Online Appendix

Local Productivity Spillovers

By Nathaniel Baum-Snow, Nicolas Gendron-Carrier, and Ronni Pavan*

November 9, 2023

* Baum-Snow: Rotman School of Management, University of Toronto, 105 St. George Street, Toronto, ON, Canada, M5S3E6. e-mail: Nate.Baum.Snow@rotman.utoronto.ca. Gendron-Carrier: Department of Economics, McGill University, 855 Sherbrooke St. West, Montreal, QC, Canada, H3A2T7. e-mail: nicolas.gendron-carrier@mcgill.ca. Pavan: Department of Economics, University of Rochester, 280 Hutchinson Road, Rochester, NY, USA, 14627. e-mail: rpavan@ur.rochester.edu.

Appendix A. The Imperfect Competition Case

This section develops structural equations that describe relationships between firm revenue or variable factor demand and peer group composition. Using these equations, we provide structural interpretations of empirical estimates. We study an environment in which the variable input share and output demand elasticity are industry-specific.

A1. Setup

With market power, each firm charges a markup over marginal cost that depends on the elasticity of demand it faces for its product. To model this phenomenon, we begin with an adapted version of the environment considered in De Loecker (2011). In this environment, consumers have CES preferences across firm-specific varieties within 2-digit industries. This yields industry-specific demand elasticities for each variety that are fixed over time. In particular, the demand faced by firm i can be written as

$$q_{i,b,k,t} = X_{k,t} p_{i,b,k,t}^{\eta_k} e^{\zeta_{i,b,k,t}}.$$

In this equation, one way of interpreting the industry-time effect $X_{k,t}$ is as capturing the following combination of industry-time specific demand shocks and an average price across varieties in industry k at time t:

$$X_{k,t} = \frac{Q_{k,t}}{P_{k,t}^{\eta_k}}$$

Alternatively, we can think of $X_{k,t}$ as representing a more reduced form demand shifter that is common to all varieties in industry k at time t. Either way, η_k is the demand elasticity faced by each firm in industry k for its product and $\zeta_{i,b,k,t}$ is an i.i.d demand shock that is uncorrelated with TFP shocks.

Profit maximization yields the following expression for the firm-year-industry specific price:

$$\text{(A1)} \quad \begin{aligned} \ln p_{i,b,k,t} = & -\frac{1}{D_k} \ln A_{i,b,k,t} + \frac{\theta_k}{D_k} \ln w_{B(b),k,t} - \frac{\theta_k}{D_k} \ln[\frac{1+\eta_k}{\eta_k}\theta_k] \\ & + \frac{1-\theta_k}{D_k}[\ln X_{k,t} + \zeta_{i,b,k,t}]. \end{aligned}$$

The denominator $D_k = -\eta_k(1-\theta_k) + \theta_k > 0$. As η_k approaches negative infinity, $\ln p_{i,b,k,t}$ goes to a constant by construction and firms have no market power. Otherwise, positive productivity shocks depress output prices. Associated negative shocks to marginal costs lead firms to increase output, moving further down marginal revenue and demand functions. That is, the more market power firms have, the greater the pass-through of positive productivity shocks to price discounts. Similarly, positive wage shocks and positive demand shocks get passed through to increased variety prices in this environment.

By definition, $\ln R_{i,b,k,t} = \ln p_{i,b,k,t} + \ln q_{i,b,k,t} = (1+\eta_k)\ln p_{i,b,k,t} + \ln X_{k,t} + \zeta_{i,b,k,t}$. Insertion of equation (A1) into this condition delivers the following general expression for revenue, which matches equation (5) in the main text. This expression also holds under perfect competition, when $\eta_k = -\infty$.

$$\begin{aligned}\ln R_{i,b,k,t} =& \frac{1+\eta_k}{\eta_k(1-\theta_k)-\theta_k}\ln A_{i,b,k,t} - \frac{\theta_k(1+\eta_k)}{\eta_k(1-\theta_k)-\theta_k}\ln w_{B(b),k,t} \\ &- \frac{\theta_k(1+\eta_k)}{D_k}\ln[\frac{1+\eta_k}{\eta_k}\theta_k] + \frac{1}{D_k}[\ln X_{k,t} + \zeta_{i,b,k,t}]\end{aligned} \tag{A2}$$

