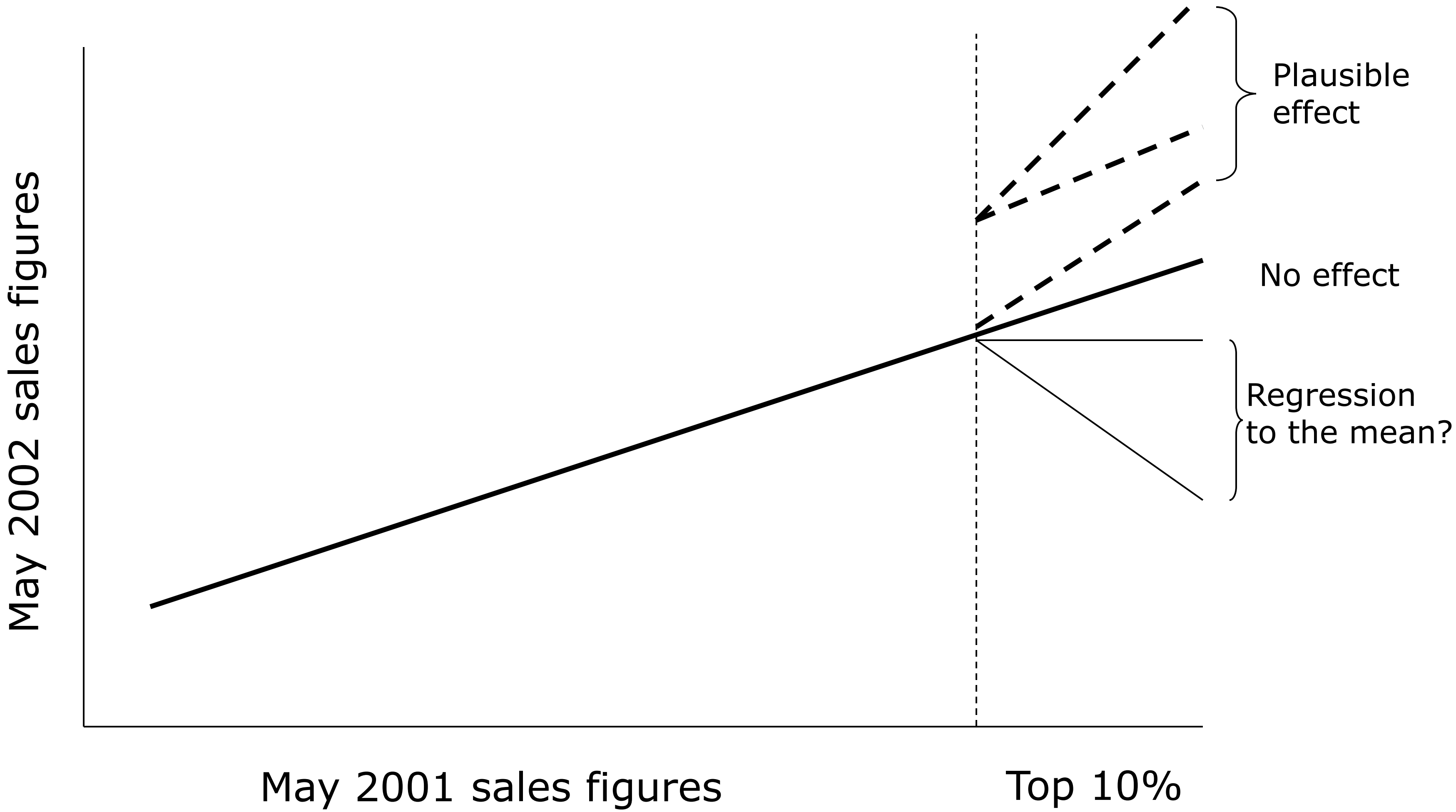
PDM Outcome 2 continued

- Genuine “treatment” effect looks more plausible
- Why should differential “maturation” lead to a crossover?
- Why should “regression to the mean” lead to crossover?
- Effect of PDM still not “proven” but case looks stronger

Regression-Discontinuity Design

- A variant of the PPNCGD
- If selection into treatment is based in a known way on pretest score then R-D design possible
- Say top 10% of sales force are given a bonus over and above their commision
- Does it affect their performance?

