Peer Institution Networks, Test-Optional Admission Policies, and STEM Major Completions

Ethan N. Lewis (Boston University)

AEA 2026 Annual Meeting

Overview

- Construct a network of private colleges based on peer institutions and use an algorithm to detect communities within it.
  - Use communities to investigate heterogeneous impacts of test-optional admissions policies.
  - Estimate population- and community-level treatment effects by comparing adopters and non-adopters within the same community.
  - Sample consists of 66 pre-COVID policy adopters; results on downstream outcomes still policy relevant.
- **Main Findings**
  - $\approx$ 1 in 5 students in "treated" cohorts do not submit an SAT/ACT score, heterogeneity across communities.
  - First/third quartiles of reported score distribution increase.
  - Student-body diversity improves.
  - **Share of graduates completing a STEM major declines, consistent with mismatch.**
    - ⋆ **Effect driven by sharp declines in Colleges and Universities on the East Coast & Elite Liberal Arts Colleges.**

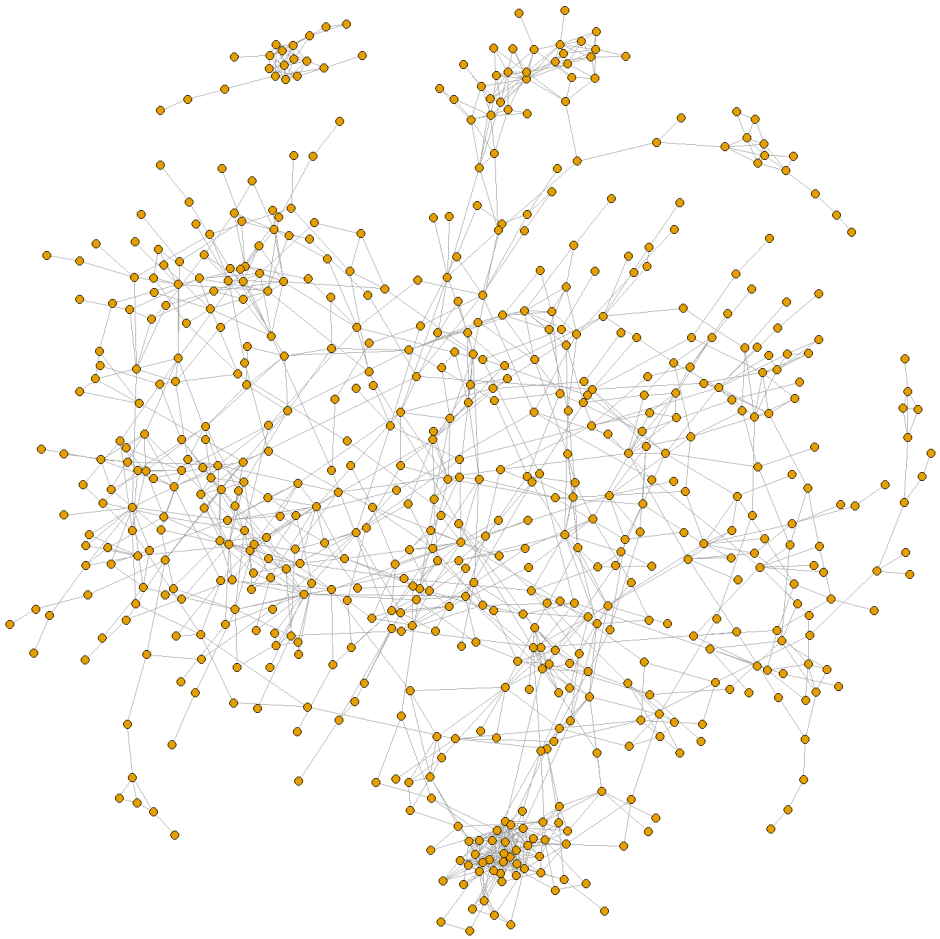
Peer Institution Network

- Networks consist of nodes (agents) and edges (connections).
- Private colleges and universities = nodes. Connected if each listed the other as a peer institution when reporting data to the NCES.
  - If p_i is the set of peers in college i's custom list. Then,

$$i \text{ is connected to } j \iff j \in p_i \ \& \ i \in p_j .$$

- **Community Detection:** Algorithmically identify clusters of densely packed nodes in the network.
  - Apply Leiden algorithm (Traag et al. (2019)) to optimally partition nodes.
  - Identifies 16 distinct communities, including communities of "Elite" National Universities and "Elite" Liberal Arts Colleges.