If the firm is a price taker, this expression matches equation (2) in the main text with no change in price by l'Hôpital's Rule. As demand for the firm's product becomes less elastic, a given change in revenue must be driven by a larger TFP shock because the firm is more constrained in its optimal increase in quantity. For example, with $\theta_k = 0.7$ and $\eta_k = -2$, a 10 percent positive observed revenue change would reflect a 13 percent increase in TFP. However, with $\eta_k = -10$ instead, the associated TFP increase needed to achieve the same change in revenue is only 4 percent. Under perfect competition, this required TFP increase is further reduced to 3.3 percent.

A2. Derivation of an Estimating Equation

As seen in equation (A2), the pass-through of TFP shocks into revenue depends both on the strength of industry-specific market power and the importance of endogenous variable factor adjustments in response to TFP shocks. Within heterogeneous peer groups, there are thus variable revenue responses to the same TFP shock, making peer effects as described by a revenue based estimation equation heterogeneous within peer groups. This heterogeneous response mixes the TFP spillover parameter γ^A with market power and variable factor share parameters η_k and θ_k. In equation (4) in the main text, the structural interpretation of the firm fixed effect is determined jointly by the firm-specific fixed effect term and the spillover term.

To see this mathematically, begin with equation (A2) and set the firm fixed effect α_i^R to equal

$-\frac{1+\eta_{k(i)}}{D_{k(i)}}\alpha_i^A$. Remaining firm-specific terms in equation (4) then have the structural interpretation

$$\gamma^R \sum_{j\in M_{b,t},\neq i}[\omega_{ij}(M_{b,t})\alpha_j^R] + \varepsilon_{i,b,k,t}^R = \frac{(1+\eta_{k(i)})}{D_{k(i)}}\gamma^A \sum_{j\in M_{b,t},\neq i}[\omega_{ij}(M_{b,t})\alpha_j^R \frac{D_{k(j)}}{1+\eta_{k(j)}}] - \frac{(1+\eta_{k(i)})}{D_{k(i)}}\varepsilon_{i,b,k,t}^A + \frac{\zeta_{i,b,k,t}}{D_{k(i)}}.$$

From this equation, it is clear that if firm i is in the same industry as all its peers, revenue spillovers γ^R directly measure TFP spillovers γ^A. However, if they are in different industries, the estimated spillover in the revenue equation γ^R mixes information about peer group composition and variable markups.

Our approach for recovering structural TFP spillovers is to adjust the dependent variable to homogenize treatment effects in estimation equations with the same form as equation (4). In particular, dividing both sides of equation (A2) by $-\frac{1+\eta_k}{D_k}$ yields the adjusted revenue measure

$$\ln \tilde{R}_{i,b,k,t} \equiv -\frac{D_k}{1+\eta_k}\ln R_{i,b,k,t} \tag{A3}$$

for use as an outcome. Substituting for $\ln A_{i,b,k,t}$ using equation (3) in the main text, we have the following alternative structural equation for adjusted revenue, in which the spillover parameter equals the TFP spillover parameter γ^A:

$$\ln \tilde{R}_{i,b,k,t} = \alpha_i^A + \tilde{\phi}_{B(b),k,t} + \gamma^A\left[\sum_{j\in M_{b,t},\neq i}\omega_{ij}(M_{b,t})\alpha_j^A\right] + \tilde{\varepsilon}_{i,b,k,t}. \tag{A4}$$

Because using adjusted revenue $\ln \tilde{R}_{i,b,k,t}$ as the dependent variable isolates firm fixed effects as the permanent firm-specific component of TFP α_i^A, the TFP spillover parameter γ^A can be directly estimated as the peer effect parameter.

The new structural interpretation of the fixed effects in equation (A4) is

$$\tilde{\phi}_{B(b),k,t} = \phi_{B(b),k,t}^A - \theta_k \ln w_{B(b),t} - \theta_k \ln\frac{\eta_k}{1+\eta_k} + \theta_k\ln\theta_k - \frac{1}{1+\eta_k}\ln X_{k,t}$$

and the error term in equation (A4) is

$$\tilde{\varepsilon}_{i,b,k,t}^R = \varepsilon_{i,b,k,t}^A - \frac{\zeta_{i,b,k,t}}{1+\eta_k}.$$

As in the perfect competition case, the fixed effects control for location fundamentals, input costs, and industry-time specific demand conditions.

