Propagation and Amplification of Local Productivity Spillovers

Xavier Giroud Simone Lenzu Quinn Maingi Holger Mueller

Online Appendix

A Additional Figures and Tables

Table A.1: GHM Replication

Column (2) presents a variant of the regression in column (1) of Table 2 using the original specification and sample selection procedure from GHM. For comparison, column (1) shows the baseline estimate from Table 5 of GHM.

	TFP	
	GHM Table 5 Column (4)	Replication
	(1)	(2)
MDP	0.0477	0.049
	(0.0231)	(0.022)
Plant FE	Yes	Yes
Industry × year FE	Yes	Yes
Case FE	Yes	Yes
R-squared	0.986	0.98
Observations	28,732	29,000

Table A.2: Treatment Effect Dynamics

This table presents estimates from event-study regressions using OLS (columns (1) and (3)) or the imputation estimator of Borusyak, Jaravel, and Spiess (2023, BJS) (columns (2) and (4)). The BJS estimates are obtained using the Stata code from Kirill Borusyak's website. MDP(0) denotes the year of the MDP opening. The base year is τ = -5. Only the main coefficients of interest are shown. Observations are weighted by plant-level employment. Standard errors are double clustered at the county and year level. The sample period is from 1977 to 1998.

	TFP			
	Local Spillover		Global Spillover	
	(1)	(2)	(3)	(4)
MDP(-4)	-0.004	-0.005	0.001	0.002
	(0.024)	(0.023)	(0.013)	(0.014)
MDP(-3)	0.006	0.004	0.004	0.001
	(0.024)	(0.024)	(0.011)	(0.014)
MDP(-2)	-0.005	-0.008	-0.003	0.001
	(0.021)	(0.023)	(0.010)	(0.012)
MDP(-1)	0.002	-0.004	0.001	0.000
	(0.020)	(0.022)	(0.009)	(0.010)
MDP(0)	0.017	0.016	0.006	0.009
	(0.018)	(0.014)	(0.008)	(0.007)
MDP(1)	0.038	0.035	0.015	0.017
	(0.018)	(0.015)	(0.008)	(0.009)
MDP(2)	0.041	0.044	0.019	0.020
	(0.020)	(0.016)	(0.010)	(0.009)
MDP(3)	0.049	0.043	0.022	0.021
	(0.022)	(0.018)	(0.010)	(0.009)
MDP(4)	0.050	0.057	0.024	0.022
	(0.023)	(0.020)	(0.011)	(0.011)
MDP(5)	0.048	0.051	0.023	0.022
	(0.022)	(0.021)	(0.011)	(0.010)
Plant FE	Yes	Yes	Yes	Yes
Industry × year FE	Yes	Yes	-	-
Industry × county × year FE	-	-	Yes	Yes
Case FE	Yes	Yes	Yes	Yes
R-squared	0.88	-	0.88	-
Observations	157,000	157,000	423,000	423,000

Table A.3: Input-Output Linkages & Knowledge Flows (Local Spillover)

This table presents variants of the regressions in columns (1) and (2) of Table 8 (columns (1) and (2)) or Table 7 (columns (3) to (5)) applied to the local productivity spillover. Only the main coefficients of interest are shown. Observations are weighted by plant-level employment. Standard errors are double clustered at the county and year level. The sample period is from 1977 to 1998.

	TFP				
	Input flows	Output flows	Same industry	Mutual R&D flows	Mutual patent citations
	(1)	(2)	(3)	(4)	(5)
MDP	0.037	0.038	0.039	0.034	0.028
	(0.017)	(0.016)	(0.017)	(0.017)	(0.016)
MDP × MDP industry	0.594	0.339	0.032	0.997	0.723
	(0.618)	(0.476)	(0.015)	(0.401)	(0.334)
Plant FE	Yes	Yes	Yes	Yes	Yes
Industry × year FE	Yes	Yes	Yes	Yes	Yes
Case FE	Yes	Yes	Yes	Yes	Yes
R-squared	0.88	0.88	0.88	0.88	0.88
Observations	157,000	157,000	157,000	157,000	157,000

Table A.4: MDP Size

This table presents variants of the regressions in column (1) of Table 2 and column (3) of Table 3 in which both terms in equation (1) are interacted with the MDP's county-level employment share at the date of entry ("MDP size"). Only the main coefficients of interest are shown. Observations are weighted by plant-level employment. Standard errors are double clustered at the county and year level. The sample period is from 1977 to 1998.

	TFP	
	Local Spillover	Global Spillover
	(1)	(2)
MDP	0.023	0.010
	(0.018)	(0.009
MDP × MDP size	0.249	0.113
	(0.116)	(0.052)
Plant FE	Yes	Yes
Industry × year FE	Yes	-
Industry × county × year FE	-	Yes
Case FE	Yes	Yes
R-squared	0.88	0.88
Observations	157,000	423,000

B Model Appendix

B.1 Derivations

B.1.1 Consumer Problem

In this section, we derive the labor supply functions (equations (10) and (13)) and average realized utility (equation (14)) as part of the solution to the consumer problem.

