

The best causal model? A provocation on Agent-Based Models as causal inference tools

Bernardo Alves Furtado¹

August 25, 2025

Institute for Applied Economic Research (Ipea), Brazil

- Understanding causality (and mechanisms) remains central to science
- Strong assumptions and the challenge of counterfactuals persist
- Real-world phenomena are complex, heterogeneous, and non-linear

- **Counterfactuals by design:** agents follow explicit rules and can be held **identical** across runs
- Treated and untreated with the **same agents and controlled environment**
- Overcoming the "parallel trends" assumption

- Difference between treated and untreated [?]

$$\tau_i = Y_i(1) - Y_i(0) \quad (1)$$

- But. One cannot observe both. Thus, the need for estimation
- Assumption: no unobserved confounders affect both

Structural Causal framework

- Formal language (graphs) to make assumptions explicit [?]
- DAGs capture relationships between variables

- Stable Unit Treatment Value Assumption (SUTVA): outcome depends only on treatment
(no social interactions, contagion, or feedback effects)
- Confounding: factors influence treatment and outcomes
treatment: exercise
outcome: heart disease
confounder: smoke
- Staggered timing: policy slowly unfolding + feedback effects:
difficult for diff-in-diff methods

- Encode causal mechanisms explicitly, **control interventions**, allows comparability [? ? ?]
- Handles heterogeneity, feedback, and non-linear dynamics
- Micro-level (agents) + meso-level structures (networks/space/household/community)

Why ABMs: advantages

- **Control by design:** no (known) hidden confounders (just random)
- Endogenous staggered treatments: policies roll out dynamically (unlike rigid DiD)
- Feedback & heterogeneity: captures social spillovers, non-linearities

Not any ABM: rigour

- Theory-Driven: rules from grounded mechanisms (no ad-hoc tweaks).
- Empirically calibrated: parameters fit to real data (e.g., PS2's Brazilian regions).
- Transparent: code/documentation for replication
- Require careful specification (theoretical + empirical), validation and plausible rules and assumptions

An illustration

- PolicySpace2: empirically grounded, spatially explicit, open-source, fully documented
- Captures interactions, feedback, and endogenous heterogeneity in treatment timing
- **Workflow:** theory (rules) → run ABM, **get simulated data** → interpret outcomes with DAG → apply econometrics (diff-in-diff) → causal claims

Takeaway message

- Given rigour, ABMs serve as heuristic, exploratory, **or causal tools**
- ABMs alone—or combined with DAG-based interpretation and econometrics—embed interactions and feedback while providing in silico laboratories for unobservable phenomena