Figure: Peer Institution Network



# of Nodes = 617, # of Edges = 1362

Figure: Communities in the Peer Institution Network

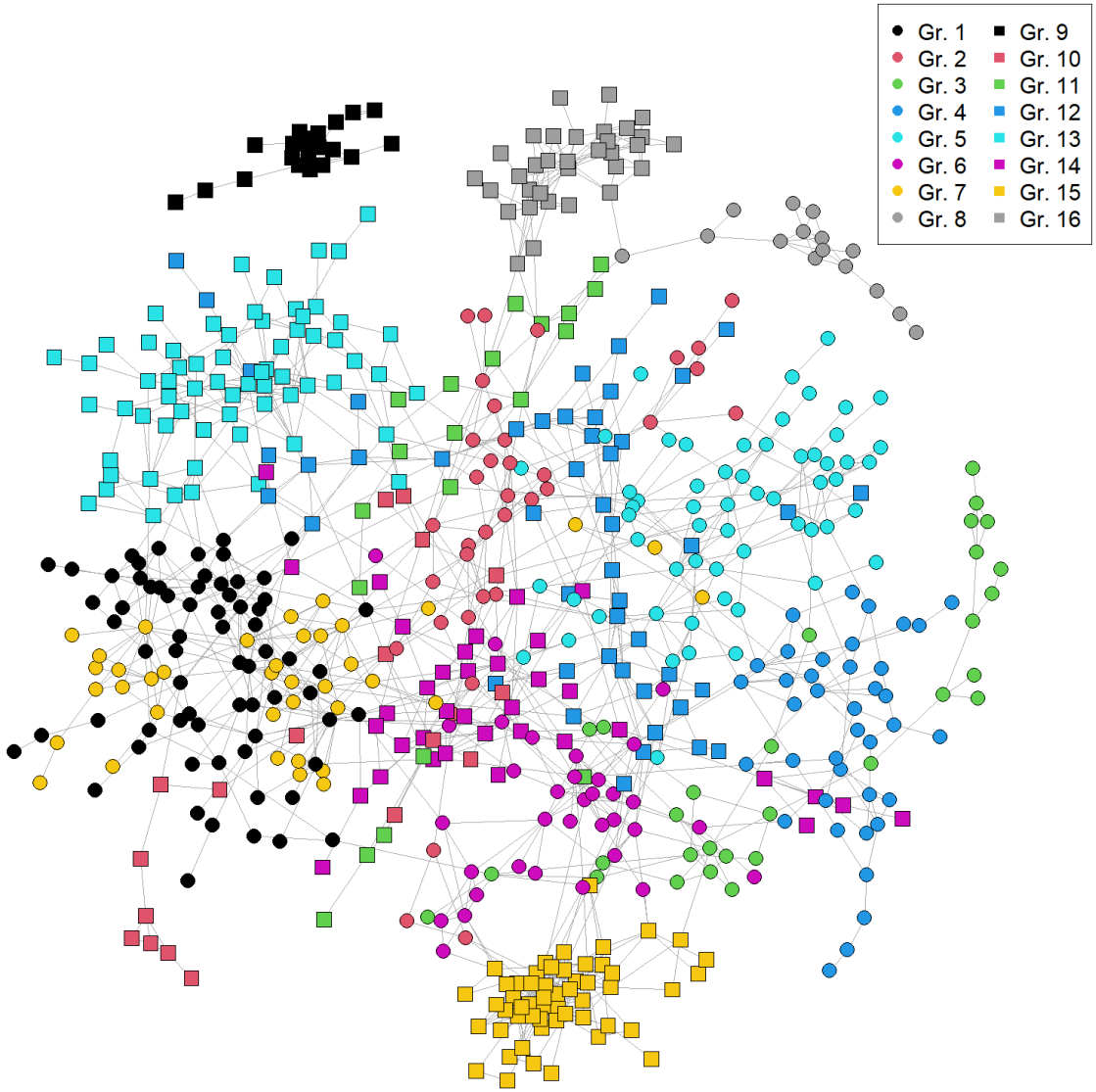
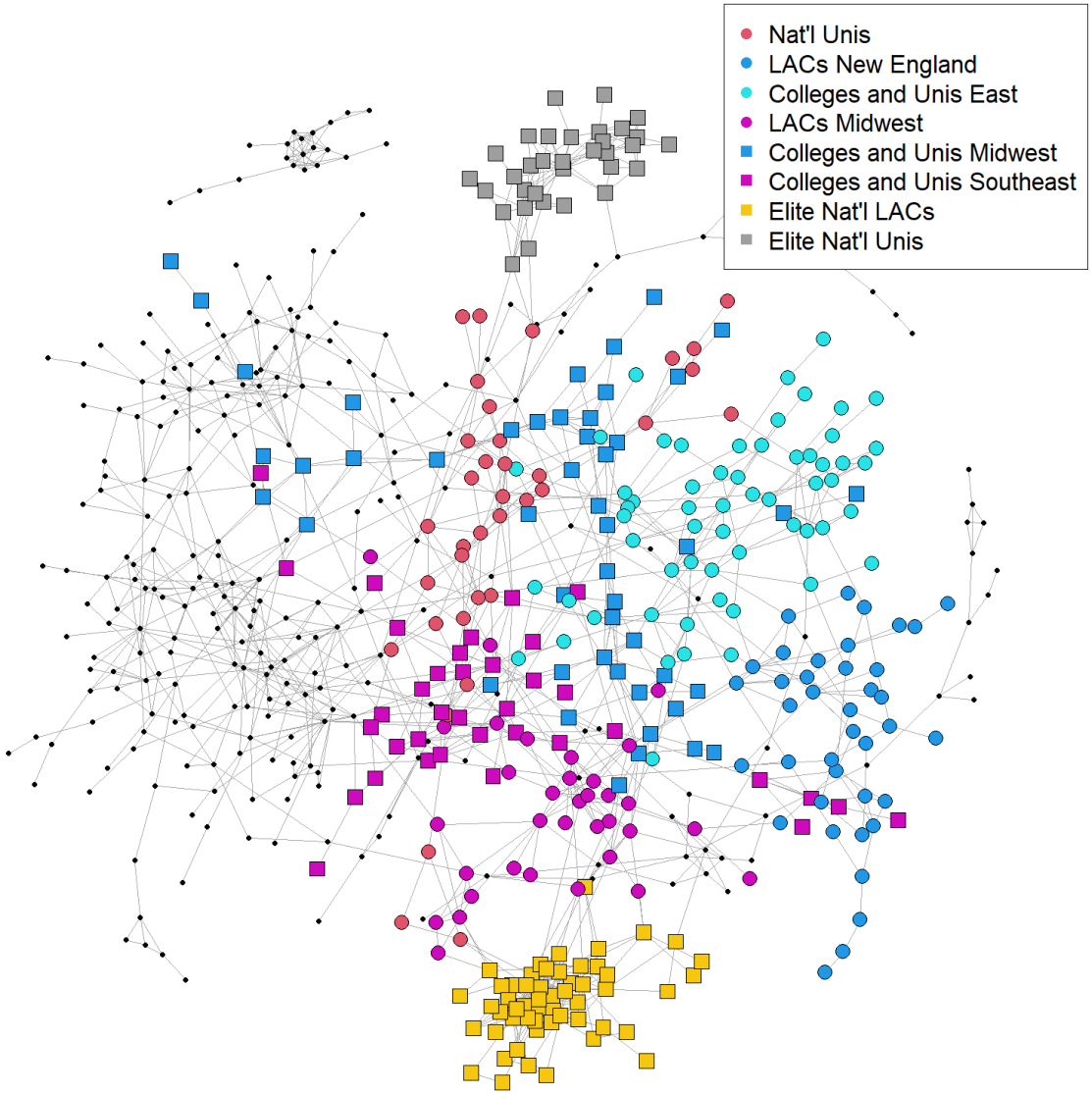


