

# BE CAREFUL WHAT YOU ASK FOR:

## Stepped wedge trials with time-varying treatment effects

Jim Hughes

University of Washington and SCHARP

Joint work with [Avi Kenny](#), Patrick Heagerty, Fan Xia, Emily Voldal

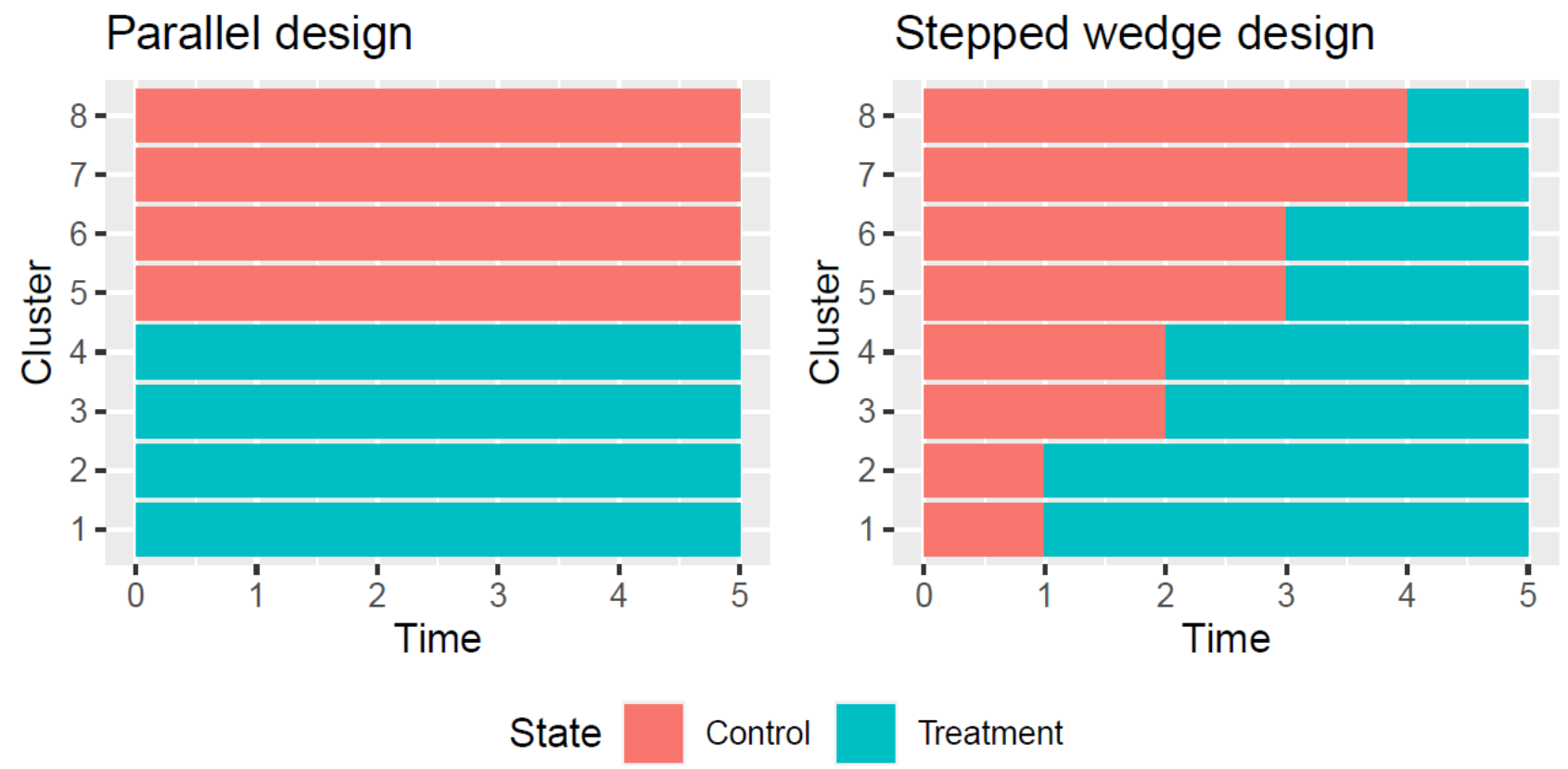


HPTN  
HIV Prevention  
Trials Network

**ANNUAL MEETING**  
2022

# Background on stepped wedge design and analysis

- Clusters randomized to when intervention is received
- Typically, measure outcome on each cluster, at each time step
- SW often used to measure effectiveness during roll-out