Regression-Discontinuity Design continued



Interrupted Time-Series Design

If you can't control who is exposed to the treatment…

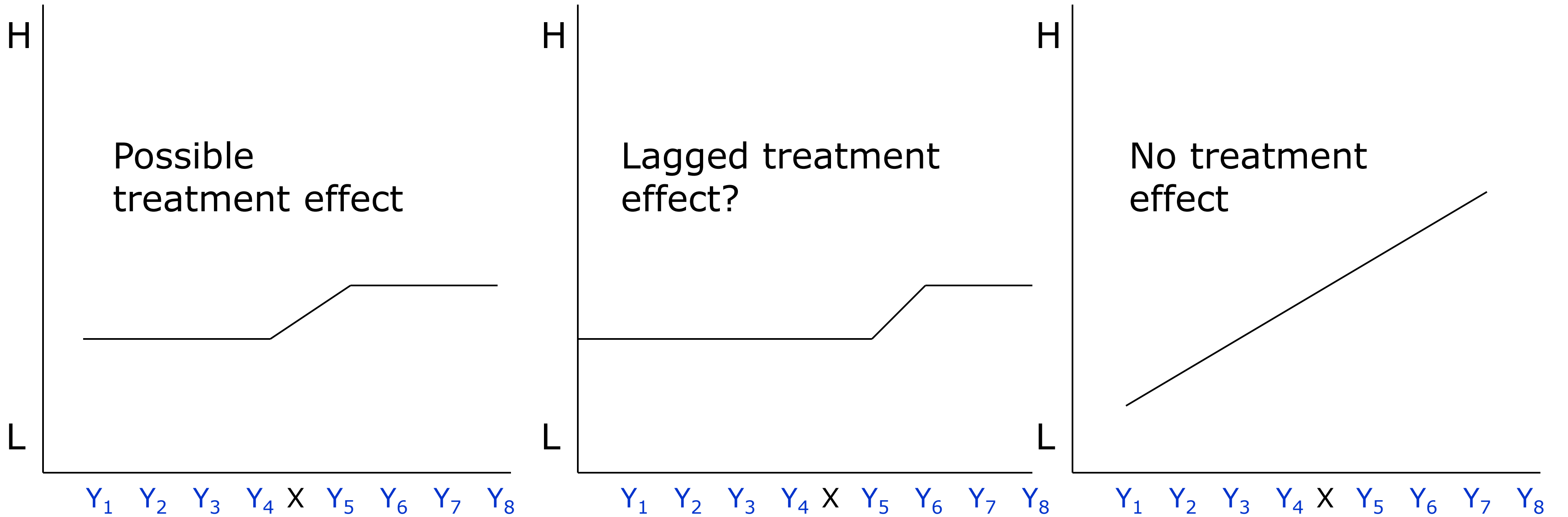
Try to control **when** observations are made

GROUP 1 Y_1 Y_2 Y_3 Y_4 X Y_5 Y_6 Y_7 Y_8

Even spacing is nice

Can rule out maturation by looking at trends in the pretests

Possible outcomes of the interrupted time series



If the process has no "memory"…

Apply treatment; Remove treatment; Apply again

GROUP 1 Y_1 Y_2 Y_3 X Y_4 X Y_5 X Y_6 Y_7 Y_8 X Y_9 X Y_{10} X Y_{11}

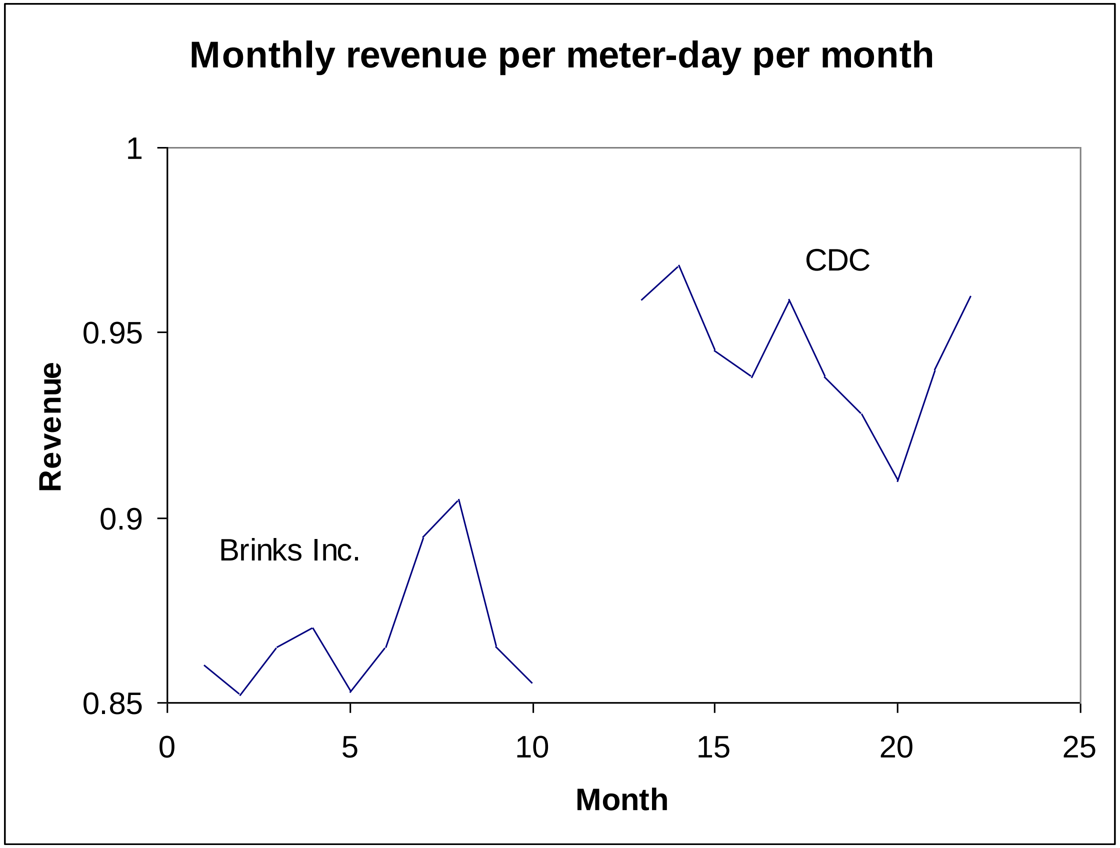
Is the effect reversible?

