Approaches to Control for Observable Selection Bias in Studies including Staggered Treatment Timing

Melissa Garrido,[1-2] Jessica Lum,[1] Daniel Asfaw,[1-2]
Steven Pizer[1-2]

Boston VA Healthcare System and Boston University School of Public Health, Boston, MA, USA

PEPReC
Partnered Evidence-based Policy Resource Center
A VA QUERI Center

Background:

- Most current methods to account for treatment effect timing heterogeneity do not allow a treatment's effect to be isolated from effects of confounders associated with both treatment timing and outcome
- Two-way fixed effects (TWFE) models produce biased or difficult to interpret estimates when effects vary with timing or duration
- A potential solution involves using difference-in-differences with inverse probability of treatment weights (IPTW) to adjust for confounding across treatment cohorts, but IPTWs lead to biased estimates in cross-sectional evaluations comparing multiple treatments
- We developed an alternative method, vector-based kernel weighting (VBKW), that produces estimates that are less biased and more efficient than IPTW in cross-sectional evaluations
- A natural extension of VBKW is to longitudinal studies in which treatment groups are defined by time of treatment receipt

Objective:

- To compare the bias and efficiency of VBKW and TWFE-based estimates in longitudinal studies

Methods:

- Simulations with 300 observations over 20 time periods (500 replications).
- We specified a static true treatment effect (0.2) and induced heterogeneity solely through an observable factor associated with treatment timing and outcomes
- We conducted simulations with mild and moderate heterogeneity (coefficient on confounder = 0.03, 0.1) and in scenarios with an even split of observations across treatment cohorts (never, early[time2], late[time10]) or relatively few never treated observations (50 never, 125 early, 125 late)

Results:

VBKW produces estimates with lower absolute mean relative bias (AMRB) than TWFE

Heterogeneity	Sample distribution across treatment times	VBKW AMRB (%)	TWFE, main effects AMRB (%)	TWFE, interaction between time & confounder AMRB (%)
Mild	Even	**2.63**	7.30	3.71
	Uneven	**0.21**	7.42	3.73
Moderate	Even	**8.76**	24.32	12.36
	Uneven	**0.70**	24.74	12.46

VBKW produced estimates with slightly larger mean interquartile ranges (IQR) than TWFE

Heterogeneity	Sample distribution across treatment times	VBKW IQR	TWFE, main effects IQR	TWFE, interaction between time & confounder IQR
Mild	Even	.0026	.0019	.0015
	Uneven	.0048	.0028	.0024
Moderate	Even	.0086	.0064	.0049
	Uneven	.0162	.0094	.0081

Conclusions:

- **In analyses involving staggered treatment timing, VBKW produces less biased estimates than TWFE. The improvements in bias outweigh the efficiency losses that accompany VBKW.**
- **When treatment effects are dynamic, we expect these differences in bias will be more pronounced. Future comparisons will include IPTWs, optimal weighting methods, and a broader range of analytic scenarios, including dynamic treatment effects.**
- **Identifying best practices to account for heterogenous treatment timing is critical to improving the rigor of evidence used to support complex interventions that cannot be randomized.**

garrido@bu.edu

Supported by NIH R56 AG068178 (PI: Garrido)