**FIGURE 1** Schematic representations of a parallel CRT versus a stepped wedge CRT design with 8 clusters.

Immediate treatment (IT)  
model:  
(Hussey & Hughes 2007)

$$Y_{ijk} = \alpha + \beta_j + \delta X_{ij} + \nu_i + e_{ijk}$$

Outcome  $Y_{ijk}$

Overall mean  $\alpha$

Time trend  $\beta_j$

Treatment effect  $\delta$

Treatment indicator  $X_{ij}$

Cluster random effect  $\nu_i$

Error  $e_{ijk}$

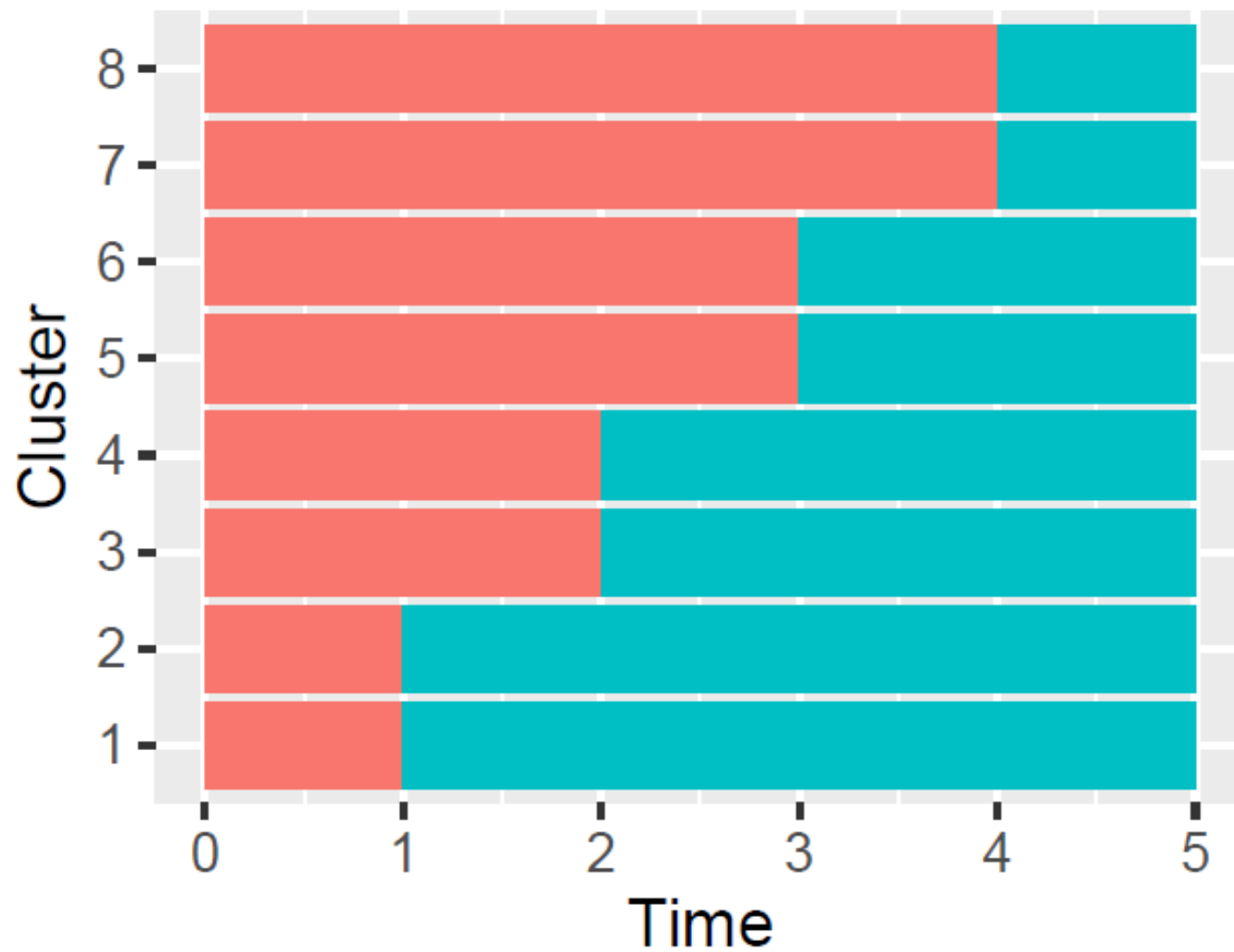
IT model: 
$$Y_{ijk} = \dots + \delta X_{ij} + \dots \quad (1)$$