Harder to take away a benefit than to introduce it!

Let the data speak for itself?

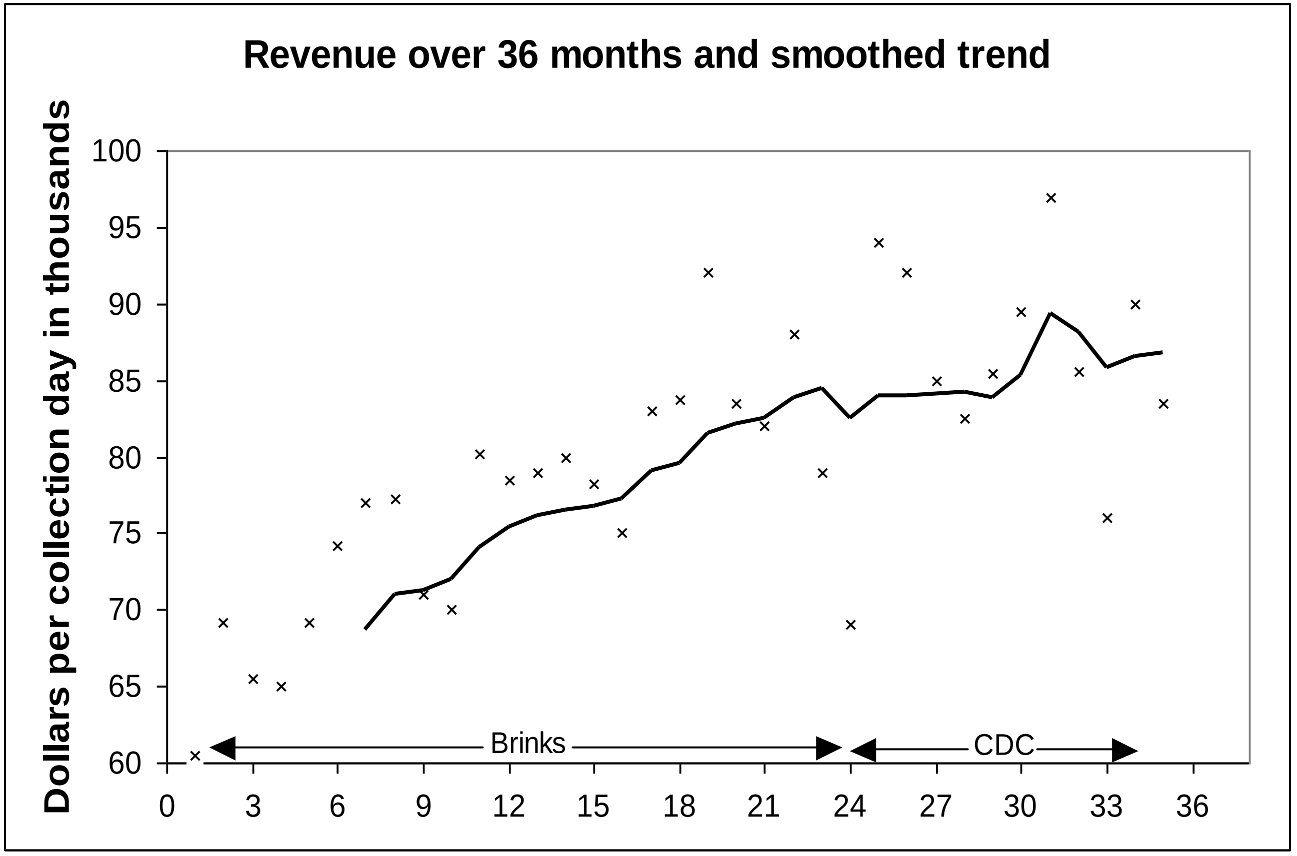
- Contract to empty parking meters in NYC given to Brinks Inc
  - Brinks employees found guilty in criminal court of “skimming”
  - NYC launched civil action against Brinks for negligence and breach of contract
- Contract given to CDC
  - In first 10 months CDC collected $1millon more than in any 10 month period when Brinks held the contract

Evidence of theft?



From Fairley, W. B. and Glenn, J.E. "A question of theft" in DeGroot, M. H., Fienberg, S. E. and Kadane, J. B., eds., *Statistics and the Law*. New York: Wiley, 1986.

Or just growth?



From Levin, B. “Comment” in DeGroot, M. H., Fienberg, S. E. and Kadane, J. B., eds., *Statistics and the Law*. New York: Wiley, 1986.