A3. Measuring Factor Shares, Markups, and TFP

Our robustness analysis that explicitly accounts for firm-specific price endogeneity requires measures of variable factor shares θ_k and demand elasticities η_k for implementation, as described in equation (A3). We calculate these objects using revenue and payments to variable and fixed inputs as observed in the data.

Using the firm level cost minimization condition, De Loecker and Eeckhout (2018) show that the firm level markup can be calculated as $\theta_k \frac{R_{i,b,k,t}}{(wL)_{i,b,k,t}}$. This relationship can be verified as being identical for all firms in industry k in the context of the more restrictive model laid out above. In particular, we have an industry level markup which is equal to $\frac{\eta_k}{1+\eta_k}$ by profit maximization.

In the data, we observe firm level revenue $R_{i,b,k,t}$ and annual payments to labor and materials. We infer payments to capital as rental and repair costs plus the book value of capital (net of amortization) times a discount rate plus industry-specific depreciation rate. We set the discount rate to be the prime business rate plus 0.04 minus the inflation rate. The prime business rate comes from the Bank of Canada (Bank of Canada, 2019). The inflation rate is calculated using the consumer price index information from Statistics Canada (Statistics Canada, 2019*b*). We infer payments to real estate as building maintenance costs plus property taxes plus rent plus the value of buildings and land (net of amortization) times a mortgage rate plus depreciation rate minus a capital gains rate. We manually collected information on commercial property tax rates for Montreal, Toronto, and Vancouver (see replication material for details). The mortgage rate is the prime rate plus 0.02. The depreciation rate is non-zero for structures only and is calculated for each 2-digit industry. We use information on flows and stocks of fixed non-residential capital by industry and type of asset to calculate depreciation rates. This information comes from Statistics Canada (Statistics Canada, 2019*a*). The capital gains rate uses the CMA level Teranet-National Bank residential home price index (Teranet-National Bank, 2019).

Using this information, we calculate the output elasticity with respect to variable factors $\theta_{k,t}$ and the markup $\frac{\eta_{k,t}}{1+\eta_{k,t}}$ at the 2-digit industry-year level. We calculate the output elasticity with respect to factor f, $\theta^f_{k,t}$, by aggregating payments to factors across all firms in each 2-digit industry-year bin, where the variable factor share $\theta_{k,t}$ is calculated as $\theta^{materials}_{k,t} + \theta^{labor}_{k,t}$. With $\theta_{k,t}$ in hand, we calculate the industry-year specific markup as

$$\frac{\eta_{k,t}}{1+\eta_{k,t}} = \theta_{k,t} \frac{\sum_i R_{i,k,t}}{\sum_i (wL)_{i,k,t}}.$$

Using this equation, we solve out for demand elasticities $\eta_{k,t}$ and average across years to recover calibrations of η_k. Our calibrations of θ_k are also averages of $\theta_{k,t}$ across years in our data.[1]

[1] We also experimented with using firm-specific markups but found them to be too noisy to be of use in estimation.

Appendix B. Estimation Details

This appendix derives the updating rules used for α_i in estimation.

B1. Case With One Peer Effect Term

We have the following generalized estimation equation which follows from equation (7) in the main text:

$$y_{i,k,b,t} = \alpha_i + \bar{\alpha} + \gamma\bar{\alpha}W_{b,t}^{-i} + \phi_{B(b),k(i),t} + \gamma \sum_{j \in M_{b,t} \setminus \{i\}} \omega_{ij}(M_{b,t})\alpha_j + \varepsilon_{i,k,b,t},$$

where $W_{b,t}^{-i} = \sum_{j \in M_{b,t} \setminus \{i\}} \omega_{ij}(M_{b,t})$. If $W_{b,t}^{-i}$ is a constant (as in the linear-in-means specification), we get initial estimates of $\alpha_i, \gamma\bar{\alpha}W_{b,t}^{-i} + \bar{\alpha} + \phi_{B(b),k(i),t}$, and γ. If $W_{b,t}^{-i}$ is not a constant, we can separately identify $\sigma = \gamma\bar{\alpha}$ and $\bar{\alpha} + \phi_{B(b),k(i),t}$. α_i is then updated using the updating rule below, derived by minimizing the associated nonlinear least square objective function.