Figure: Policy Relevant Communities



Data and Methodology

Final Dataset: Panel of private colleges (2001–2019) with 66 treated units (adopted 2005–2016) and 139 never treated. Summary Statistics

Outcome = (Log of) STEM BA Completions / Total BA Completions (led 4 years).

(Mostly) standard DiD/Event-Study framework with staggered timing:

1. **Population-level dynamic effects:** Regress outcome on treatment leads and lags plus institution and *community-time* FEs.
2. **Community-level static effects:** Regress outcome on post-treatment indicator $\times$ community dummies plus institution and *community-time* FEs.

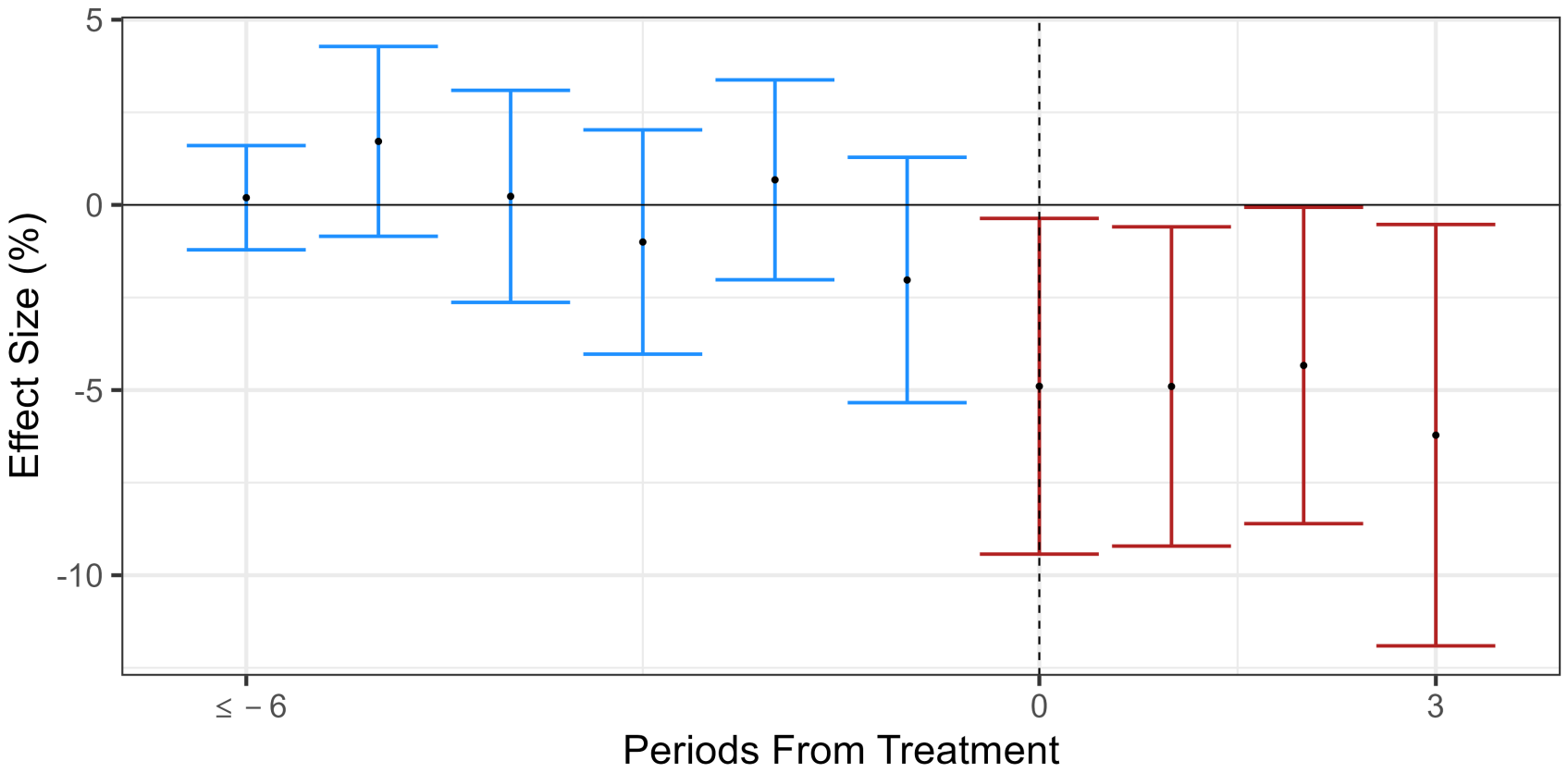
Estimate (1) and (2) using 2-stage DiD procedure from Gardner et al. (2024):

- Estimate FEs in stage 1 with pre-treatment data only, then regress outcomes minus estimated FEs on treatment indicators in stage 2.

