Geography and Firm Boundaries

Holger Mueller

NYU, NBER, CEPR & ECGI

DGF Keynote Speech

September 2023

- What does it mean if two otherwise unrelated productive units (e.g., two manufacturing plants) are part of the same firm?

- This question forms the basis of an academic literature that is by now almost 100 years old: the literature on firm boundaries (a.k.a. "theory of the firm").

- This keynote: revisit the classic question posed by the literature on the boundaries of the firm through the lens of economic geography.

➢ "What does it mean if two otherwise unrelated productive units in different geographic regions are part of the same firm?"

Boundaries of the Firm

Boundaries of the Firm

- Coase (1937): "If production could be carried out without any organisation at all, well might we ask, why is there any organisation?"
  - Why are there firms ("Theory of the Firm")?
  - What difference does it make whether productive units (e.g., manufacturing plants) are part of the same firm or operate as separate stand-alone entities?
  - ➢ What determines the boundaries of the firm?

- Coase (1937): transaction costs.
  - Difficult to write and enforce fully contingent contracts that specify what should happen in all future situations ("contractual incompleteness").
  - ➢ Firms emerge in response to this inefficiency because a firm's owner can simply direct employees what do ("authority," control rights").
  - But what exactly are the (transaction) costs of contractual incompleteness that are mitigated in firms?

Boundaries of the Firm

- Williamson (1975, 1985), Klein, Crawford, and Alchian (1978): hold-up problem.
  - Contractual incompleteness: ex-post opportunism if parties make relationship-specific investments ex ante.
  - ➢ Anticipating ex-post holdup: parties underinvest in relationship-specific investments.
  - But how exactly is hold-up behavior mitigated within the firm?
- Grossman and Hart (1986), Hart and Moore (1990) (GHM): property-rights approach.
  - Bargaining power in ex-post (hold-up) renegotiations derives from asset ownership.
  - Asset ownership protects against hold-up and thus provides incentives to make ex-ante relationship-specific investments.
  - ➢ Party whose investment is relatively more important should own assets.
  - Is GHM a good model to think about real firms?

Boundaries of the Firm

- Reality check: control over firm's assets resides with headquarters (HQ):
  - … even though HQ itself is not the owner of the assets (shareholders are).
  - … even though HQ itself makes no relationship-specific investments.
  - "The Grossman-Hart Moore framework … predicts that control should be allocated to parties whose relationship-specific investments are most important to the relationship. Yet headquarters is given control, even though it does not really make such investments" (Bolton and Scharfstein, 1989).

- (Realistic) theory of firm boundaries should have:
  - Centralized decision-making by HQ.
  - HQ, while not the owner, has effective control over firm's assets.
  - HQ controls investment, giving more funds to some firm units and less to others.

Boundaries of the Firm

- "Capital-allocation-centric" view of firm boundaries (Alchian, 1969; Williamson, 1975; Stein, 1997):
  - "Loosely speaking, a collection of assets should optimally reside under the roof of a single firm to the extent that the firm's internal capital market can do a more efficient job of allocating capital to these assets than would the external capital market, if the assets were located in distinct firms" (Stein, 2003).
- Creates interdependence among otherwise unrelated productive units.
  - In other words, the units are linked because they are part of the same firm and HQ has authority to move resources from one unit to another.
  - "Thus, for example, if a company owns two unrelated divisions A and B, and the appeal of investing in B suddenly increases, the argument would seem to imply that investment in A would decline—even if it is positive NPV at the margin—as corporate headquarters channels relatively more of its scarce resources toward B" (Stein, 1997).

Geography

Geography

- Taking as given "interdependence among otherwise unrelated productive units," geography comes into play if the units are in different locations.
  - Shock to one firm unit affects other firm units in distant locations.
  - ➢ Shocks propagate across regions through firms' internal networks.
- If local shocks propagate across regions, they no longer cancel out in the aggregate (even if the shocks are idiosyncratic).
  - ➢ Can explain origins of aggregate fluctuations, similar to propagation of shocks through input-output/production (= supply-chain) networks (Acemoglu et al., 2012).
- Policy implications: local policy makers are unlikely to internalize spillover effects of local shocks on other regions.
  - ➢ Local industrial policies are inefficient from a national/welfare perspective.

Geography

- Remainder of talk: explore role of within- (or intra-) firm networks in propagating shocks across regions in more depth.

- How exactly do shocks to individual firm units affect other firm units in distant locations? Study two different scenarios:

I. Competition for scarce resources (e.g., firm-wide budget constraint). Central to notion of internal capital markets (Williamson, Stein).
   - (Negative) cash-flow shock (e.g., drop in local consumer demand).
   - (Positive) productivity shock.

II. Sharing of (non-rival) resources (e.g., knowledge, innovation).

I. Competition for Scarce Resources

Competition for Scarce Resources

- Binding firm-wide resource constraint:
  - Financial/budget constraint.
  - Other resource constraint (e.g., managerial attention).
- Role of HQ is to allocate scarce resources across regional firm units ("establishments") to maximize overall firm profits.
  - Establishment: e.g., industrial plant, restaurant, department store, supermarket.

➢ If resource constraint binds, shock to one unit triggers resource reallocation.
  - ➢ Affects other firm units in distant locations ("interdependence").

- How exactly are other firm units (in distant locations) affected? Need theoretical framework to guide empirical analysis:
  - (Negative) cash-flow shock (e.g., drop in local consumer demand).
  - (Positive) productivity shock.