Equation (3) implies that the CDF of indirect utility is given by:

$$\mathbb{P}\left(\bigcap_{n\in\mathcal{N}}\bigcap_{sj\in\mathcal{E}_n}\left\{(1+d)\,b_{nsj}\frac{w_{nsj}}{P_n^{\alpha}R_n^{1-\alpha}}\leq t_{nsj,}\right\}\right)$$
$$= exp\left\{-\sum_{n\in\mathcal{N}}\left(\sum_{sj\in\mathcal{E}_n}(B_nB_s)^{\frac{1}{1-\rho}}\left((1+d)\,\frac{w_{nsj}}{P_n^{\alpha}R_n^{1-\alpha}}\right)^{\frac{\epsilon}{1-\rho}}t_{nsj}^{-\frac{\epsilon}{1-\rho}}\right)^{1-\rho}\right\}. \qquad \text{(A.1)}$$

Using the notation of Lind and Ramondo (2023), the multivariate Fréchet CDF in equation (A.1) is generated by the Archimedean copula:

$$G\left(\left\{x_{nsj}\right\}_{n\in\mathcal{N}:sj\in\mathcal{E}_n}\right)=\sum_{n\in\mathcal{N}}\left(\sum_{sj\in\mathcal{E}_n}x_{nsj}^{\frac{1}{1-\rho}}\right)^{1-\rho}$$

with scale parameters $T_{nsj} = (1+d)\,B_nB_s\left(\frac{w_{nsj}}{P_n^{\alpha}R_n^{1-\alpha}}\right)^{\epsilon}$. By Lemma O.5 in Lind and Ramondo (2023), average equilibrium utility is given by:

$$\bar{U} = (1+d)\,\Gamma\left(\frac{\epsilon-1}{\epsilon}\right)\left[\sum_{n\in\mathcal{N}}B_n\left(\frac{W_n^b}{P_n^{\alpha}R_n^{1-\alpha}}\right)^{\epsilon}\right]^{\frac{1}{\epsilon}}.$$

Moreover, by Lemma O.6 (part 1) in Lind and Ramondo (2023), it can be shown that (supply-side) plant labor shares are given by:

$$\frac{l_{nsj}^S}{\bar{L}}\equiv\mathbb{P}\left\{\text{consumer } v \text{ chooses plant } sj \text{ in location } n\right\}=\frac{B_nB_s^{\frac{1}{1-\rho}}\left(\frac{w_{nsj}^{\frac{1}{1-\rho}}W_n^{b^{-\frac{\rho}{1-\rho}}}}{P_n^{\alpha}R_n^{1-\alpha}}\right)^{\epsilon}}{\sum_{i\in\mathcal{N}}B_i\left(\frac{W_i^b}{P_i^{\alpha}R_i^{1-\alpha}}\right)^{\epsilon}}. \qquad \text{(A.2)}$$

Summing equation (A.2) over all plants within the same location and sector and substituting in equation (12) gives:

$$\frac{L_{ns}}{\bar{L}} = \frac{B_n B_s^{\frac{1}{1-\rho}} \left(\frac{\left(W_{ns}^b\right)^{\frac{1}{1-\rho}} \left(W_n^b\right)^{-\frac{\rho}{1-\rho}}}{P_n^\alpha R_n^{1-\alpha}} \right)^\epsilon}{\sum_{i \in \mathcal{N}} B_i \left(\frac{W_i^b}{P_i^\alpha R_i^{1-\alpha}} \right)^\epsilon}.$$

Finally, dividing equation (A.2) by equation (10) and substituting in equation (12) gives:

$$\frac{l_{nsj}^S}{L_{ns}} = \frac{w_{nsj}^{\frac{\epsilon}{1-\rho}}}{\sum_{k \in \mathscr{E}_{ns}} w_{nsk}^{\frac{\epsilon}{1-\rho}}}.$$

B.1.2 Producer Problem

In this section, we derive the labor demand function (equation (15)) and the Free on Board Price (equation (17)) as part of the solution to the producer problem.

We begin by characterizing final goods producers' demand for intermediate goods. Let p_{nisj} denote the price which the final goods producer in location n and sector s pays to plant $\{i, s, j\}$ to receive one unit of its intermediate good. Moreover, define the price indices $\bar{p}_{nis} := \left(\sum_{j \in \mathscr{E}_{is}} p_{nisj}^{1-\omega} \right)^{\frac{1}{1-\omega}}$ and $\hat{p}_{ns} := \left(\sum_{i \in \mathcal{N}_s} \bar{p}_{nis}^{1-\eta} \right)^{\frac{1}{1-\eta}}$. Standard nested CES demand results imply that expenditure shares are given by:

$$\frac{q_{nisj} p_{nisj}}{q_{ns} \hat{p}_{ns}} = \left(\frac{p_{nisj}}{\bar{p}_{nis}} \right)^{1-\omega} \left(\frac{\bar{p}_{nis}}{\hat{p}_{ns}} \right)^{1-\eta}. \tag{A.3}$$

We next present the solution to the plant's problem. Our pricing assumption is that plants take the price indices $\bar{p}_{nis}$ and $\hat{p}_{ns}$ as given. Standard results then imply that plants set a constant net markup over marginal cost given by $\mu_{isj} = \frac{1}{\omega-1}$. Combining this with Cobb-Douglas production implies that, in equilibrium:

$$p_{nisj} = \frac{\omega}{\omega - 1} z_{isj}^{-1} w_{isj}^{\gamma_s} \left(p_{is}^m\right)^{1-\gamma_s} \tau_{ni}. \tag{A.4}$$

Plugging equation (A.4) into equation (A.3) and manipulating yields:

$$\frac{X_{isj}}{\sum_{k \in \mathscr{E}_{is}} X_{isk}} = \frac{z_{isj}^{\omega-1} w_{isj}^{\gamma_s(1-\omega)}}{\sum_{k \in \mathscr{E}_{is}} z_{isk}^{\omega-1} w_{isk}^{\gamma_s(1-\omega)}},$$

where X_{isj} denotes plant $\{i, s, j\}$'s revenue. Constant expenditure share on labor then implies that labor demand is given by:

$$\frac{l_{isj}^{D}}{L_{is}} = \frac{z_{isj}^{\omega-1} w_{isj}^{\gamma_s(1-\omega)-1}}{\sum_{k\in\mathscr{E}_{is}} z_{isk}^{\omega-1} w_{isk}^{\gamma_s(1-\omega)-1}}.$$