Estimating Equations

STEM Majors

Figure: Proportion of Graduates w/ a STEM Major (Eq. 1)



Population Trends | Carnegie Communities

Cross-Community Heterogeneity

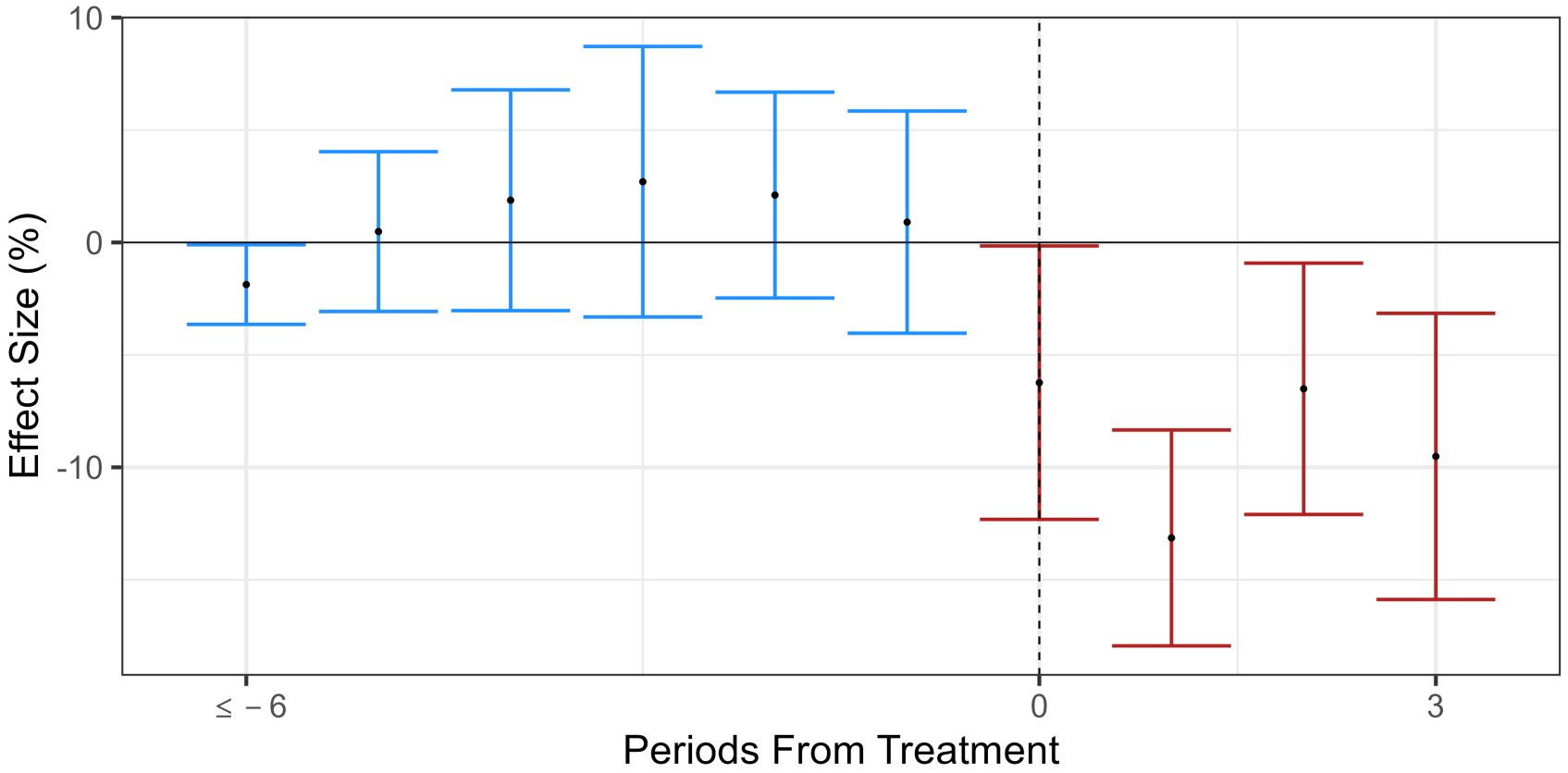
Table: Effects on Log Graduate STEM Share (Eq. 2)

	Overall	URM	non-URM
Less Selective Nat'l Unis	0.0565	0.0191	0.0568
	(0.1363)	(0.0749)	(0.1471)
LACs - New England	-0.0917	0.0901	-0.0979
	(0.0845)	(0.1970)	(0.0884)
Colleges and Unis - East	**-0.1571*****	**-0.2646***	**-0.1511*****
	(0.0590)	(0.1423)	(0.0541)
LACs - Midwest	0.0019	0.1151	-0.0070
	(0.0395)	(0.0879)	(0.0384)
Colleges and Unis - Midwest	-0.0973	0.0373	-0.1016*
	(0.0602)	(0.1922)	(0.0591)
Colleges and Unis - Southeast	0.0394	0.0430	0.0267
	(0.0350)	(0.1283)	(0.0401)
Elite Nat'l LACs	**-0.0837*****	**-0.2300*****	**-0.0664****
	(0.0314)	(0.0786)	(0.0314)
Elite Nat'l Unis	0.0539	0.0516	0.0597
	(0.0878)	(0.0892)	(0.0884)
Prob. All Effects Equal	$<$0.01	0.019	$<$0.01
Observations	2,714	2,714	2,714