Competition for Scarce Resources

Theory

Competition for Scarce Resources

- Firm operates in n regions, indexed by $i = 1, \ldots, n$.
- Regional firm units (e.g., plants) produce output using labor input with production function $\phi_i f_i(L_i)$, where $f_i' > 0$, $f_i'' < 0$, and where ϕ_i is productivity shifter.
- Regional firm units take output prices p_i and factor input prices w_i as given.
- Labor input in period t generates output in period $t + 1$, which is discounted using discount factor δ.
- Factor input costs are funded out of firm's current (period t) cash flow. Firm's budget constraint is $\sum_i w_i L_i \leq \sum_i C_i$ (total payroll $\leq$ total cash flow), where C_i is (current) cash flow produced by regional firm unit i.
- Factor input choices and funding decisions made centrally by HQ, which has authority to move budgets across regional firm units to maximize overall firm value (Williamson, 1975; Gertner, Scharfstein, and Stein, 1994; Stein, 1997).

➢ Budget constraint is at overall firm level, not at regional firm unit level!

Competition for Scarce Resources

- HQ solves:

$$\max_{L_i,\lambda} \delta \sum_i p_i \phi_i f(L_i) - \sum_i w_i L_i + \lambda[\sum_i C_i - \sum_i w_i L_i]$$

- Kuhn-Tucker conditions:

$$\delta p_i \phi_i f'(L_i) = (1+\lambda) w_i \ \forall i$$

$$\sum_i w_i L_i \leq \sum_i C_i$$

$$\lambda[\sum_i C_i - \sum_i w_i L_i] = 0; \lambda \geq 0$$

- Optimality condition. For any two regional firm units i and j it must hold that:

$$\boxed{\frac{\delta p_i \phi_i f'(L_i)}{w_i} = 1 + \lambda = \frac{\delta p_j \phi_j f'(L_j)}{w_j}}$$

➢ A marginal dollar of funds must have same value at each regional firm unit!

Competition for Scarce Resources

- Financially unconstrained firm ($\lambda = 0$):
  - Optimality condition: HQ allocates labor input to each region until MRP of labor, $\delta p_i \phi_i f'(L_i)$, equals wage, w_i.
  - ➢ Labor input in each region is at first-best optimum.
  - Negative cash-flow shock in region *j*: budget constraint is slack, so no effect on labor input anywhere (no "real" effects).
  - Positive productivity shock in region *j*: labor in region *j* becomes more productive, so labor input there increases. As budget constraint is slack, we can easily fund this new hiring, no effect on labor input in other regions.
- Financially constrained firm ($\lambda > 0$):
  - Optimality condition: HQ allocates labor input to each region until MRP of labor, $\delta p_i \phi_i f'(L_i)$, equals wage plus a wedge, $w_i(1 + \lambda)$.
  - ➢ Labor input in each region is below first-best optimum.

Competition for Scarce Resources

- Negative cash-flow shock in region *j* (assuming $\lambda > 0$):
  - Negative cash-flow shock in region *j* raises shadow value of marginal dollar of funds, $1 + \lambda$, and thus tightens firm's overall budget constraint:

$$\frac{d\lambda}{dC_j} = \frac{1}{\sum_i \frac{w_i^2}{\delta p_i \phi_i f''(L_i)}} < 0$$

  - HQ adjusts production in each region to ensure that optimality condition remains satisfied.
  - Since regional firm units exhibit decreasing returns to scale, $f_i'' < 0$, labor input must decline in all regions, including other regions $i \neq j$:

$$\frac{dL_i}{dC_j} = \frac{w_i}{\delta p_i \phi_i f''(L_i)} \frac{d\lambda}{dC_j} > 0 \;\; \forall\, i$$

➢ If firm is financially constrained, negative cash-flow shock in one region leads to reduction in labor everywhere: HQ optimally "spreads" shock across regions.

Competition for Scarce Resources

- Positive productivity shock in region *j* (assuming $\lambda > 0$):
  - Labor in region *j* becomes more productive, so labor input there increases:

$$\frac{dL_j}{d\phi_j} = \frac{w_j f'(L_j)}{\phi_j f''(L_j)} \left[\frac{-\sum_{i \neq j} \frac{w_i}{\delta p_i \phi_i f''(L_i)}}{\sum_i \frac{w_i^2}{\delta p_i \phi_i f''(L_i)}} \right] > 0$$

  - If firm is financially constrained, increase in labor in region *j* must be "funded" by withdrawing (scarce) resources from other regions => labor input in other regions $i \neq j$ declines ("labor reallocation"):

$$\frac{w_j f'(L_j)}{\phi_j f''(L_j)} \left[\frac{\frac{w_i}{\delta p_i \phi_i f''(L_i)}}{\sum_i \frac{w_i^2}{\delta p_i \phi_i f''(L_i)}} \right] < 0 \;\; \forall\, i \neq j$$

➢ If firm is financially constrained, positive productivity shock in region *j* leads to "winner picking:" labor in region *j* increases at expense of other regions $i \neq j$.

Competition for Scarce Resources

Empirical Predictions (Labor)

		Region j	Regions i ≠ j
Negative Cash-Flow Shock	No FC	-	-
	FC	⬇️	⬇️
Positive Productivity Shock	No FC	⬆️	-
	FC	⬆️	⬇️

Testing the Theory

Negative Cash-Flow Shocks

(Giroud and Mueller, 2019)

Competition for Scarce Resources

- Study how local "consumer demand shocks" affect employment of large, multi-establishment, multi-region firms in non-tradable sector.
- Non-tradable industries: e.g., restaurants, supermarkets, retail stores.
  - Rely on local consumer demand (ideally suited to study local shocks).
- Local consumer demand shocks: drop in local house prices during Great Recession.
  - House price collapse caused sharp drop in local consumer spending (Mian, Rao, and Sufi, 2013; Stroebel and Vavra, 2019).
  - Negatively impacts local firms' cash flows ("cash-flow shock").
- Mian and Sufi (2014) & Giroud and Mueller (2017): non-tradable establishments experienced larger employment losses in regions with larger house price drops.
- Q: do local consumer demand shocks also affect firms' other establishments in distant regions?
  - Recall: non-tradable industries => no direct effect on distant regions!