Finally, combining equation (A.4) with standard nested CES demand results gives:

$$\tilde{p}_{is} = \frac{\omega}{\omega-1}\left(\sum_{j\in\mathscr{E}_{is}}\left(\frac{w_{isj}^{\gamma_s}}{z_{isj}W_{is}^{\gamma_s}}\right)^{1-\omega}\right)^{\frac{1}{1-\omega}} W_{is}^{\gamma_s}\left(p_{is}^{m}\right)^{1-\gamma_s} \equiv \mathcal{MC}_{is} W_{is}^{\gamma_s}\left(p_{is}^{m}\right)^{1-\gamma_s}.$$

B.2 Proofs

B.2.1 Equilibrium Characterization and Uniqueness

In this section, we provide the proof of Proposition 1 and show the isomorphism between our model and an augmented version of Caliendo and Parro (2015).

Proof of Proposition 1

We divide the proof into two sub-claims.

Claim 1. Given firm networks and parameters, if $\zeta\left(1+\theta\left(\max_j\left|\mathcal{E}_j\right|-1\right)\right) < 1$, there exists a unique vector of plant-level knowledge which satisfies the system of equations governed by equation (8).

Proof. Define the function:

$$g_{isj}\left(\{k_{ntk}\}_{n\in\mathcal{N};tk\in\mathscr{E}_n}\right) := \left(\sum_{tk\in\mathscr{E}_n} k_{ntk}\right)^{\zeta(1-\theta)}\left(\Pi_{n\in\mathcal{E}_j}\left(\sum_{tk\in\mathscr{E}_n} k_{ntk}\right)^{\zeta\theta}\right).$$

It is straightforward to see that g_{isj} is positively homogeneous of degree $\zeta\left(1+\theta\left(\left|\mathcal{E}_j\right|-1\right)\right)$ and that $\frac{\partial g_{isj}}{\partial k_{ntk}} \geq 0$. Combining this with Euler's Homogeneous Function Theorem gives $\sum_{n\in\mathcal{N}}\sum_{tk\in\mathscr{E}_n}\left|\frac{\partial ln g_{isj}}{\partial ln k_{nk}}\right| \leq \zeta\left(1+\theta\left(\max_j\left|\mathcal{E}_j\right|-1\right)\right)$. Finally, note that $k_{isj} = \sum_{n\in\mathcal{N}}\sum_{tk\in\mathscr{E}_n}\mathbf{1}_{ntk=isj}\, g_{ntk}\left(\{k_{ntk}\}_{n\in\mathcal{N};tk\in\mathscr{E}_n}\right)$. Applying Remark 1 from Allen, Arkolakis, and Li (2022) completes the proof. □

Claim 2. For each location-sector pair is, given parameter values, plant-level knowledge, and fundamental productivity, there exist unique labor shares and relative wages which satisfy equations (13) and (15).

Proof. Define the function:

$$\mathcal{D}_{isj}\left(\{w_{isk}\}_{k\in\mathcal{E}_{is}}\right) := \frac{z_{isj}^{\omega-1} w_{isj}^{\gamma_s(1-\omega)-1}}{\sum_{k\in\mathcal{E}_{is}} z_{isk}^{\omega-1} w_{isk}^{\gamma_s(1-\omega)-1}} - \frac{w_{isj}^{\frac{\epsilon}{1-\rho}}}{\sum_{k\in\mathcal{E}_{ns}} w_{isk}^{\frac{\epsilon}{1-\rho}}}. \tag{A.5}$$

If equation (A.5) characterizes an excess demand system with the gross substitution property, equilibrium labor shares and relative wages are unique. This requires four properties: (i) $\mathcal{D}_{isj}$ is continuous, (ii) $\mathcal{D}_{isj}$ is homogeneous of degree zero, (iii) $\sum_{j\in\mathcal{E}_{is}} \mathcal{D}_{isj} = 0$, and (iv) $\mathcal{D}_{isj}$ exhibits gross substitution in wages. It is straighforward to see that properties (i) and (ii) hold. For property (iii), note that $\sum_{j\in\mathcal{E}_{is}} \frac{z_{isj}^{\omega-1} w_{isj}^{\gamma_s(1-\omega)-1}}{\sum_{k\in\mathcal{E}_{is}} z_{isk}^{\omega-1} w_{isk}^{\gamma_s(1-\omega)-1}} = 1$ and $\sum_{j\in\mathcal{E}_{is}} \frac{w_{isj}^{\frac{\epsilon}{1-\rho}}}{\sum_{k\in\mathcal{E}_{is}} w_{isk}^{\frac{\epsilon}{1-\rho}}} = 1$. Finally, for gross substitution, it is sufficient to show that $\frac{\partial \mathcal{D}_{isj}}{\partial w_{isk}} > 0 \; \forall k \in \mathcal{E}_{is} \setminus \{j\}$. To see that this holds, note that:

$$\frac{\partial}{\partial w_{isk}} \frac{z_{isj}^{\omega-1} w_{isj}^{\gamma_s(1-\omega)-1}}{\sum_{k\in\mathcal{E}_{is}} z_{isk}^{\omega-1} w_{isk}^{\gamma_s(1-\omega)-1}} = \underbrace{-\left(\gamma_s\left(1-\omega\right)-1\right)}_{>0} \underbrace{\frac{z_{isk}^{\omega-1} w_{isk}^{\gamma_s(1-\omega)-2} z_{isj}^{\omega-1} w_{isj}^{\gamma_s(1-\omega)-1}}{\left(\sum_{k\in\mathcal{E}_{is}} z_{isk}^{\omega-1} w_{isk}^{\gamma_s(1-\omega)-1}\right)^2}}_{>0}$$