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

STEM Major Decline at Elite LACs – All Students

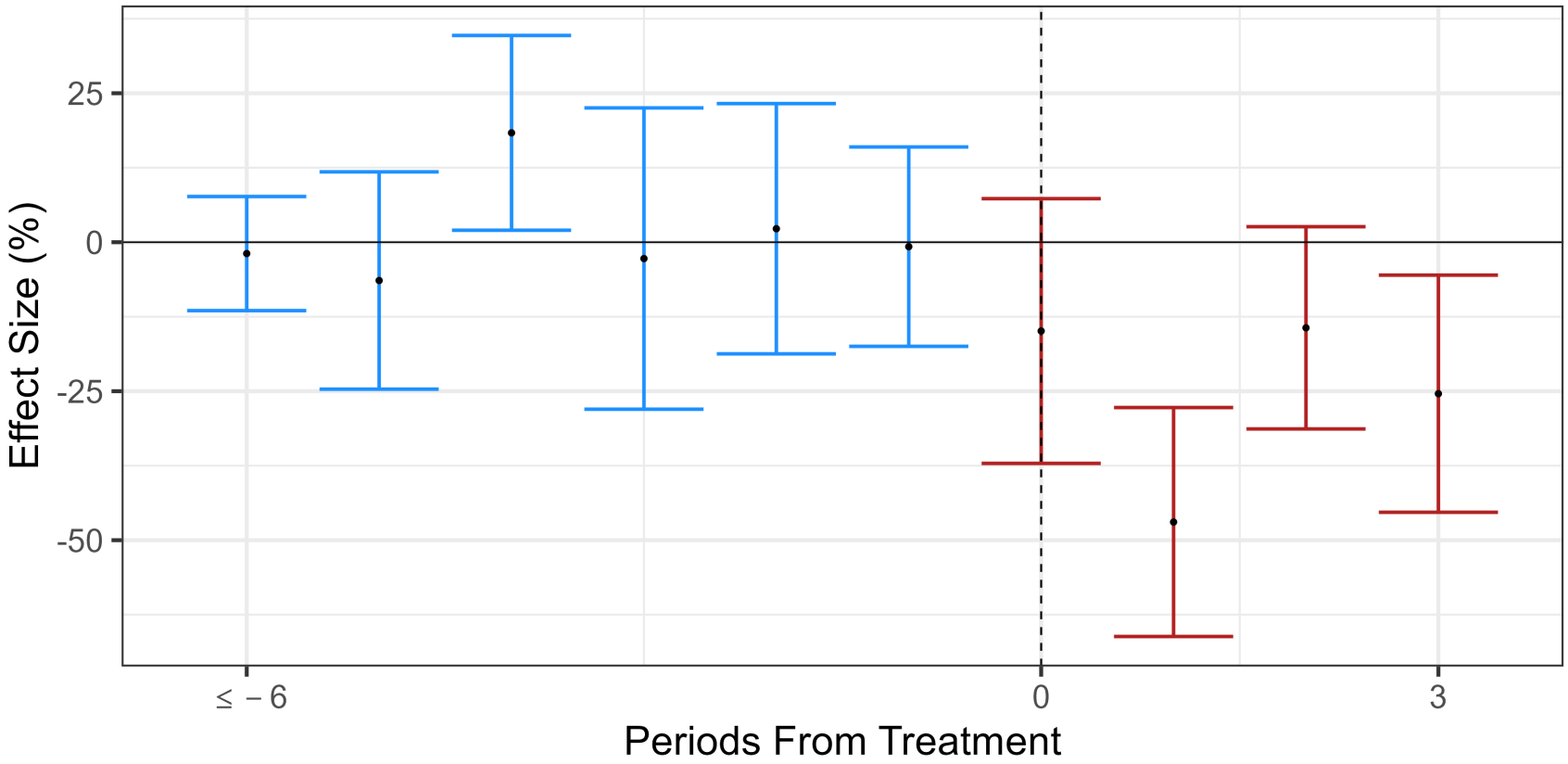
Figure: Proportion of Graduates w/ a STEM Major - Elite LACs



Colleges and Unis – East

URM STEM Major Decline at Elite LACs

Figure: Proportion of URM Graduates w/ a STEM Major - Elite LACs



Colleges and Unis – East

Conclusion

Thank You!

Questions/Comments: ethannl@bu.edu

Summary Statistics

Table: Selectivity, Enrollment, and Graduate STEM Share by Community and Treatment Status

Community	SAT 25th %ile		Admit Rate		FTFT Enrollment		Graduate STEM %	
	T	NT	T	NT	T	NT	T	NT
Less Selective National Universities	1046.00	1078.74	75.73	66.14	1500.54	1208.55	12.92	16.18
	(63.01)	(56.71)	(8.24)	(16.99)	(596.26)	(553.05)	(5.12)	(7.74)
Liberal Arts Colleges – New England	942.70	907.40	71.44	74.95	483.33	549.11	9.22	10.80
	(105.83)	(53.04)	(10.71)	(15.74)	(214.24)	(380.02)	(5.22)	(17.82)
Colleges and Universities – East	1044.93	932.89	64.79	74.76	865.87	534.17	10.70	8.84
	(74.07)	(76.69)	(11.03)	(12.23)	(368.72)	(366.21)	(3.88)	(4.54)
Liberal Arts Colleges – Midwest	1049.97	1031.22	69.68	76.66	387.27	465.77	18.00	18.34
	(68.13)	(62.45)	(11.88)	(8.09)	(76.02)	(169.16)	(6.82)	(3.62)
Colleges and Universities – Midwest	1091.00	931.84	69.43	76.74	444.07	331.36	19.38	12.78
	(40.28)	(118.28)	(12.42)	(12.97)	(157.57)	(159.38)	(3.86)	(6.48)
Colleges and Universities – Southeast	1016.20	1007.00	81.31	78.29	574.10	566.85	14.12	14.10
	(42.59)	(62.47)	(4.33)	(9.88)	(113.97)	(239.31)	(5.30)	(7.78)
Elite National Liberal Arts Colleges	1157.15	1251.21	51.60	41.04	589.33	520.49	15.84	20.08
	(73.07)	(59.27)	(14.44)	(15.94)	(105.25)	(160.30)	(2.86)	(5.90)
Elite National Universities	1218.35	1281.45	44.59	32.44	1316.87	1839.05	17.88	28.08
	(34.19)	(86.61)	(3.93)	(16.37)	(640.51)	(933.01)	(8.58)	(20.78)
Full Sample	1050.78	1080.60	66.35	63.39	639.17	771.48	14.30	16.28
	(107.59)	(164.42)	(14.45)	(22.77)	(414.68)	(694.65)	(6.46)	(12.40)

Note: Summary statistics are calculated using data from 2001–2005. Standard deviations are in parentheses. FTFT = First-time full-time. Graduate STEM % is the percent of graduates majoring in a STEM field. T denotes statistics conditional on eventual treatment; NT denotes never-treated units.