Competition for Scarce Resources

Consumer spending & employment decline
Mian, Rao, and Sufi, 2013;
Mian and Sufi, 2014;
Giroud and Mueller, 2017

Competition for Scarce Resources

Competition for Scarce Resources

Competition for Scarce Resources

- No FC: no effect on either local establishments (Giroud and Mueller, 2017) or distant establishments of same firm.
- FC: 10% drop in local house prices associated with
  - 1.1% drop in employment at local establishments.
  - 0.3% drop in employment at distant establishments of same firm.
- Identification is within ZIP code × industry x year:
  - Compare establishments within same ZIP code and industry—e.g., local Aldi and Lidl—that face same local demand shock but are differentially exposed to demand shocks in other regions because their parent firms have different regional networks.
  - ➢ FE: results not driven by common shocks between regions in which parent firm operates => regional spillovers through firms' internal networks of establishments!
- Regional spillovers through firms' internal networks also matter for aggregate employment.
  - Employees laid off due to shocks in other regions are not re-employed by local firms => shocks not only propagate across establishments but also across regions.

Competition for Scarce Resources

Empirical Predictions (Labor)

		Region j	Regions i ≠ j
Negative Cash-Flow Shock	No FC	-	-
	FC	⬇️	⬇️
Positive Productivity Shock	No FC	⬆️	-
	FC	⬆️	⬇️

Competition for Scarce Resources

Empirical Predictions (Labor)

		Region j	Regions i ≠ j
Negative Cash-Flow Shock	No FC	- ✓	- ✓
	FC	⬇️ ✓	⬇️ ✓
Positive Productivity Shock	No FC	⬆️	-
	FC	⬆️	⬇️

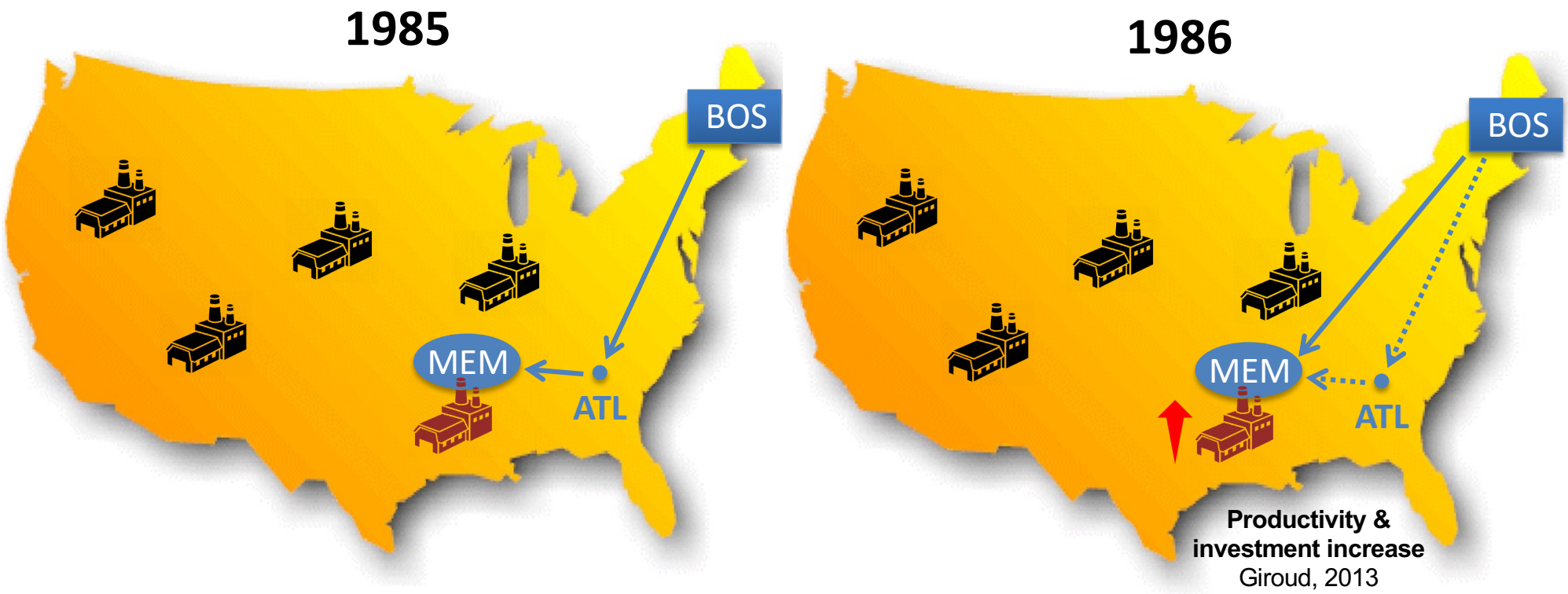
Positive Productivity Shocks

(Giroud and Mueller, 2015)