and

$$\frac{\partial}{\partial w_{isk}} \frac{w_{isj}^{\frac{\epsilon}{1-\rho}}}{\sum_{k\in\mathcal{E}_{ns}} w_{isk}^{\frac{\epsilon}{1-\rho}}} = -\underbrace{\frac{\epsilon}{1-\rho}}_{<0} \underbrace{\frac{w_{isj}^{\frac{\epsilon}{1-\rho}} w_{isk}^{\frac{\epsilon}{1-\rho}-1}}{\left(\sum_{k\in\mathcal{E}_{ns}} w_{isk}^{\frac{\epsilon}{1-\rho}}\right)^2}}_{>0},$$

which implies that $\frac{\partial \mathcal{D}_{isj}}{\partial w_{isk}} > 0$. □

Isomorphism to Caliendo and Parro

We finally show that the across-location equilibrium conditions (equations (10) and (21)), which pin down $\{\boldsymbol{L}, \boldsymbol{W}\}$, are isomorphic to corresponding conditions in a version of

Caliendo and Parro (2015, CP) augmented with mobile labor across regions, idiosyncratic preferences over locations and sectors, a land market, and local agglomeration economies. Below we state the augmentations to CP, present the equilibrium conditions of the augmented CP model, and state the isomorphism between the equilibrium conditions of our model and the corresponding conditions of the augmented CP model. For ease of comparison, our notation closely follows CP.

Consider the following augmentations to CP: (i) workers are mobile across regions and sectors; (ii) utility for atomistic consumer ν is given by $u_{ns\nu} = b_{ns\nu} C_\nu^{1-\alpha} h_\nu^\alpha$, where n and s denote location and sector, respectively, with consumption goods as in CP and a land market as described in equation (9); (iii) idiosyncratic preferences are given by the distribution $\mathbb{P}\left(\bigcap_{n\in\mathcal{N}}\bigcap_{s\in\mathcal{S}_n}\{b_{ns}\leq t_{ns}\}\right) = exp\left\{-\sum_{n\in\mathcal{N}}\left(\sum_{s\in\mathcal{S}_n}(B_{ns})^{\frac{1}{1-\rho}} t_{ns}^{-\frac{\epsilon}{1-\rho}}\right)^{1-\rho}\right\}$; and (iv) location-sector specific productivity scale parameters are given by $\lambda_i^s = \tilde{\lambda}_i^s L_i^{\beta_s}$, where $L_i^{\beta_s}$ represents classical, local agglomeration economies and $\tilde{\lambda}_i^s$ is exogenous "fundamental productivity." We can state the equilibrium conditions (with s, t corresponding to sectors and i, n corresponding to regions) for this augmented model as:

$$c_i^s = \Upsilon_i^s \left(w_i^s\right)^{\gamma_i^s} \Pi_{t=1}^S \left(P_i^s\right)^{\gamma_i^{t,s}}, \tag{A.6}$$

$$P_n^s = A_s \left(\sum_{i=1}^N \lambda_i^s \left(c_i^s \kappa_{ni}^s\right)^{-\theta_j}\right)^{-\frac{1}{\theta_j}}, \tag{A.7}$$

$$\pi_{ni}^s = \frac{\lambda_i^s \left(c_i^s \kappa_{ni}^s\right)^{-\theta_j}}{\sum_{h=1}^N \lambda_h^s \left(c_h^s \kappa_{nh}^s\right)^{-\theta_j}}, \tag{A.8}$$

$$X_n^s = \sum_{t=1}^S \gamma_n^{j,t} \sum_{i=1}^N X_i^k \left(\frac{\pi_{in}^t}{1+\tau_{in}^k}\right) + \alpha_n^j I_n, \tag{A.9}$$

$$\frac{L_{ns}}{\bar{L}} = \frac{B_{ns}^{\frac{1}{1-\rho}} \left(\frac{\left(w_n^s\right)^{\frac{1}{1-\rho}} \left(W_n^b\right)^{-\frac{\rho}{1-\rho}}}{P_n^\alpha R_n^{1-\alpha}}\right)^\epsilon}{\sum_{t=1}^S \sum_{i=1}^N B_{it}^{\frac{1}{1-\rho}} \left(\frac{\left(w_n^s\right)^{\frac{1}{1-\rho}} \left(W_n^b\right)^{-\frac{\rho}{1-\rho}}}{P_n^\alpha R_n^{1-\alpha}}\right)}. \tag{A.10}$$

Equations (A.6)-(A.9) correspond to equations (2), (4), (6), and (7) in CP, respectively;

equation (A.10) is an additional equilibrium condition resulting from the additional forces in the augmented model.[41] Given parametric restrictions, equations (A.6) and (A.7) are isomorphic to equations (17) and (19), equation (A.8) is isomorphic to equation (20), equation (A.9) is isomorphic to equation (21), and equation (A.10) is isomorphic to equation (10).

B.2.2 Inversion Uniqueness

In this section, we provide the proof of Proposition 2. We divide the proof into four steps.