Back

Methodology

(Mostly) standard DiD framework with staggered timing, want to estimate

$$y_{ct} = \sum_{k=-a}^{b} \tau_k \mathbf{1}\{t - T_c = k\} + \gamma_c + \underbrace{\gamma_{gt}}_{\text{Community-Time FEs}} + \varepsilon_{ct} \tag{1}$$

and

$$y_{ct} = \sum_{g \in \mathcal{G}} \tau_g \mathbf{1}\{Post \wedge c \in g\} + \gamma_c + \gamma_{gt} + \varepsilon_{ct}, \tag{2}$$

Where c indexes colleges, t indexes years, and $\mathcal{G}$ is the set of all communities. τ_k in (1) are population-level, dynamic ATTs, and τ_g in (2) are community-level, static ATTs.

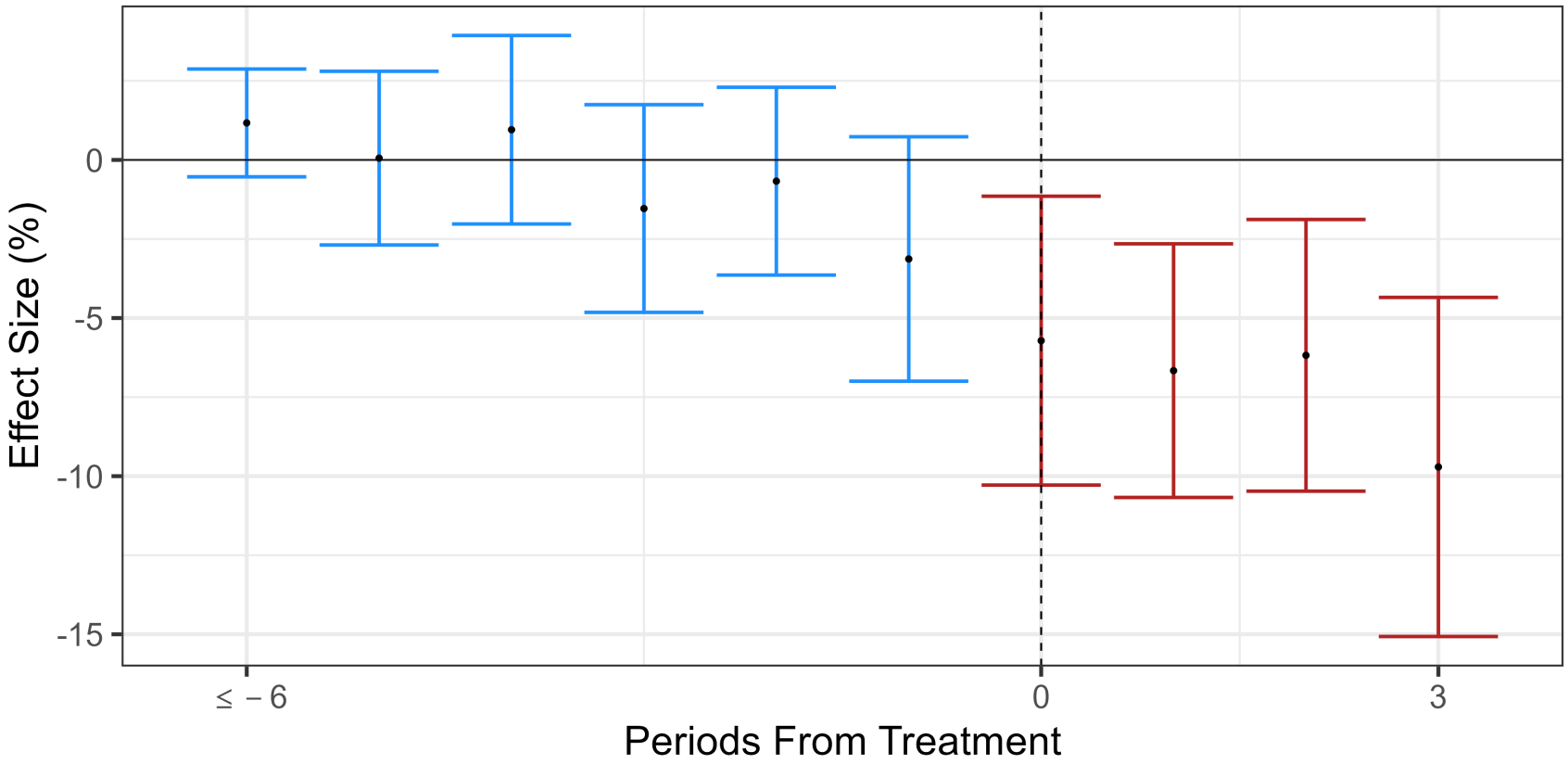
Estimate (1) and (2) using 2-stage DiD procedure from Gardner et al. (2024).

- Estimate $\hat{\gamma}_c$ and $\hat{\gamma}_{gt}$ from pre-treatment data, estimate $\hat{\tau}_{k/g}$ by regressing $y_{ct} - \hat{\gamma}_c - \hat{\gamma}_{gt}$ on relevant treatment indicators.