ETI model: 
$$Y_{ijk} = \dots + \delta(s_{ij})X_{ij} + \dots \quad (2)$$

(Kenny et al, 2022)

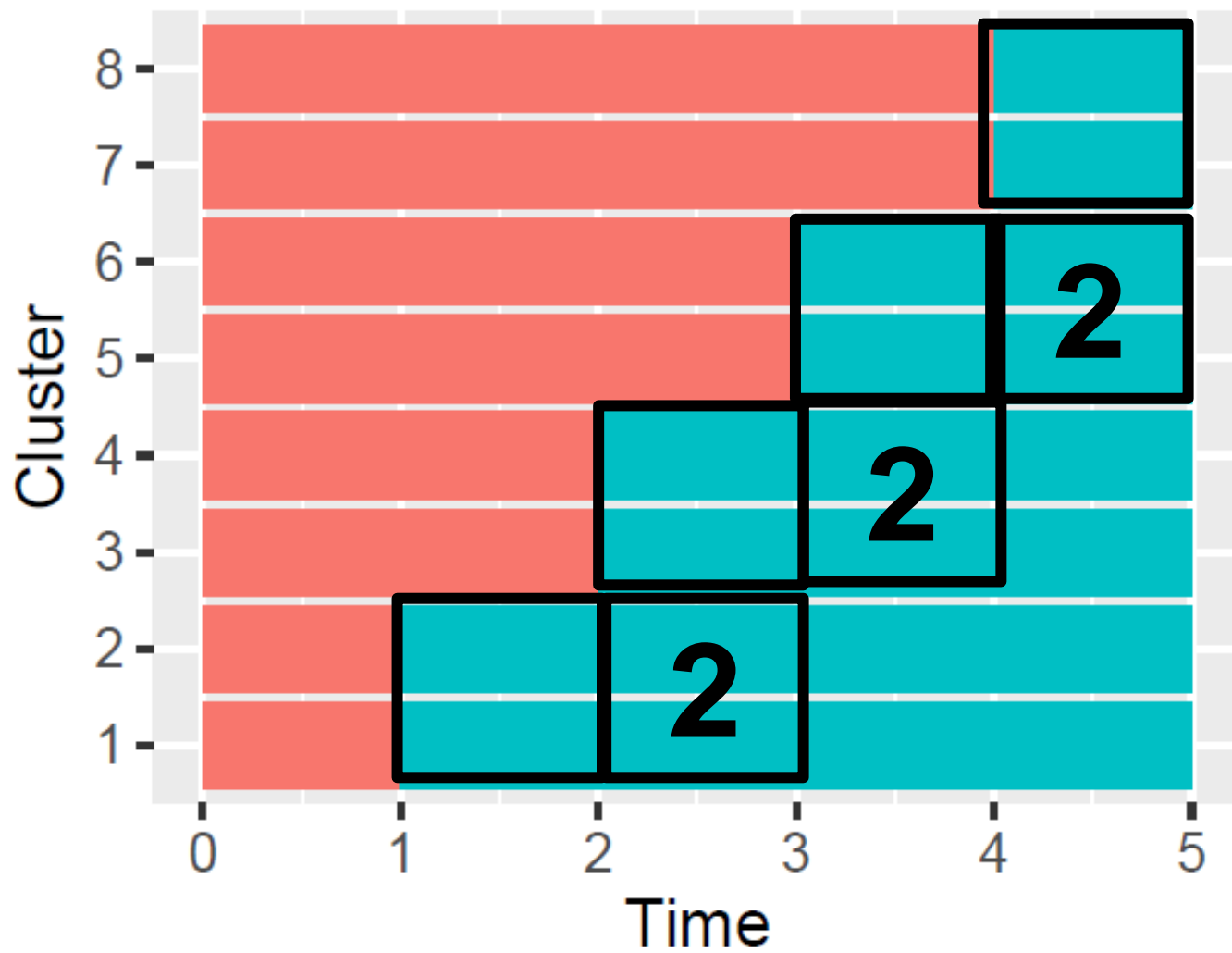
- In model (2), the treatment effect  $\delta$  is a function of *exposure time*  $s_{ij}$  (time since intervention start).