Competition for Scarce Resources

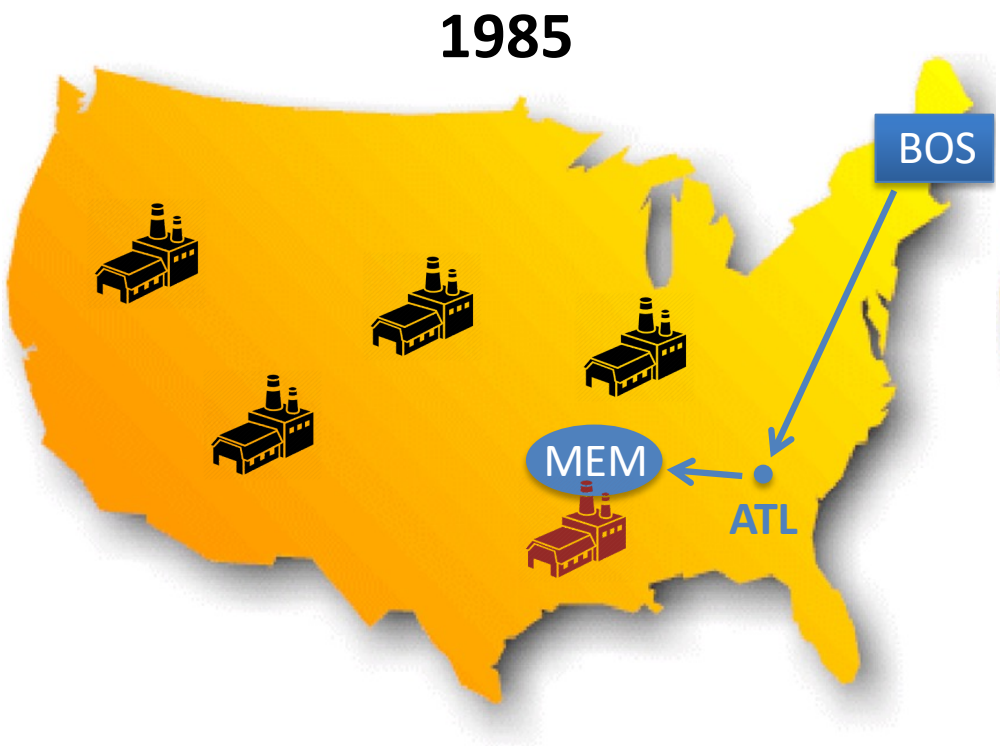
- Study how plant-specific productivity shocks affect investment and employment of large, multi-plant, multi-region manufacturing firms.
- Plant-specific productivity shocks: introduction of new airline routes that reduce travel time between HQ and ("treated") plants.
  - Travel time reduction makes it easier for HQ to monitor plant, give advice, share knowledge, etc.
  - ➢ Giroud (2013): productivity (and investment) at treated plant increases.
- Q: do plant-specific productivity shocks affect firms' other plants in distant regions?
  - Distant plants: experience no travel time reduction/productivity shock themselves.