Back

Population Trends

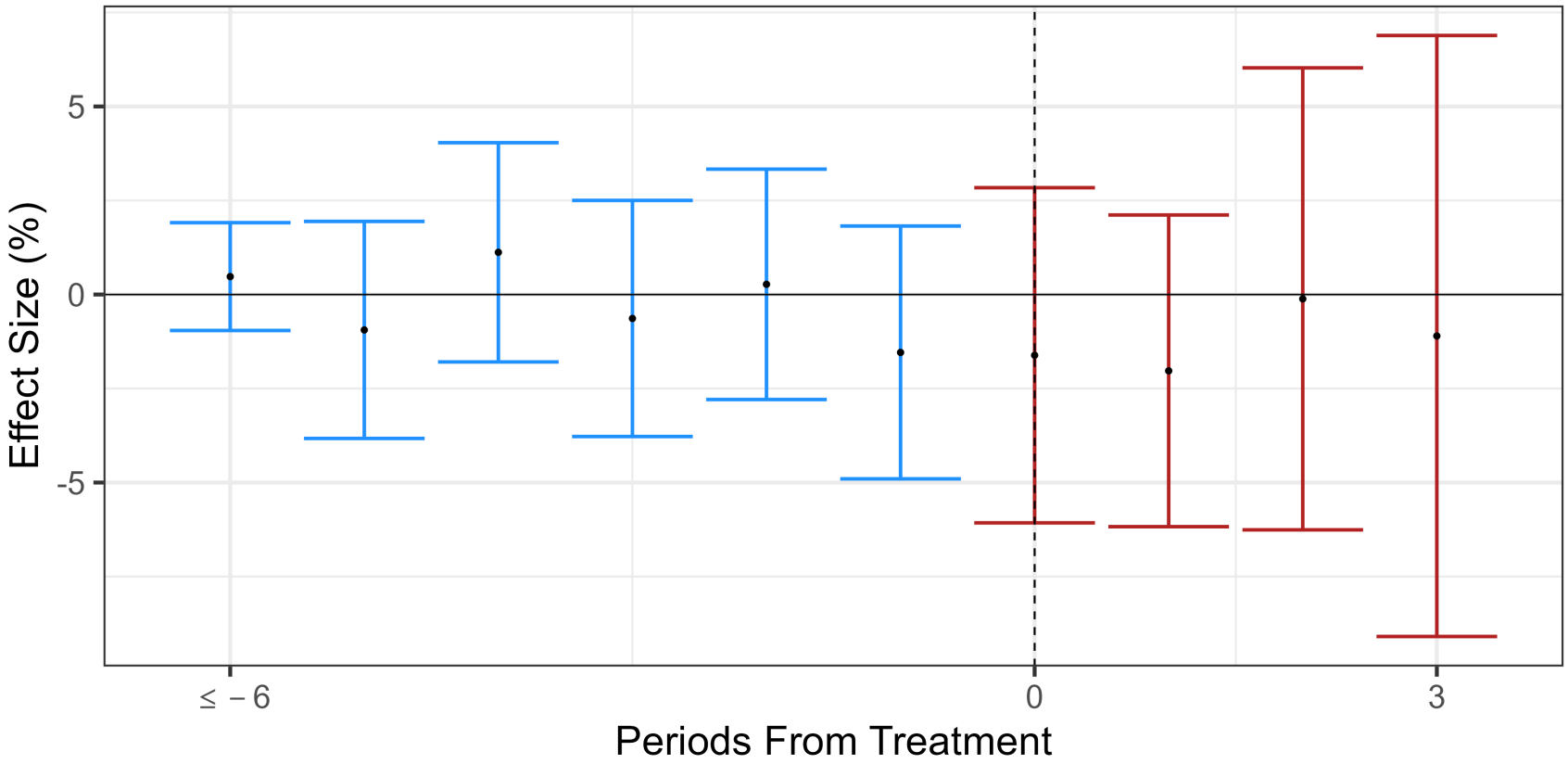
Figure: Proportion of Graduates w/ a STEM Major



Back

Trends Based on Alt. Classification

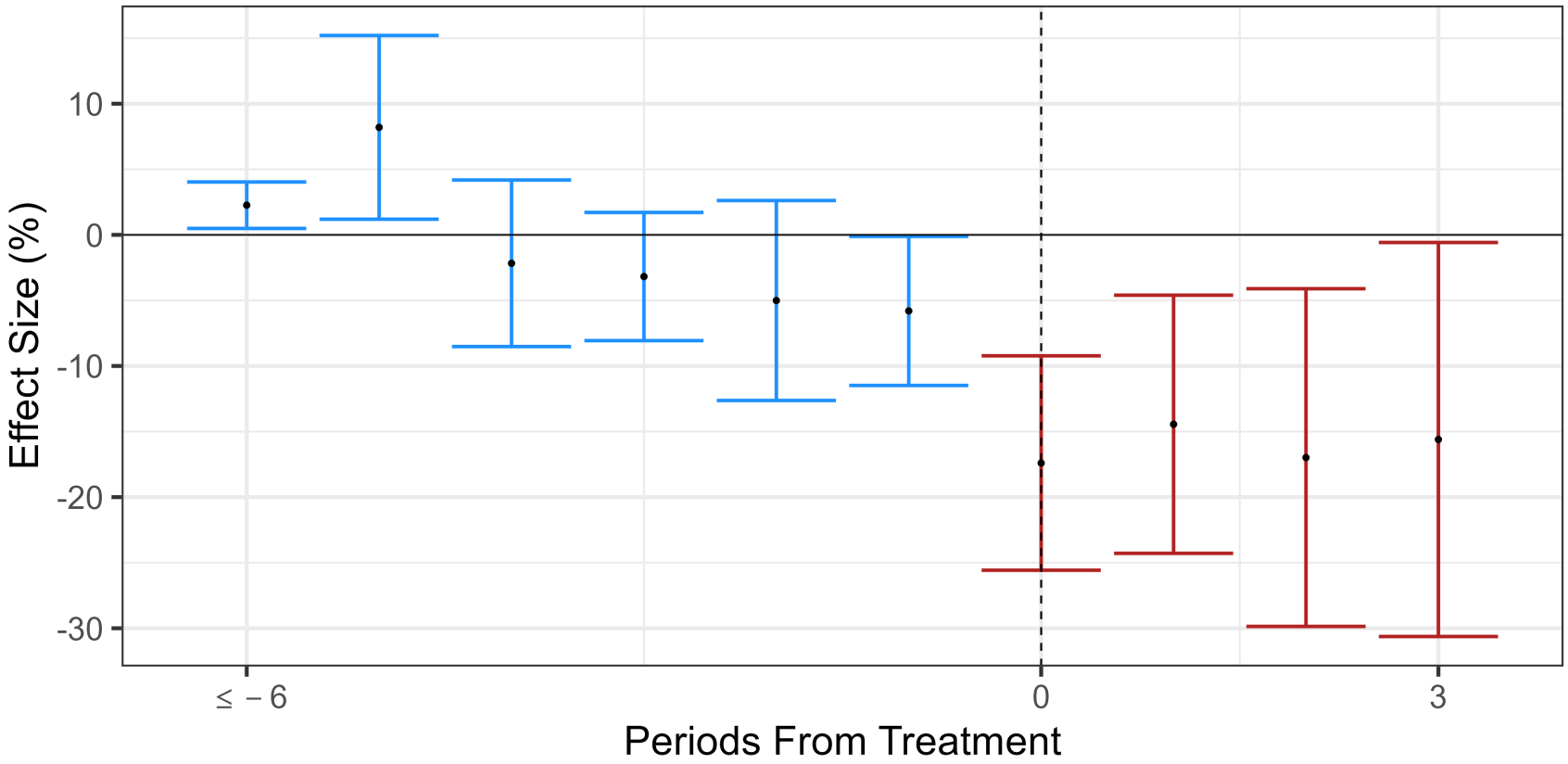
Figure: Proportion of Graduates w/ a STEM Major