Step 1: Within-County-Sector Inversion We first provide an analytical inversion for within-county relative wages and productivity. Given plant-level employment, equation (13) can be manipulated to show that:

$$\frac{w_{isj}}{W_{is}} = l_{isj}^{\frac{1-\rho}{\epsilon}} \frac{\sum_{k \in \mathscr{E}_{is}} l_{isk}}{\sum_{k \in \mathscr{E}_{is}} l_{isk}^{1+\frac{1-\rho}{\epsilon}}} \tag{A.11}$$

Thus, we can invert plant-level employment to recover plant-level relative wages, which we denote by $\tilde{w}_{isj} := \frac{w_{isj}}{W_{is}}$. Define $z_{isj}^* := l_{isj}^{\frac{1}{\omega-1}} \tilde{w}_{isj}^{\frac{1-\gamma_s(1-\omega)}{\omega-1}}$. Given relative wages and employment, it can be shown that for any $\kappa_{is} \in \mathbb{R}_{++}$, the set of productivity parameters given by $\left\{\kappa_{is} z_{isj}^*\right\}_{j \in \mathscr{E}_{is}}$ ensures that equation (15) holds for observed labor shares. We refer to z_{isj}^* as "relative productivity"; we pin down κ_{is} later in Step 3.

Step 2: Sectoral Amenity Scale Parameters and Wages

Claim 3. Given plant-level employment and wages, parameters, and recovered markups, there exist unique (up to a normalization) sets of sector-specific amenity scale prameters and county-sector wages that rationalize the observed data as an equilibrium of the model.

Proof. Summing equation (21) within each sector, plugging in equation (10), and manipulating yields the system $O_s = \sum_{t \in \mathcal{S}} O_t \phi_{ts}$, for $s_{ns} := \frac{L_{ns}^{\frac{\epsilon+1-\rho}{\epsilon}} \left(\tilde{W}_{ns}^b\right)^{-1} B_s^{-\epsilon}}{\sum_{t \in \mathcal{S}_n} L_{nt}^{\frac{\epsilon+1-\rho}{\epsilon}} \left(\tilde{W}_{nt}^b\right)^{-1} \tilde{B}_t^{-\epsilon}}$, $W_{ns}^b := \left(\sum_{j \in \mathscr{E}_{ns}} \tilde{w}_{nsj}^{\frac{\epsilon}{1-\rho}}\right)^{\frac{1-\rho}{\epsilon}}$, $O_s := \frac{\sum_{n \in \mathcal{N}_s} s_{ns} W_n L_n}{\sum_{t \in \mathcal{S}} \sum_{n \in \mathcal{N}_t} s_{nt} W_n L_n}$, and $\phi_{ts} := \frac{\gamma_s}{\gamma_t} \frac{(\omega-1)}{\omega} \left(\kappa_s \left(\gamma_t + \frac{1}{(\omega-1)}\right) + \delta_{ts}\left(1-\gamma_t\right)\right)$.

[41] The above equations correspond to all of the equilibrium conditions in CP, except equation (9), which accounts for trade imbalances. In our model, trade imbalances are fully captured in equation (21), so we omit stating the additional equilibrium condition for brevity. Also, while our model does not feature tarrifs, plant-level markups enter equilibrium conditions in the same way.

Linearity implies that the system has at most one solution with $\sum_{s\in\mathcal{S}} O_s = 1$. All that remains is to show is therefore that there is a unique inversion from $\{O_s\}_{s\in\mathcal{S}}$ to $\{B_s\}_{s\in\mathcal{S}}$. To show this, define:

$$\mathscr{O}_s\left(\{B_t\}_{t\in\mathcal{S}}\right) := \frac{\sum_{n\in\mathcal{N}_s} \frac{L_{ns}^{\frac{\epsilon+1-\rho}{\epsilon}}\left(\tilde{W}_{ns}^b\right)^{-1}B_s^{-\epsilon}}{\sum_{t\in\mathcal{S}_n} L_{nt}^{\frac{\epsilon+1-\rho}{\epsilon}}\left(\tilde{W}_{nt}^b\right)^{-1}\tilde{B}_t^{-\epsilon}} w_n L_n}{\sum_{n\in\mathcal{N}} w_n L_n}, \tag{A.12}$$

$$\mathcal{D}_s\left(\{B_t\}_{t\in\mathcal{S}}\right) := \mathscr{O}_s\left(\{B_t\}_{t\in\mathcal{S}}\right) - O_s. \tag{A.13}$$

If the system of equations defined by equation (A.13) represents an excess demand system with the gross substitution property, then it has a unique (up to scale) solution. This requires (i) $\mathcal{D}_s$ is continuous, (ii) $\mathcal{D}_s$ is homogeneous of degree zero, (iii) $\sum_s \mathcal{D}_s = 0$, and (iv) $\mathcal{D}_s$ exhibits gross substitution in amenity scale parameters. Properties (i) and (ii) follow immediately from inspection of equation (A.12). For property (iii), clearly $\sum_{s\in\mathcal{S}} O_s = 1$; we also have:

$$\begin{aligned}\sum_{s\in\mathcal{S}} \mathscr{O}_s\left(\{B_t\}_{t\in\mathcal{S}}\right) &= \sum_{s\in\mathcal{S}} \frac{\sum_{n\in\mathcal{N}_s} \frac{L_{ns}^{\frac{\epsilon+1-\rho}{\epsilon}}\left(\tilde{W}_{ns}^b\right)^{-1}B_s^{-\epsilon}}{\sum_{t\in\mathcal{S}_n} L_{nt}^{\frac{\epsilon+1-\rho}{\epsilon}}\left(\tilde{W}_{nt}^b\right)^{-1}\tilde{B}_t^{-\epsilon}} w_n L_n}{\sum_{n\in\mathcal{N}} w_n L_n} \\ &= \frac{\sum_{n\in\mathcal{N}} \frac{\sum_{s\in\mathcal{S}_n} L_{ns}^{\frac{\epsilon+1-\rho}{\epsilon}}\left(\tilde{W}_{ns}^b\right)^{-1}B_s^{-\epsilon}}{\sum_{t\in\mathcal{S}_n} L_{nt}^{\frac{\epsilon+1-\rho}{\epsilon}}\left(\tilde{W}_{nt}^b\right)^{-1}\tilde{B}_t^{-\epsilon}} w_n L_n}{\sum_{n\in\mathcal{N}} w_n L_n} = \frac{\sum_{n\in\mathcal{N}} w_n L_n}{\sum_{n\in\mathcal{N}} w_n L_n} = 1.\end{aligned}$$