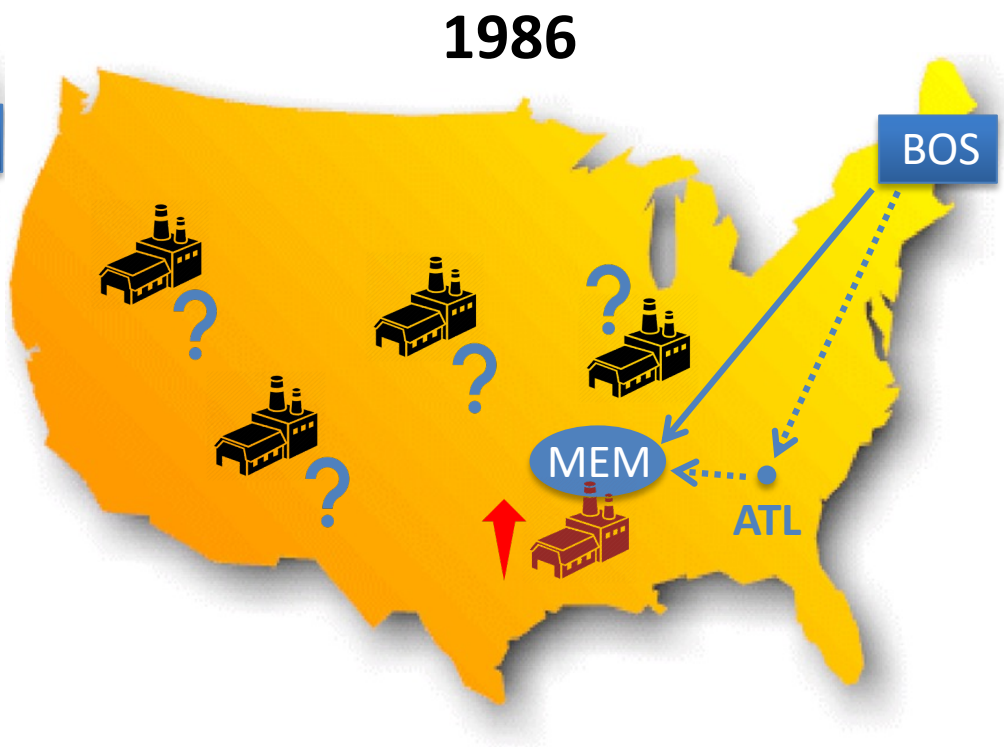
Competition for Scarce Resources



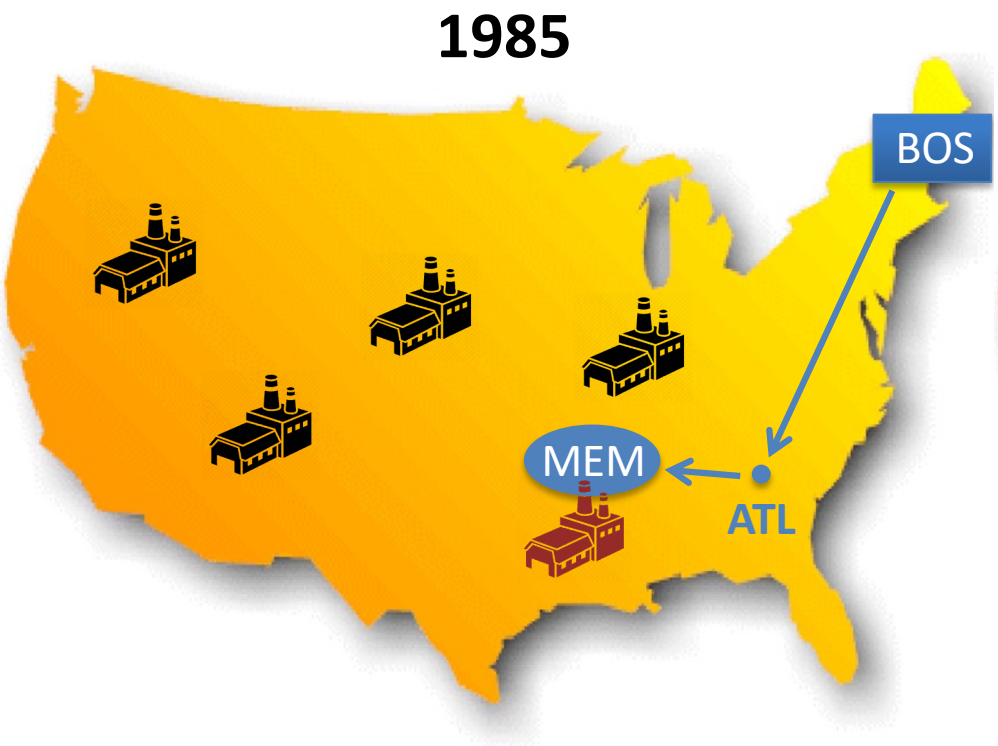
1985: fastest way to fly from BOS to MEM is Delta flight with stopover in ATL
1986: Northwest introduced direct flight from BOS to MEM

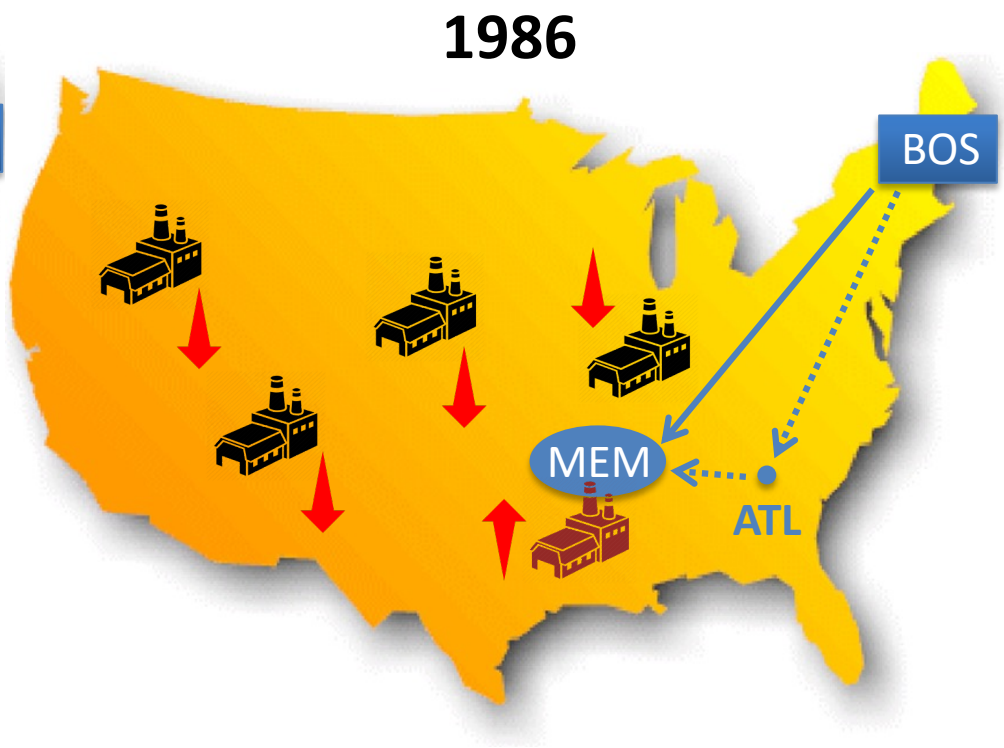
Competition for Scarce Resources





Competition for Scarce Resources





Competition for Scarce Resources

- No FC: investment and employment at treated plant increase. No effect on other plants of same firm.
- FC: investment and employment at other plants decline. Decline matches increase at treated plant.
  - Investment at all other plants combined drops by $186,000 compared to $179,000 increase at treated plant.
  - Employment at all other plants combined drops by 6 employees compared to increase of 5 employees at treated plant.
- Identification is within MSA × year:
  - Compare "other" plants with local plants in same MSA that are differentially (or not at all) exposed to travel time reduction in their respective firm networks.
  - ➢ FE: results not driven by common shocks between regions in which parent firm operates => regional spillovers through firms' internal plant-level networks!
- Zero-sum game? No! Reallocation of scarce resources (capital & labor) from other plants to (more productive) treated plant benefits firm overall.