## Stepped wedge design

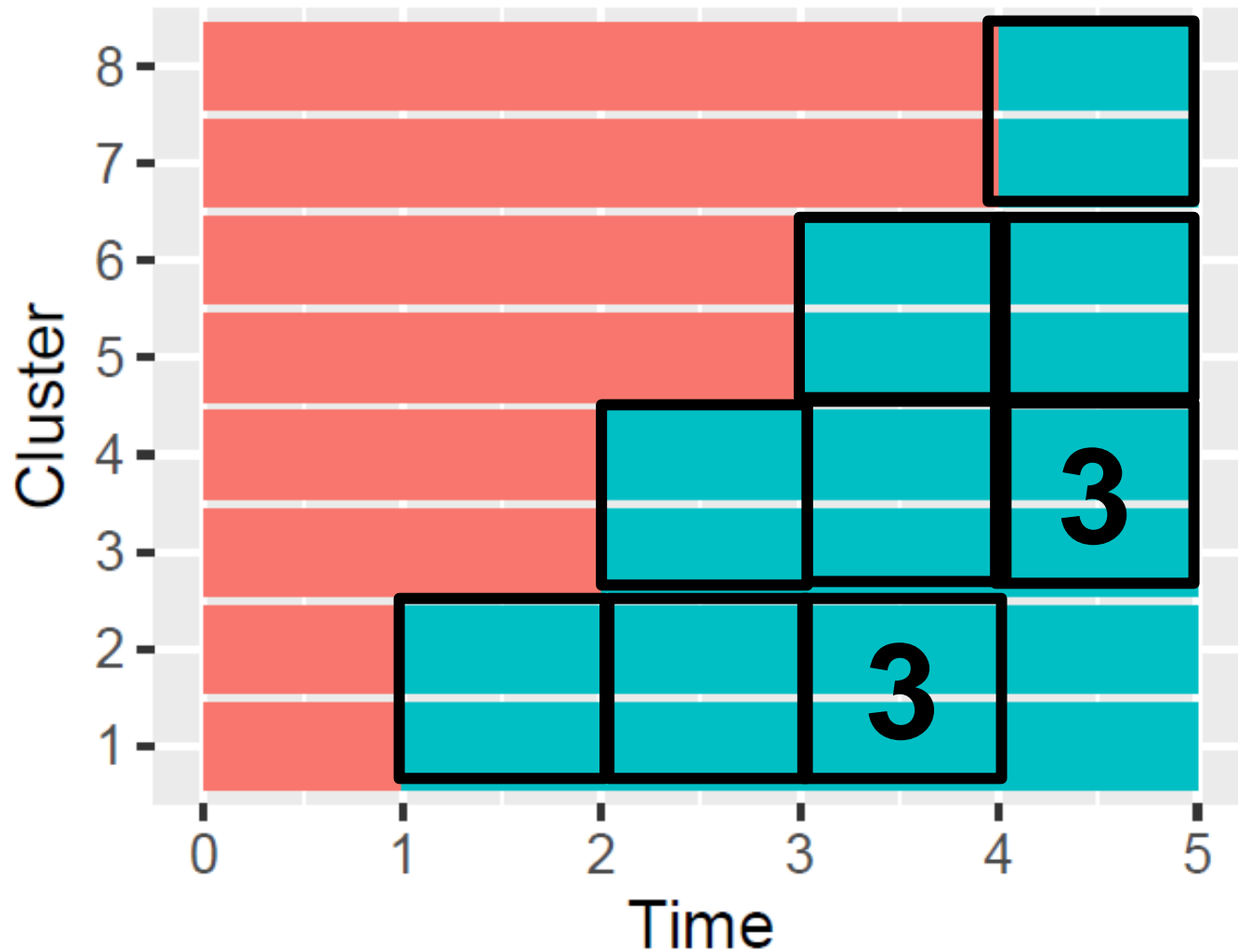




## Stepped wedge design

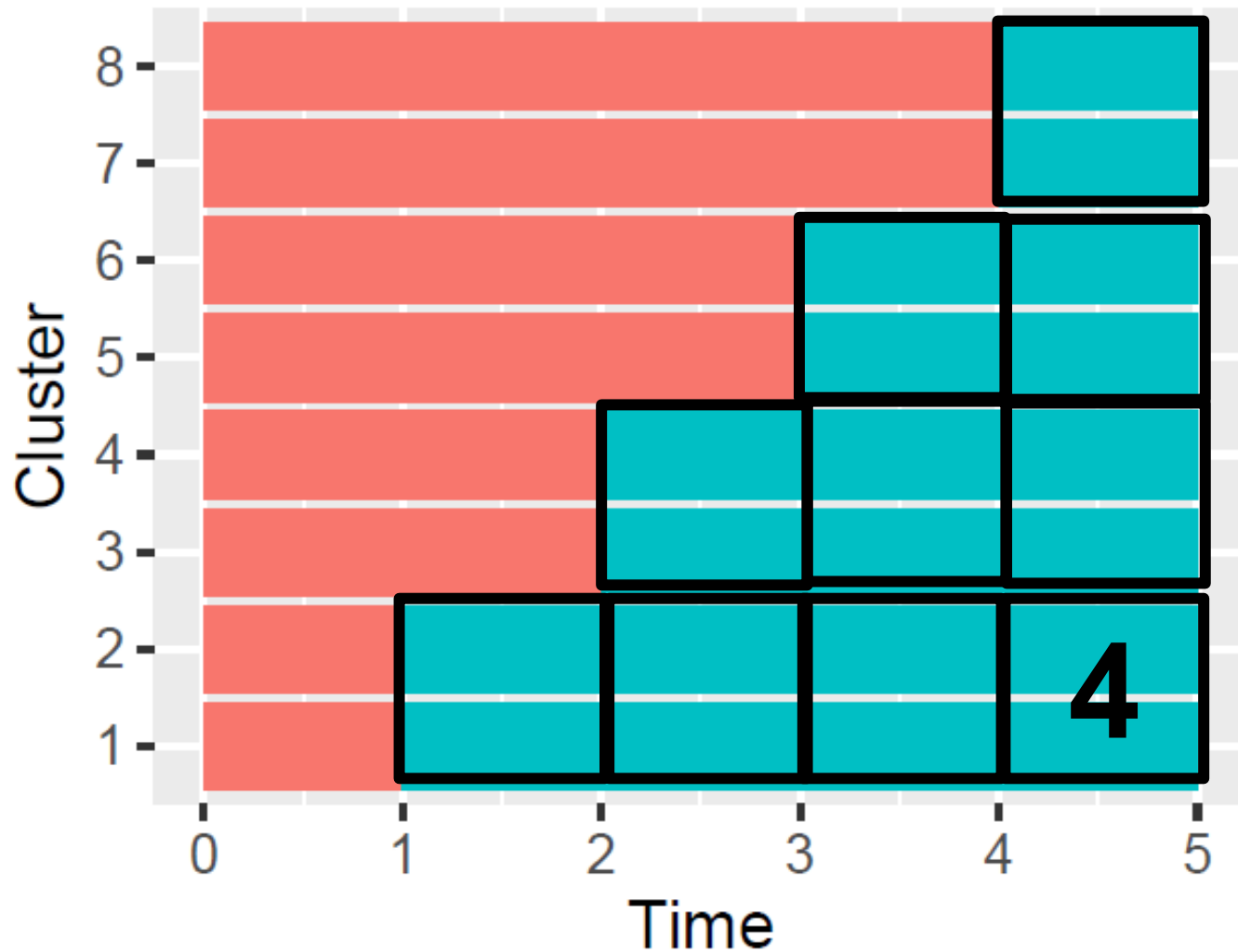


# Stepped wedge design





# Stepped wedge design



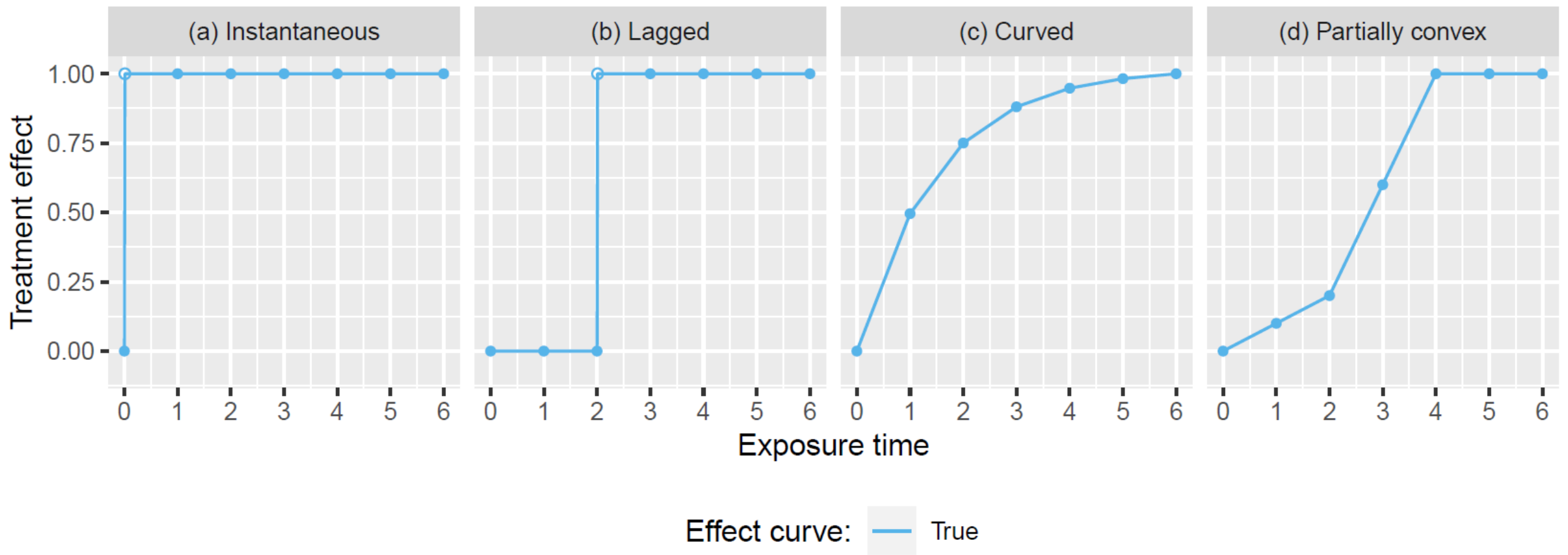