For property (iv), define $O_{ns} := \left(\frac{L_{ns}^{\frac{\epsilon+1-\rho}{\epsilon}}\left(\tilde{W}_{ns}^b\right)^{-1}B_s^{-\epsilon}}{\sum_{t\in\mathcal{S}_n} L_{nt}^{\frac{\epsilon+1-\rho}{\epsilon}}\left(\tilde{W}_{nt}^b\right)^{-1}\tilde{B}_t^{-\epsilon}}\right)$. It is sufficient to show that $\forall n \in \mathcal{N}$, $\frac{\partial lnO_{ns}}{\partial lnB_{s'}} \geq 0$ and $\exists n \in \mathcal{N} : \frac{\partial lnO_{ns}}{\partial lnB_{s'}} > 0 \; \forall s' \in \mathcal{S} \backslash \{s\}$. We have that $\frac{\partial lnO_{ns}}{\partial lnB_{s'}} = \epsilon s_{ns} s_{ns'} \geq 0$. Since all sectors overlap in at least one location in the economy, the inequality is strict for at least one location. Finally, equation (10) pins down $\{W_{ns}\}_{n\in\mathcal{N}:s\in\mathcal{S}_n}$. □

Step 3: Plant-Level Productivities

Claim 4. Given observed data, parameters, and recovered fundamentals, there exists a unique set of fundamental productivities (up to a normalization for each sector $s \in \mathcal{S}$) that is consistent with an equilibrium of the model.

Proof. Consider sector s, and define $y_{ns} := \frac{\omega}{\omega-1}\frac{W_{ns}L_{ns}}{\gamma_s}$ and $y_{nst} := W_{nt}L_{nt}\left((1+d)\,\kappa_s + \delta_{ts}\frac{1-\gamma_t}{\gamma_t}\right)$. Moreover, define the function:

$$D_{ns}\left(\{\tilde{p}_{is}\}_{i\in\mathcal{N}_s}\right) = y_{ns} - \sum_{i\in\mathcal{N}}\sum_{t\in\mathcal{S}_i}\left(\frac{\tau_{in}\tilde{p}_{ns}}{p_{is}}\right)^{1-\eta} y_{ist}. \tag{A.14}$$

If D_{ns} characterizes an excess demand system with the gross substitution property, there is a unique (up-to scale) set of Free on Board prices satisfying equation (A.14). This requires (i) D_{ns} is continuous, (ii) D_{ns} is homogeneous of degree zero, (iii) $\sum_{n\in\mathcal{N}_s} D_{ns} = 0$, and (iv) D_{ns} exhibits gross substitution in Free on Board prices. Properties (i) and (ii) follow immediately by inspection. Property (iii) requires that $\sum_{n\in\mathcal{N}_s} y_{ns} = \sum_{i\in\mathcal{N}}\sum_{t\in\mathcal{S}_i} y_{ist}$, which follows from the proof of Claim 3. For property (iv), it is sufficient to show that $\frac{\partial}{\partial\tilde{p}_{n's}}\sum_{i\in\mathcal{N}}\sum_{t\in\mathcal{S}_i}\left(\frac{\tau_{in}\tilde{p}_{ns}}{p_{is}}\right)^{1-\eta} y_{ist}$ has a constant sign for $n' \neq n$. We have:

$$\frac{\partial}{\partial\tilde{p}_{n's}}\sum_{i\in\mathcal{N}}\sum_{t\in\mathcal{S}_i}\left(\frac{\tau_{in}\tilde{p}_{ns}}{p_{is}}\right)^{1-\eta} y_{ist} = (1-\eta)\sum_{i\in\mathcal{N}}\sum_{t\in\mathcal{S}_i}\underbrace{\left(\frac{\tau_{in}\tilde{p}_{ns}}{p_{is}}\right)^{-\eta}\tau_{in}\frac{\partial\frac{\tilde{p}_{ns}}{p_{is}}}{\partial\tilde{p}_{n's}} y_{ist}}_{<0}.$$

Thus, all cross-partials have the same sign (which depends on the value of η), which is sufficient for gross substitution.

It remains to show that within-industry fundamental productivities are unique up to scale. Let $\left\{\tilde{p}^*_{is}\right\}_{i\in\mathcal{N}:s\in\mathcal{S}_i}$ be an (arbitrarily scaled) set of Free On Board prices that solves equation (A.14). Final goods prices in sector s and location n are then given $p^*_{ns} = \left(\sum_{i\in N}\tau_{ni}^{1-\eta}\left(\tilde{p}^*_{is}\right)^{1-\eta}\right)^{\frac{1}{1-\eta}}$, which in turn pins down materials prices $\left\{p^{m*}_{is}\right\}_{i\in\mathcal{N}_s}$. It then follows that $\mathcal{MC}^*_{is} = \frac{\tilde{p}^*_{is}}{\left(p^{m*}_{is}\right)^{1-\gamma_{is}}W_{is}^{\gamma_s}}$. Suppose we re-scale sector s's prices by k_s. Then, $\tilde{p}'_{is} = k_s\tilde{p}^*_{is}$, $p'_{ns} = k_s p^*_{ns}$, and $p^{m'}_{is} = p^{m*}_{is}\Pi_{t\in\mathcal{S}}k_t^{\delta_{ts}}$, so $\mathcal{MC}'_{is} = \mathcal{MC}^*_{is}\left(\Pi_{t\in\mathcal{S}}k_t^{\delta_{ts}}\right)^{\gamma_s-1}$. Note that $\left(\Pi_{t\in\mathcal{S}}k_t^{\delta_{ts}}\right)^{\gamma_s-1}$ is constant within-sector; therefore, $\mathcal{MC}_{is}$ is unique up to scale within each sector.