Competition for Scarce Resources

Empirical Predictions (Labor)

		Region j	Regions i ≠ j
Negative Cash-Flow Shock	No FC	- ✓	- ✓
	FC	⬇️ ✓	⬇️ ✓
Positive Productivity Shock	No FC	⬆️	-
	FC	⬆️	⬇️

Competition for Scarce Resources

Empirical Predictions (Labor)

		Region j	Regions i ≠ j
Negative Cash-Flow Shock	No FC	- ✓	- ✓
	FC	⬇️ ✓	⬇️ ✓
Positive Productivity Shock	No FC	⬆️ ✓	- ✓
	FC	⬆️ ✓	⬇️ ✓

II. Sharing of (Non-Rival) Resources

Sharing of (Non-Rival) Resources

- Non-rival resources: e.g., knowledge, ideas, innovations.
  - Non-rivalry: consumption of resource does not prevent others from consuming it.
- Local (knowledge, innovation) "shock" at one regional firm unit => HQ can use resource also at other firm units in distant locations.
  - Markusen (1984): "Once an innovation is made, it can be incorporated into any number of additional plants without reducing the marginal product of that innovation in existing plants."
  - ➢ Propagation and amplification of local shocks (same resource is used multiple times).
- HQ has no reason to share resource with other firms.
  - If anything, incentives not to share (e.g., competitors).
  - ➢ Shocks propagate to distant locations, but only within firm boundaries.

Sharing of (Non-Rival) Resources

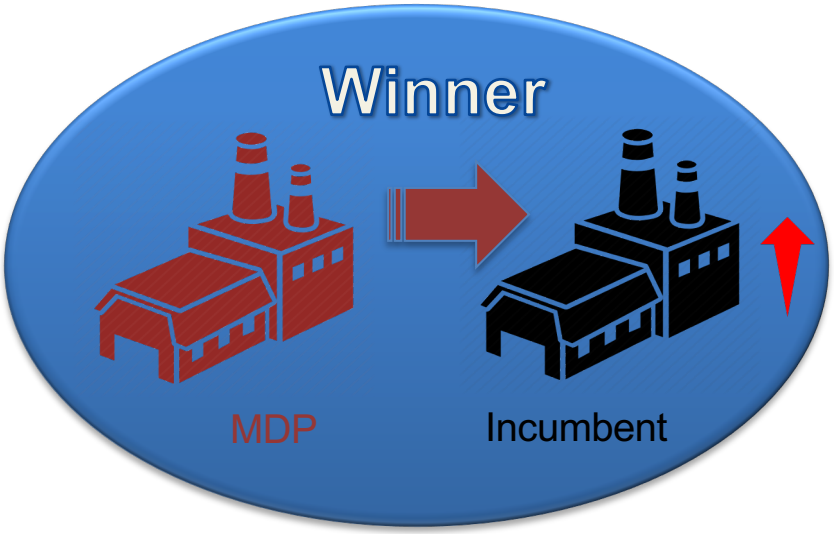
“Knowledge Shocks”

(Giroud, Lenzu, Maingi, & Mueller, 2023)

Sharing of (Non-Rival) Resources

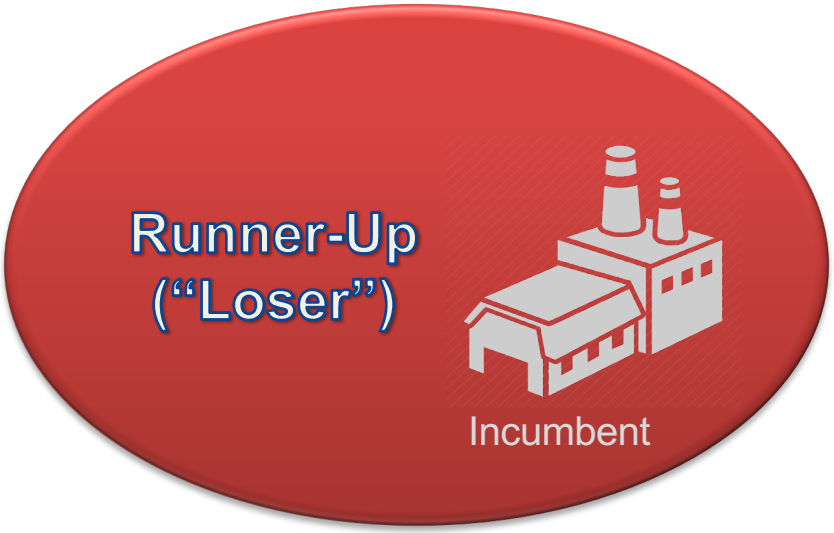
- Study how local "knowledge shocks" affect productivity and employment of large, multi-plant, multi-region manufacturing firms.
- Local knowledge shocks: openings of large manufacturing plants ("Million Dollar Plants" or MDPs) in a county.
  - MDP openings raise productivity of local incumbent plants (Greenstone, Hornbeck, and Moretti, 2010).
  - Mechanism: knowledge spillover.
- Hypothesis (Markusen): once the knowledge spills over (from the MDP) to a firm's local plant, HQ can freely share it also with other plants of the firm.
- Q: do local knowledge shocks (also) raise productivity of firms' other plants in distant locations?