The nonlinear least square estimator minimizes the following objective function:

$$\sum_{i \in I} \sum_{t \in T_i} \Big(y_{i,k,b,t} - \alpha_i - \bar{\alpha} - \gamma\bar{\alpha} \sum_{j \in M_{b,t} \setminus \{i\}} \omega_{ij}(M_{b,t}) - \phi_{B(b),k(i),t} - \gamma \sum_{j \in M_{b,t} \setminus \{i\}} \omega_{ij}(M_{b,t})\alpha_j \Big)^2$$

For the linear-in-means specification, $\omega_{ij}(M_{b,t}) = \frac{1}{|M_{b,t}|-1}$ and for the agglomeration specification, $\omega_{ij}(M_{b,t}) = 1$.

The first-order condition with respect to α_i is:

$$\begin{aligned} 0 = &-2 \sum_{t \in T_i} \Big(y_{i,k,b,t} - \alpha_i - \bar{\alpha} - \gamma\bar{\alpha}W_{b,t}^{-i} - \phi_{B(b),k(i),t} - \gamma \sum_{j \in M_{b,t} \setminus \{i\}} \omega_{ij}(M_{b,t})\alpha_j \Big) \\ &- 2 \sum_{t \in T_i} \sum_{j \in M_{b,t} \setminus \{i\}} \Big(y_{j,k,b,t} - \alpha_j - \bar{\alpha} - \gamma\bar{\alpha}W_{b,t}^{-j} - \phi_{B(b),k(j),t} - \gamma \sum_{j' \in M_{b,t} \setminus \{j\}} \omega_{jj'}(M_{b,t})\alpha_{j'} \Big) \gamma\omega_{ji}(M_{b,t}). \end{aligned}$$

Solving for α_i (Step 1/3):

$$\begin{aligned} T_i\alpha_i = &\sum_{t \in T_i} \Big(y_{i,k,b,t} - \bar{\alpha} - \gamma\bar{\alpha}W_{b,t}^{-i} - \phi_{B(b),k(i),t} - \gamma \sum_{j \in M_{b,t} \setminus \{i\}} \omega_{ij}(M_{b,t})\alpha_j \Big) \\ &+ \sum_{t \in T_i} \sum_{j \in M_{b,t} \setminus \{i\}} \Big(y_{j,k,b,t} - \alpha_j - \bar{\alpha} - \gamma\bar{\alpha}W_{b,t}^{-j} - \phi_{B(b),k(j),t} - \gamma \sum_{j' \in M_{b,t} \setminus \{i,j\}} \omega_{jj'}(M_{b,t})\alpha_{j'} \Big) \gamma\omega_{ji}(M_{b,t}) \\ &- \sum_{t \in T_i} \sum_{j \in M_{b,t} \setminus \{i\}} \gamma^2 \omega_{ji}(M_{b,t})^2 \alpha_i \end{aligned}$$

Solving for α_i (Step 2/3):

$$T_i\alpha_i + \sum_{t\in T_i}\sum_{j\in M_{b,t}\setminus\{i\}} \gamma^2\omega_{ji}(M_{b,t})^2\alpha_i =$$
$$\sum_{t\in T_i}\left(y_{i,k,b,t} - \bar{\alpha} - \gamma\bar{\alpha}W_{b,t}^{-i} - \phi_{B(b),k(i),t} - \gamma\sum_{j\in M_{b,t}\setminus\{i\}}\omega_{ij}(M_{b,t})\alpha_j\right)$$
$$+\sum_{t\in T_i}\sum_{j\in M_{b,t}\setminus\{i\}}\left(y_{j,k,b,t} - \alpha_j - \bar{\alpha} - \gamma\bar{\alpha}W_{b,t}^{-j} - \phi_{B(b),k(j),t} - \gamma\sum_{j'\in M_{b,t}\setminus\{i,j\}}\omega_{jj'}(M_{b,t})\alpha_{j'}\right)\gamma\omega_{ji}(M_{b,t})$$