We finally show that the recovered values of $\mathcal{MC}_{is}$ pin down the values of z_{isj}. Manipulating the definition of $\mathcal{MC}_{is}$ and z_{isj} gives:

$$z_{isj} = z^*_{isj}\frac{\mathcal{MC}^*_{is}}{\frac{1}{\omega-1}\left(\sum_{j\in\mathscr{E}_{is}}\left(z^*_{isj}\right)^{\omega-1}\tilde{w}_{isj}^{\gamma_s(1-\omega)}\right)^{\frac{1}{1-\omega}}}. \tag{A.15}$$

Note that equation (A.15) pins down κ_{is}, which was introduced in Step 1. Finally, combining Claim 1 with observed employment gives the unique decomposition of z_{isj} into $\tilde{z}_{isj}$, L_i^{β}, and k_{isj}. □

Step 4: County-Level Amenity Scale Parameters

Claim 5. Given observed data, parameters, and recovered fundamentals, there exists a unique (up to a normalization) set of county amenity scale parameters that rationalizes the observed data as an equilibrium of the model.

Proof. The proof is identical to the proof of Proposition 3 in Redding (2016). □

C Estimation Appendix

C.1 Model Economy

To construct our model economy, we combine public Census data with information from confidential Census data (1982 CMF) to match key characteristics of MC firms, MC plants, and MC plant-level networks, as described below.[42]

First, we use public Census data to measure county-level exogenous fundamentals. We gather data on land areas and geographical coordinates from the U.S. Gazetteer and TIGER/Line Shapefiles. We use these data to compute (haversine) geographical distances, which in our model pin down trade costs, $\tau_{ni} = dist_{ni}^{\psi}$.

Second, for each county, we obtain data on manufacturing employment and wages as well as the number of plants in each sector from the 1982 County Business Patterns (CBP). We augment the CBP with imputed data from Eckert et al. (2021) to fill in missing values for county-sector level employment. Overall, our model economy has 312,633 plants. We consider 14 sectors based on 2-digit SIC codes: Food and Tobacco (SIC 20-21), Textiles, Apparel, and Leather (SIC 22-23, 31), Lumber and Furniture (SIC 24-25), Paper and Printing (SIC 26-27), Chemicals (SIC 28), Petroleum and Coal (SIC 29), Rubber and Plastics (SIC 30), Minerals (SIC 32), Primary Metals (SIC 33), Fabricated Metals (SIC 34), Industrial Machinery (SIC 35), Electronics (SIC 36), Transportation Equipment (SIC 37), and Miscellaneous Manufacturing (SIC 38-39).

[42] Computational constraints make it necessary to run the estimation outside the Census Bureau.

Third, for each county-sector cell, we use the (augmented) CBP in combination with confidential Census data to simulate plant-level employment and whether a plant is SC or MC. While this information is not directly observable in public data, the CBP provides for each county-sector cell a breakdown of the number of plants by employment size, grouped into 12 granular size buckets.[43] Using confidential Census data, we compute for each size bucket: i) the share of MC plants, ii) mean SC plant employment, and iii) mean MC plant employment. For each size bucket, we then fit a (truncated) Pareto distribution to plant-level employment separately for SC and MC plants (24 Pareto distributions). We use these distributions together with the MC plant share to simulate, for each county-sector cell, plant-level employment and, given employment, whether a plant is SC or MC. The end result is a rich plant-level data set where, as in confidential Census data, we know for each plant its county, sector, employment, plant ID, and SC/MC affiliation.

Finally, using confidential Census data, we generate the (non-parametric) joint distribution of firm size and spatial dispersion in firms' plant-level networks. We assign MC plants to firms matching this joint distribution. Firm size is the number of plants; spatial dispersion corresponds to the fraction of MC firms that have all of their plants in the same state, same Census region, two Census regions, three Census regions, or all four Census regions. (Using a more granular partitioning does not materially affect the estimates.) Matching the joint distribution of firm size and spatial dispersion in firms' plant-level networks accounts for observed patterns in the data, where larger MC firms tend to be much more geographically dispersed than smaller MC firms.

C.2 Parameter Estimation

Let $\boldsymbol{\Theta} := \{\zeta, \theta, \eta, \epsilon, \omega, \rho\}$ denote the set of estimated parameters. We employ a standard indirect inference objective function, $Q(\boldsymbol{\Theta}) = (\boldsymbol{m} - \boldsymbol{m}(\Theta))' \, \Xi \, (\boldsymbol{m} - \boldsymbol{m}(\Theta))$ and $\hat{\boldsymbol{\Theta}} = \arg\min_{\boldsymbol{\Theta}} Q(\boldsymbol{\Theta})$ for data moments $\boldsymbol{m}$, corresponding model moments $\boldsymbol{m}(\Theta)$, and positive definite weighting matrix Ξ.[44] The process to compute $\boldsymbol{m}(\boldsymbol{\Theta})$ consists of three steps.

Step 1. Pre-Shock Equilibrium and Recovery of Fundamentals. Given parameter values and

[43] The size buckets are: 1-4 employees, 5-9 employees, 10-19 employees, 20-49 employees, 50-99 employees, 100-249 employees, 250-499 employees, 500-999 employees, 1000-1499 employees, 1500-2499 employees, 2500-4999 employees, and 5000+ employees.