Local Productivity (Knowledge) Spillover



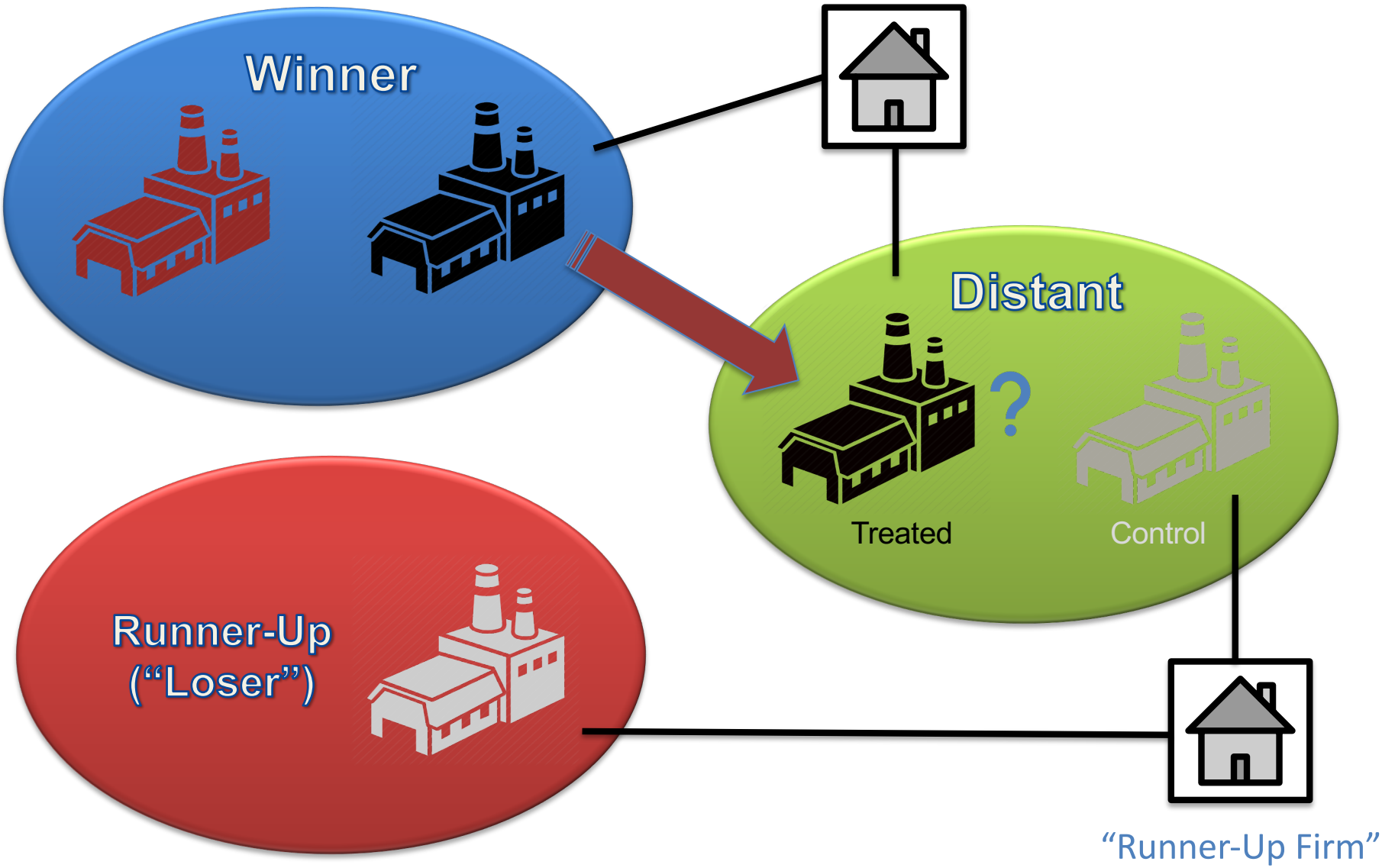
Winner county: where MDP located

Productivity increases
Greenstone, Hornbeck, and Moretti, 2010



Runner-up county: where MDP "almost" located (counterfactual)

Global Productivity (Knowledge) Spillover



Sharing of (Non-Rival) Resources

- Local productivity spillover: MDP openings raise productivity of local incumbent plants by 4% and employment by 3.5%.
- Global productivity spillover: productivity at distant plants of firms exposed to local spillover (i.e., firms with plants in winner county) increases by 1.8% and employment increases by 1.6%.
- Identification is within county × industry × year:
  - Compare distant plants of firms exposed to local spillover with other plants in same county and industry that belong to "runner-up firms."
  - ➢ FE: results not driven by common shocks between winner county and distant county => regional spillovers through firms' internal plant-level networks!
- Mechanism? Within-firm knowledge sharing!
  - Global productivity spillover does not decay with geographical distance.
  - Spillover is stronger if MDP and distant plant are in same industry or in (knowledge-based) industries characterized by mutual R&D flows or mutual patent citations.
  - (NB: global spillover is unrelated to firm-wide financial constraints.)

Beyond Reduced-Form Evidence

Concluding Remarks

Concluding Remarks

- Theory of the firm: what does it mean if two otherwise unrelated productive units are part of the same firm?
- "Resource-allocation-centric" view of firm boundaries: productive units that are part of the same firm share common resources.
  - Capital, managerial attention: (rival).
  - Knowledge, innovations (non-rival).
- HQ has authority to move resources across firm units.
  - ➢ Interdependence among otherwise unrelated productive units.
- Geography: interdependence => linkages across regions.
  - ➢ Local shock at one firm unit affects other firm units in distant locations.
  - ➢ Shocks propagate across regions through firms' internal networks.
- Macro implications: local shocks can spread and amplify.
- Policy implications: local policy makers unlikely to internalize spillovers on other regions; policies inefficient from a national/welfare perspective.