Solving for α_i (Step 3/3):

$$\alpha_i =$$
$$\frac{1}{\left(T_i + \gamma^2\sum_{t\in T_i}\sum_{j\in M_{b,t}\setminus\{i\}}\omega_{ji}(M_{b,t})^2\right)}\times$$
$$\sum_{t\in T_i}\Bigg[\left(y_{i,k,b,t} - \bar{\alpha} - \gamma\bar{\alpha}W_{b,t}^{-i} - \phi_{B(b),k(i),t} - \gamma\sum_{j\in M_{b,t}\setminus\{i\}}\omega_{ij}(M_{b,t})\alpha_j\right)$$
$$+\gamma\sum_{j\in M_{b,t}\setminus\{i\}}\left(y_{j,k,b,t} - \alpha_j - \bar{\alpha} - \gamma\bar{\alpha}W_{b,t}^{-j} - \phi_{B(b),k(j),t} - \gamma\sum_{j'\in M_{b,t}\setminus\{i,j\}}\omega_{jj'}(M_{b,t})\alpha_{j'}\right)\omega_{ji}(M_{b,t})\Bigg]$$

In the linear-in-means specification with basic weights $\omega_{ij}(M_{b,t}) = \frac{1}{|M_{b,t}|-1}$, this expression is:

$$\alpha_i =$$
$$\frac{1}{\left(T_i + \gamma^2\sum_{t\in T_i}\frac{1}{|M_{b,t}|-1}\right)}\times$$
$$\sum_{t\in T_i}\Bigg[\left(y_{i,k,b,t} - \bar{\alpha}(1-\gamma) - \phi_{B(b),k(i),t} - \frac{\gamma}{|M_{b,t}|-1}\sum_{j\in M_{b,t}\setminus\{i\}}\alpha_j\right)$$
$$+\frac{\gamma}{|M_{b,t}|-1}\sum_{j\in M_{b,t}\setminus\{i\}}\left(y_{j,k,b,t} - \alpha_j - \bar{\alpha}(1-\gamma) - \phi_{B(b),k(j),t} - \frac{\gamma}{|M_{b,t}|-1}\sum_{j'\in M_{b,t}\setminus\{i,j\}}\alpha_{j'}\right)\Bigg]$$

In the agglomeration model with basic weights $\omega_{ji}(M_{b,t}) = 1$, this expression is:

$$\alpha_i = \frac{1}{\left(T_i + \gamma^2 \sum_{t\in T_i}(|M_{b,t}|-1)\right)} \times$$

$$\sum_{t\in T_i}\left[\left(y_{i,k,b,t} - \bar{\alpha}(1-\gamma) - \gamma\bar{\alpha}M_{b,t} - \phi_{B(b),k(i),t} - \gamma \sum_{j\in M_{b,t}\setminus\{i\}} \alpha_j\right)\right.$$

$$\left. + \gamma \sum_{j\in M_{b,t}\setminus\{i\}} \left(y_{j,k,b,t} - \alpha_j - \bar{\alpha}(1-\gamma) - \gamma\bar{\alpha}M_{b,t} - \phi_{B(b),k(j),t} - \gamma \sum_{j'\in M_{b,t}\setminus\{i,j\}} \alpha_{j'}\right)\right]$$

B2. *Horse race*

We carry out the analogous process for the horse race specification. For estimation, we replace $\gamma_{\text{Agg}} \sum_{j\in M_{b,t},\neq i} \alpha_j + \ddot{\sigma}(|M_{b,t}|-1)$ in the baseline horse race estimation equation (8) in the main text with its generalized counterpart $\gamma_{\text{W}} \sum_{j\in M_{b,t},\neq i} \alpha_j \omega_{ij}(M_{b,t})$ $+\ddot{\sigma}_{\text{W}} \sum_{j\in M_{b,t},\neq i} \omega_{ij}(M_{b,t})$. Define two weights, one for each element of the horse race