[44] Since, at the estimated parameter values, our model moments exactly match the data moments, the choice of weighting matrix is irrelevant.

observed fundamentals, we recover the unobserved fundamentals, plant-level knowledge, and plant-level wages that rationalize the data as an equilibrium of the model. The inversion process follows the step-by-step procedure described in Online Appendix B.2.

Definition 1. *Given parameter values and fundamentals, the pre-shock equilibrium is given by* $\{\boldsymbol{k}(\Theta), \boldsymbol{l}, \boldsymbol{w}(\Theta), \boldsymbol{L}, \boldsymbol{W}(\Theta)\}$.

We denote the dependence of endogenous objects, $\{\boldsymbol{k}, \boldsymbol{w}, \boldsymbol{W}\}$, on estimated parameters as a reminder that the unobserved endogenous objects are functions of the parameters to be estimated. Proposition 2 shows that, given parameter values, the vector of endogenous objects $\{\boldsymbol{W}, \boldsymbol{w}, \boldsymbol{k}\}$ that is consistent with an equilibrium of the model is unique.

Step 2. Post-Shock Equilibrium. The second step involves re-computing the endogenous objects after perturbing the pre-shock equilibrium by adding an MDP to the winner county. We use characteristics of the actual MDPs to discipline corresponding fundamentals in the model. Specifically, we set the MDPs' fundamental productivities to match the actual MDPs' county-level employment shares immediately after entry and compute MDP parent firm knowledge as described in the model.[45] Holding all other fundamentals constant, we compute the new equilibrium of the model, which we refer to as "post-shock equilibrium." We perform this step for each of the 47 MDP cases (i.e., we use the actual winner and runner-up counties from the data).

Definition 2. *Given parameter values and fundamentals, the post-shock equilibrium is given by* $\{\boldsymbol{k}'_c(\Theta), \boldsymbol{l}'_c(\Theta), \boldsymbol{w}'_c(\Theta), \boldsymbol{L}'_c(\Theta), \boldsymbol{W}'_c(\Theta)\}$.

Step 3. Difference-in-Differences Regressions. To generate model moments that correspond to the data moments, we stack the pre-shock equilibrium together with the 47 post-shock equilibria and estimate local and global spillover difference-in-differences regressions akin to those in our reduced-form analysis. As in our reduced form, the local spillover sample includes all plants in the winner and corresponding runner-up counties (except for the MDPs and any plants owned by the MDPs' parent firms). We measure outcome variables y_{iskct} (in logs) at the plant-case-time level, where *isk* indicates plants, as specified in our

[45] In our estimated model, larger MDPs generate stronger local productivity spillovers, which is consistent with the reduced-form evidence in Table A.4.

model, c indicates cases, and time t corresponds to pre- and post-MDP entry. We estimate the following specification:

$$y_{iskct} = \xi_{isk} + \xi_{cst} + \beta_1(Winner_i \times Post_{ct}) + \varepsilon_{iskct},$$

where the dependent variable is either plant-level productivity, plant-level employment, or plant-level wages. The specification is exactly the same as in our reduced form, except that we interact the case fixed effects with the industry $\times$ time fixed effects to account for the stacked nature of the MDP experiments.[46] The estimated parameter $\hat{\beta}_1$ provides us with three model moments—local semi-elasticities of plant-level employment, wages, and productivity to the MDP openings. To generate the remaining three model moments—global semi-elasticities of plant-level employment, wages, and productivity to the MDP openings—we estimate the same specification as in our reduced-form global spillover analysis, except that we again interact the case fixed effects with the industry $\times$ county $\times$ year fixed effects to account for the stacked nature of the experiments.

Finally, to estimate standard errors for the structural parameters, we employ a GMM standard error estimator for just-identified models. This takes the form of $\hat{\Omega} = \left(G'\hat{S}^{-1}G\right)^{-1}$, where G denotes the Jacobian of moments to parameters, and $\hat{S}$ denotes the sample moment covariance matrix. We calculate G numerically. To calculate $\hat{S}$, we estimate all six central reduced-form regressions (local and global TFP, wage, and employment responses to the MDP openings) jointly as one regression by concatenating the regression samples, interacting all right-hand side variables with "local" or "global" indicators $\times$ corresponding outcome variable (TFP, wages, employment) indicators, and otherwise mimicking the reduced-form regression specifications.[47] The standard errors for the structural parameters are the square roots of the diagonal elements of $\hat{\Omega}$.

References

Allen, Treb, Costas Arkolakis, and Xiangliang Li, 2022, On the Equilibrium Properties of Network Models with Heterogeneous Agents, mimeo, Dartmouth College.

[46] The coefficient on $Post_{ct}$ is absorbed by the interacted fixed effects.

[47] This method of jointly estimating all six regressions as one regression generates the same coefficient estimates as running the six regressions independently but allows us to compute the off-diagonal elements of $\hat{S}$.

Borusyak, Kirill, Xavier Jaravel, and Jann Spiess, 2023, Revisiting Event Study Designs: Robust and Efficient Estimation, forthcoming, Review of Economic Studies.

Caliendo, Lorenzo, and Fernando Parro, 2015, Estimates of the Trade and Welfare Effects of NAFTA, Review of Economic Studies 82, 1–44.

Eckert, Fabian, Teresa Fort, Peter Schott, and Natalie Yang, 2021, Imputing Missing Values in the US Census Bureau's County Business Patterns, NBER Working Paper 26632.

Lind, Nelson, and Natalia Ramondo, 2023, Trade with Correlation, American Economic Review 103, 317–353.

Redding, Stephen, 2016, Goods Trade, Factor Mobility and Welfare, Journal of International Economics 101, 148–167.