# The problem of a time-varying treatment effect

IT model: 
$$Y_{ijk} = \dots + \delta X_{ij} + \dots \quad (1)$$

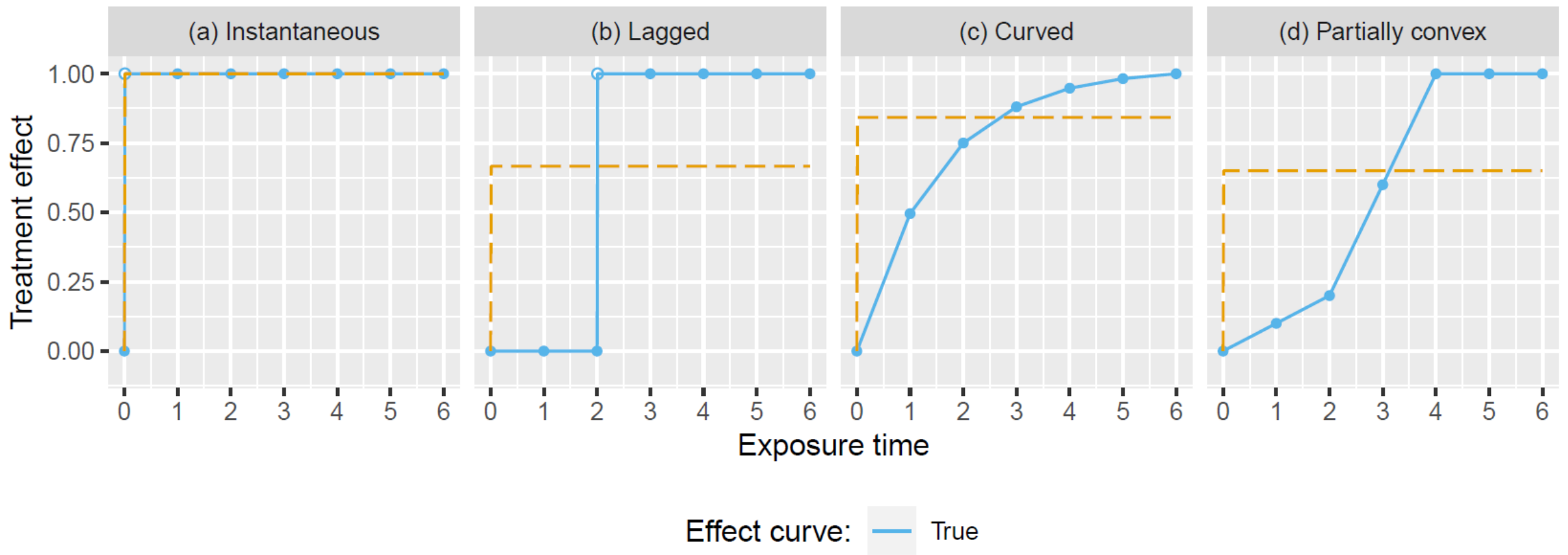
ETI model: 
$$Y_{ijk} = \dots + \delta(s_{ij})X_{ij} + \dots \quad (2)$$
  
(Kenny et al, 2022)

- In model (2), the treatment effect  $\delta$  is a function of *exposure time*  $s_{ij}$  (time since intervention start).
- What happens if data are generated according to (2) but analyzed with (1)?