$$W^{-i}_{q,b,t} = \sum_{j\in M_{b,t}\setminus\{i\}} \omega_s(k(i), k(j), M_{b,t})$$

where $q \in \{m, s\}$. The nonlinear least square estimator minimizes the following objective function:

$$\sum_{i\in I}\sum_{t\in T_i}\left(y_{i,k,b,t} - \alpha_i - \bar{\alpha} - \gamma_s\bar{\alpha}W^{-i}_{s,b,t} - \gamma_m\bar{\alpha}W^{-i}_{m,b,t} - \phi_{k(i),B(b),t}\right.$$

$$\left. - \gamma_s \sum_{j\in M_{b,t}\setminus\{i\}} \omega_s(k(i), k(j), M_{b,t})\alpha_j - \gamma_m \sum_{j\in M_{b,t}\setminus\{i\}} \omega_m(k(i), k(j), M_{b,t})\alpha_j\right)^2$$

The first-order condition with respect to α_i:

$$0 = -2\sum_{t\in T_i}\left(y_{i,k,b,t} - \alpha_i - \bar{\alpha} - \gamma_s\bar{\alpha}W^{-i}_{s,b,t} - \gamma_m\bar{\alpha}W^{-i}_{m,b,t} - \phi_{k(i),B(b),t}\right.$$

$$\left. - \gamma_s \sum_{j\in M_{b,t}\setminus\{i\}} \omega_s\left(k\left(i\right), k\left(j\right), M_{b,t}\right)\alpha_j - \gamma_m \sum_{j\in M_{b,t}\setminus\{i\}} \omega_m\left(k\left(i\right), k\left(j\right), M_{b,t}\right)\alpha_j\right)$$

$$
\begin{aligned}
&= -2 \sum_{t \in T_i} \sum_{j \in M_{b,t} \setminus \{i\}} \Bigg[\Bigg(y_{j,k,b,t} - \alpha_j - \bar{\alpha} - \gamma_s \bar{\alpha} W^{-j}_{s,b,t} - \gamma_m \bar{\alpha} W^{-j}_{m,b,t} - \phi_{k(j),B(b),t} \\
&- \gamma_s \sum_{j' \in M_{b,t} \setminus \{j\}} \omega_s(k(j), k(j'), M_{b,t}) \alpha_{j'} - \gamma_m \sum_{j' \in M_{b,t} \setminus \{j\}} \omega_m(k(j), k(j'), M_{b,t}), M_{b,t}) \alpha_{j'} \Bigg) \\
&\times \Big(\gamma_s \omega_s(k(j), k(i), M_{b,t}) + \gamma_m \omega_m(k(j), k(i), M_{b,t}) \Big) \Bigg]
\end{aligned}
$$

Solving for α_i (Step 1/2):

$$
\begin{aligned}
&T_i \alpha_i + \alpha_i \sum_{t \in T_i} \sum_{j \in M_{b,t} \setminus \{i\}} \Big(\gamma_s \omega_s(k(j), k(i), M_{b,t}) + \gamma_m \omega_m(k(j), k(i), M_{b,t}) \Big)^2 \\
&= \sum_{t \in T_i} \Bigg(y_{i,k,b,t} - \bar{\alpha} - \gamma_s \bar{\alpha} W^{-i}_{s,b,t} - \gamma_m \bar{\alpha} W^{-i}_{m,b,t} - \phi_{k(i),B(b)} - \gamma_s \sum_{j \in M_{b,t} \setminus \{i\}} \omega_s(k(i), k(j), M_{b,t}) \alpha_j \\
&- \gamma_m \sum_{j \in M_{b,t} \setminus \{i\}} \omega_m(k(i), k(j), M_{b,t}) \alpha_j \Bigg) \\
&+ \sum_{t \in T_i} \sum_{j \in M_{b,t} \setminus \{i\}} \Bigg[\Bigg(y_{j,k,b,t} - \alpha_j - \bar{\alpha} - \gamma_s \bar{\alpha} W^{-j}_{s,b,t} - \gamma_m \bar{\alpha} W^{-j}_{m,b,t} - \phi_{k(j),B(b)} \\
&- \gamma_s \sum_{j' \in M_{b,t} \setminus \{i,j\}} \omega_s(k(j), k(j'), M_{b,t}) \alpha_{j'} - \gamma_m \sum_{j' \in M_{b,t} \setminus \{i,j\}} \omega_m(k(j), k(j'), M_{b,t}) \alpha_{j'} \Bigg) \\
&\times \Big(\gamma_s \omega_s(k(j), k(i), M_{b,t}) + \gamma_m \omega_m(k(j), k(i), M_{b,t}) \Big) \Bigg]
\end{aligned}
$$