Back

STEM Major Decline at Colleges and Unis – East

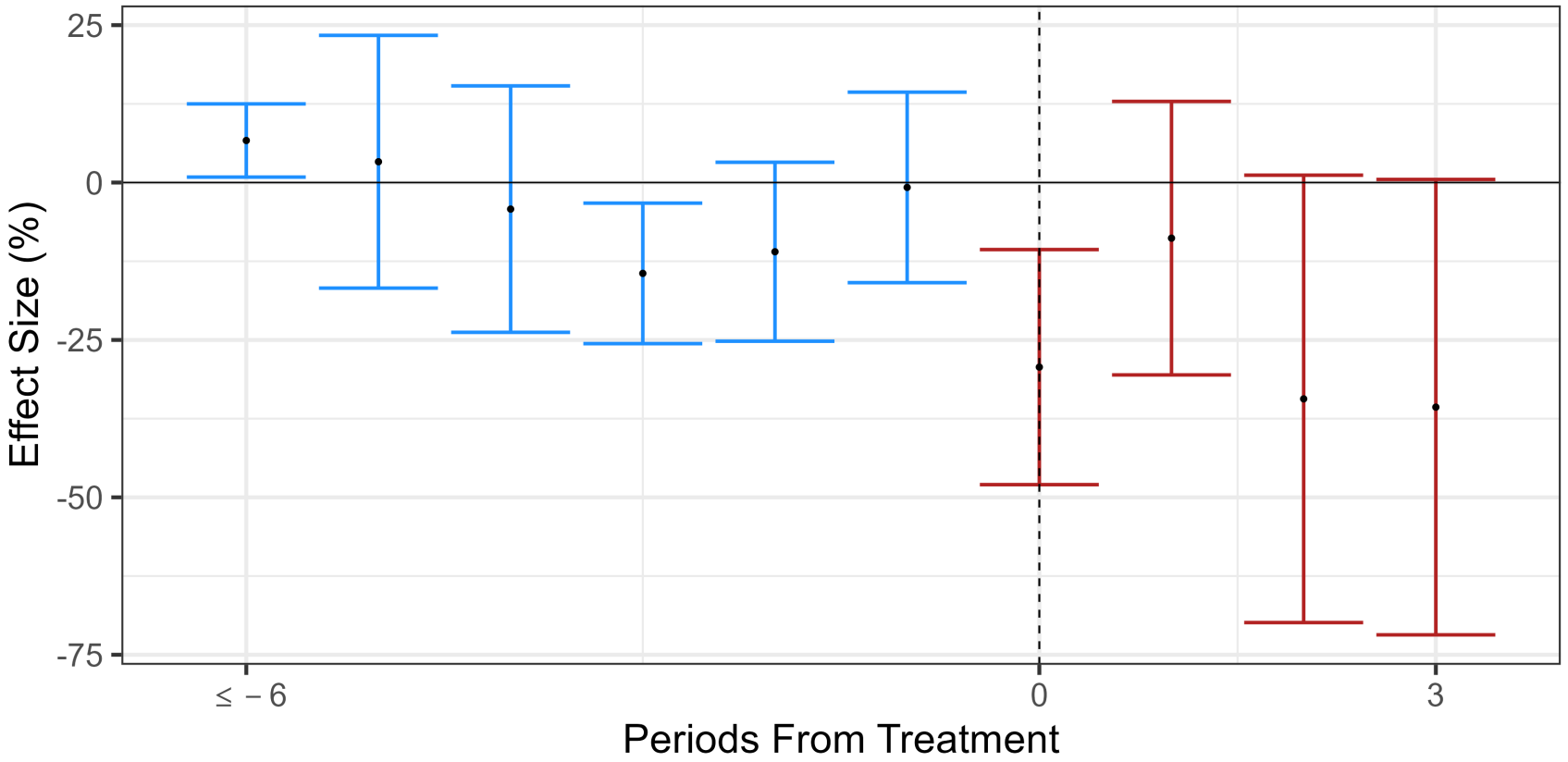
Figure: Prop. of Graduates w/ a STEM Major - Colleges and Unis East



Back

URM STEM Major Decline at Colleges and Unis – East

Figure: Prop. of URM Graduates w/ a STEM Major - Colleges and Unis East



Back

References I

Gardner, J., Thakral, N., Tô, L. T., and Yap, L. (2024). Two-stage differences in differences. *Working Paper*.

Traag, V. A., Waltman, L., and van Eck, N. J. (2019). From louvain to leiden: guaranteeing well-connected communities. *Scientific Reports*, 9(5233):1–12.