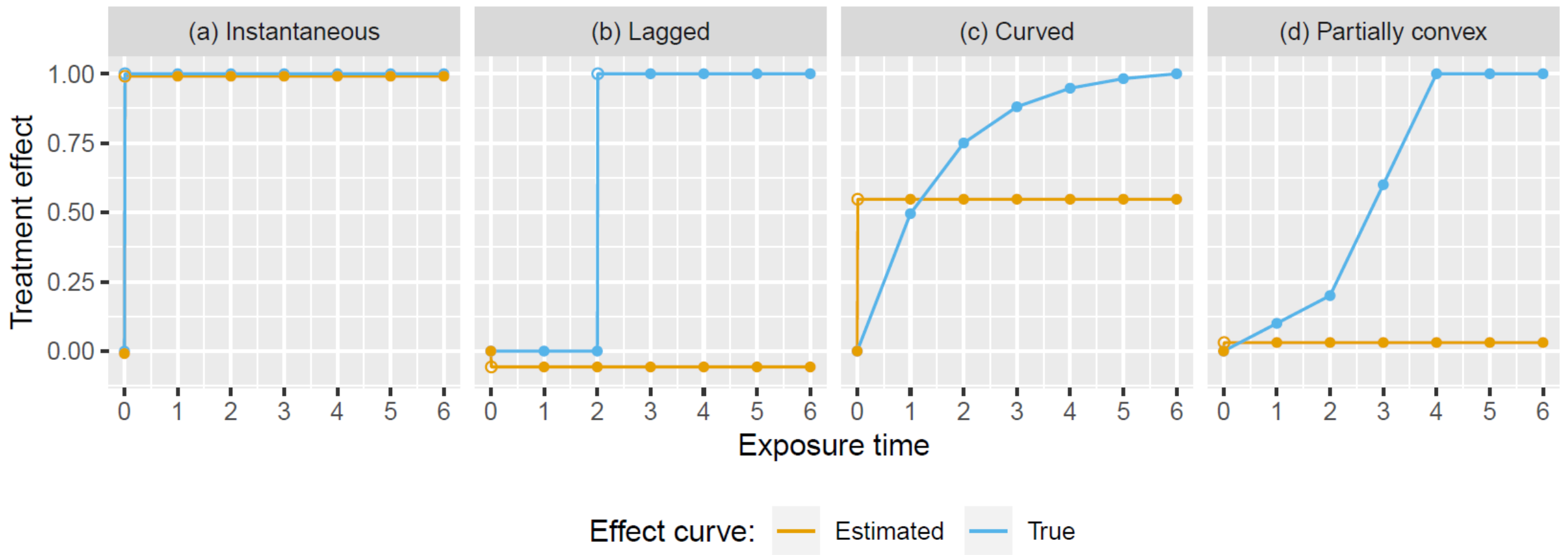
# The problem of a time-varying treatment effect



# The problem of a time-varying treatment effect



# The problem of a time-varying treatment effect



So what do we do???

1. Think hard about what you mean by “Treatment effect”
  - Treatment effect at a particular time?
  - Average treatment effect over an interval?
  - Average treatment effect after a lag?
2. Avoid modelling assumptions

## ETI model

$$Y_{ijk} = \dots + \delta(s_{ij})X_{ij} + \dots$$

$$\text{ATE estimator: } \hat{\Psi}_{[s_1, s_2]} = \frac{1}{s_2 - s_1 + 1} \sum_{r=s_1}^{s_2} \hat{\delta}(r)$$

$$\text{PTE estimator: } \hat{\Psi}_s = \hat{\delta}(s)$$



## Key findings:

- Estimate from IT model is biased for ATE/PTE in all cases, except when IT model is true
- ATE/PTE estimator is unbiased in all cases
- ATE/PTE is less “efficient” (bigger standard error) than IT model estimator when IT model is true

## Conclusions:

- In stepped wedge studies, be careful fitting models that assume immediate, constant treatment effect
- Think carefully about how you want to define the “treatment effect”
- In most cases, we recommend constructing a robust estimate of the treatment effect based on the  $\hat{\delta}(s)$