Concluding Remarks

- Propagation of shocks through firms' internal networks (still) relatively understudied relative to other propagation mechanisms.
  - Trade flows, labor migration (QSMs).
  - Credit supply shocks through banking/financial networks.
  - Input-output (= supply chain) production networks.
- Empirical challenges:
  - Granular establishment-level data (e.g., confidential US Census data).
  - Well-identified local (or establishment-specific) shocks.
- Theory? "Resource-allocation-centric" models of firm boundaries based on resource allocation by HQ are highly stylized (e.g., Stein, 1997).
  - No geography!
  - Spatial GE models in which firms choose optimal number, size, and locations of establishments extremely difficult to solve (e.g., Oberfield and Rossi-Hansberg, 2023).

Thank You!

Beyond Reduced-Form Evidence

Structural Estimation

- Reduced-form empirical evidence can:
  - Identify a causal "effect" (here: spatial propagation of local productivity shocks).
  - Shed light on underlying mechanism(s).
- However, many questions which reduced-form evidence cannot answer:
  - Aggregate (e.g., welfare) effects?
  - Comparison with other propagation mechanisms (e.g., supply-chain networks).
  - Policy counterfactuals (e.g., industrial policies to attract plants).
- Need quantitative spatial GE model!

Quantitative Spatial Model (QSM) with Knowledge Sharing & Within-Firm Plant-Level Networks

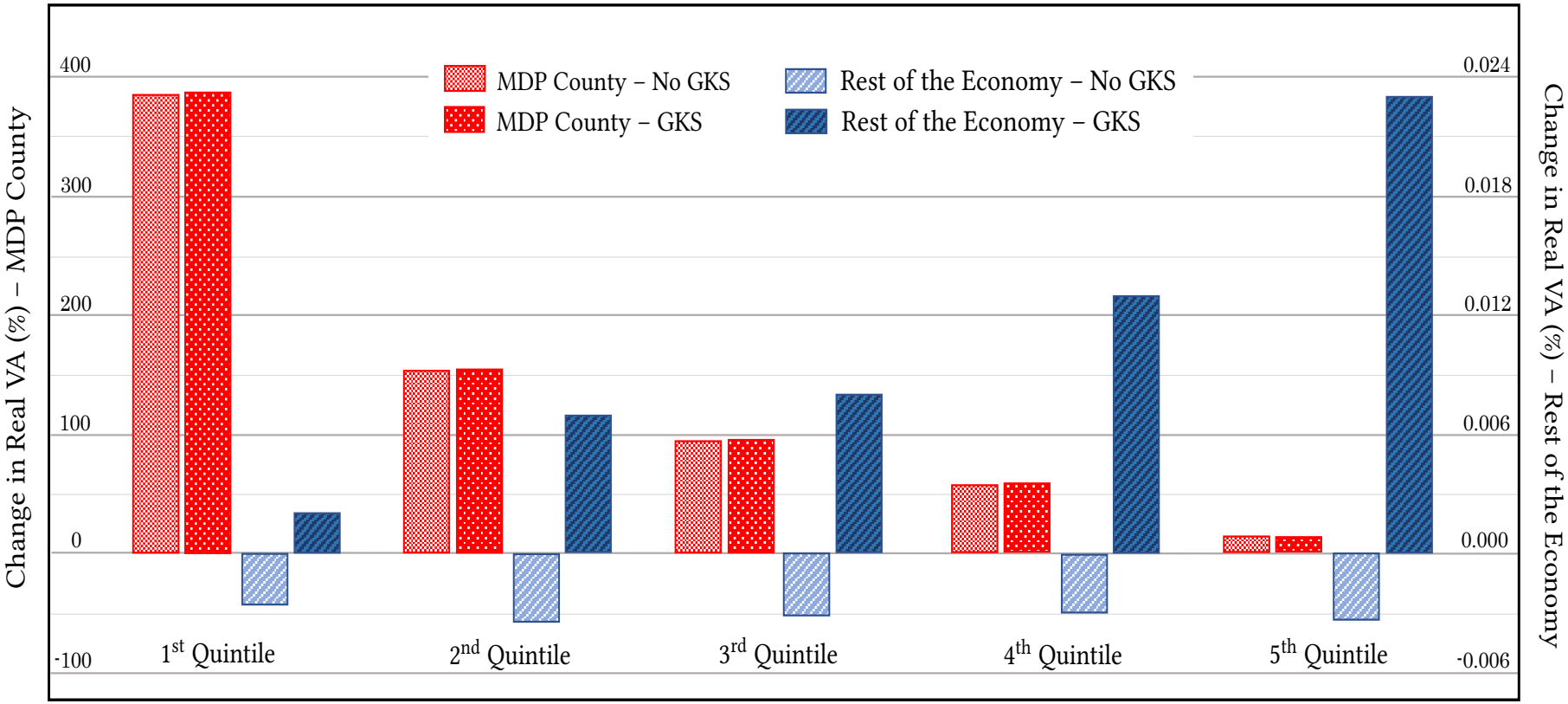
(Giroud, Lenzu, Maingi, & Mueller, 2023)

Structural Estimation

- Build QSM where plants of large multi-region firms are linked through shared knowledge.
- Estimate model by targeting as moments reduced-form DiD estimates: employment, wage, and productivity responses to MDP openings.
- Generate corresponding model-based DiD estimates:
  - Build model economy that closely mirrors geography of production networks in U.S. Census data (312,633 plants).
  - Within model economy, simulate actual MDP openings.
  - Estimate plant-level DiD regressions that mirror those in reduced-form analysis.
- Estimated model matches all targeted moments exactly!

Structural Estimation

Simulating (counterfactual) MDP openings in more or less developed regions



Quintiles: sort U.S. counties into quintiles by population ("development")

Conclusion