# Acknowledgments

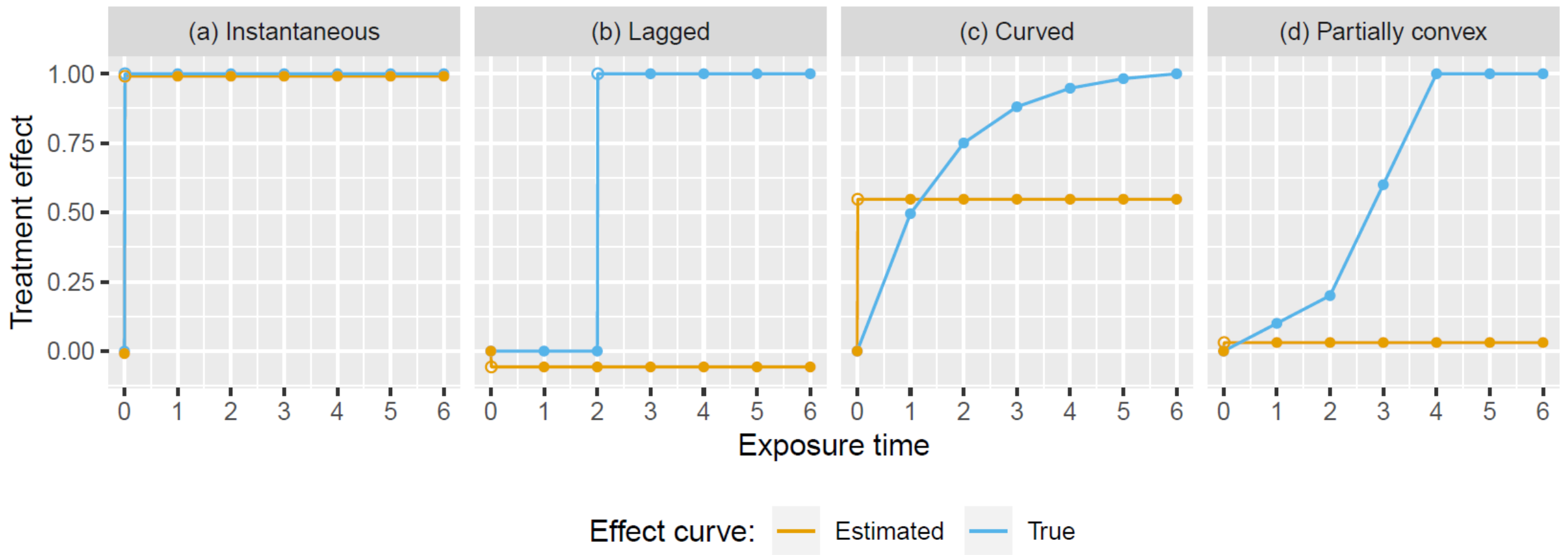
- Overall support for the HIV Prevention Trials Network (HPTN) is provided by the National Institute of Allergy and Infectious Diseases (NIAID), Office of the Director (OD), National Institutes of Health (NIH), National Institute on Drug Abuse (NIDA), the National Institute of Mental Health (NIMH), and the Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD) under Award Numbers UM1AI068619-15 (HPTN Leadership and Operations Center), UM1AI068617-15 (HPTN Statistical and Data Management Center), and UM1AI068613-15 (HPTN Laboratory Center).
- The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

# The problem of a time-varying treatment effect

Why does this happen?

- The usual IT estimator can be written as  $\hat{\delta} = \sum_s w(s) \hat{\delta}(s)$
- $w(s)$  are weights that sum to 1.0
- BUT  $w(s)$  can be  $>1$  or  $<0$  !
  - This can occur when you combine multiple correlated estimators (of the same parameter) that have variable precision
- In the examples given,  $w(1) > 1$  and  $w(6) < 0$

# The problem of a time-varying treatment effect



# Treatment effect is not constant - Power

- Specify the SW design via design matrix  $\mathbf{X}$  (including time-varying treatment effect)
- Use GLMM framework to specify variance  $\Sigma$
- Specify the estimator  $\hat{\Psi} \equiv \frac{1}{s_2 - s_1 + 1} \sum_{r=s_1}^{s_2} \hat{\delta}_r$

$$Var(\hat{\Psi}) = H^T (\mathbf{X}^T \Sigma^{-1} \mathbf{X})^{-1} H$$

For testing  $H_0: \Psi = 0$ ,

$$Power(\Psi) = \Phi \left( \sqrt{\frac{\Psi^2}{Var(\hat{\Psi})}} - Z_{1-\frac{\alpha}{2}} \right)$$