Solving for α_i (Step 2/2):

$$\alpha_i =$$
$$\frac{1}{T_i + \sum_{t\in T_i}\sum_{j\in M_{b,t}\setminus\{i\}}\Big(\gamma_s\omega_s(k(j),k(i),M_{b,t}) + \gamma_m\omega_m(k(j),k(i),M_{b,t})\Big)^2}\times$$
$$\sum_{t\in T_i}\Bigg[\Bigg(y_{i,k,b,t} - \bar{\alpha} - \gamma_s\bar{\alpha}W^{-i}_{s,b,t} - \gamma_m\bar{\alpha}W^{-i}_{m,b,t} - \phi_{k(i),B(b)}$$
$$-\gamma_s\sum_{j\in M_{b,t}\setminus\{i\}}\omega_s(k(i),k(j),M_{b,t})\alpha_j - \gamma_m\sum_{j\in M_{b,t}\setminus\{i\}}\omega_m(k(i),k(j),M_{b,t})\alpha_j\Bigg)$$
$$+\sum_{j\in M_{b,t}\setminus\{i\}}\Bigg[\Bigg(y_{j,k,b,t} - \alpha_j - \bar{\alpha} - \gamma_s\bar{\alpha}W^{-j}_{s,b,t} - \gamma_m\bar{\alpha}W^{-j}_{m,b,t} - \phi_{k(j),B(b)}$$
$$-\gamma_s\sum_{j'\in M_{b,t}\setminus\{i,j\}}\omega_s(k(j),k(j'),M_{b,t})\alpha_{j'} - \gamma_m\sum_{j'\in M_{b,t}\setminus\{i,j\}}\omega_m(k(j),k(j'),M_{b,t})\alpha_{j'}\Bigg)\times$$
$$\Big(\gamma_s\omega_s(k(j),k(i),M_{b,t}) + \gamma_m\omega_m(k(j),k(i),M_{b,t})\Big)\Bigg]\Bigg]$$

Appendix C. Proofs of Consistency

This appendix analyzes the consistency of an estimator of spillovers between firms based on the minimization of the squared prediction errors. The proofs shown here mimic the proof of Theorem 1 in Arcidiacono et al. (2012) (AFGK), in particular the first four lemmas where they show the consistency of their estimator. Throughout this section, firm i in peer group n at time t is characterized by a fixed effect α_i and a shock $\epsilon_{i,t,n}$. We analyze:

1) The consistency of an estimator where spillovers of firm j to firm i in peer group n are weighted by a known weight $\omega_{i,j,n}$;
2) The consistency of a horse race estimator where both agglomeration spillovers ($\omega_{i,j,n} = 1$) and linear-in-mean spillovers ($\omega_{i,j,n} = \frac{1}{|M_{n,t}|-1}$, where $M_{n,t}$ denotes the set of firms in peer group n at time t) operate simultaneously;
3) The lack of consistency and bias of the estimator when the number of groups N goes to infinity, but the peer group has a fixed time dimension, in particular $T = 2$, and the shocks $\epsilon_{i,t,n}$ are autocorrelated;
4) The consistency of the estimator when both N and T go to infinity and the shocks are autocorrelated.

Throughout this section, we maintain most of AFGK's assumptions:

(i) $E(\epsilon_{i,t,n}\epsilon_{j,t,k}) = 0$ for all $j \neq i$ and $n \neq k$.

(ii) $E(\epsilon_{i,t,n}\alpha_j) = 0$ for all i, j, t, n.

(iii) $E\left(\alpha_{in}^4\right) < \infty$ for all i, n.

(iv) $E(\epsilon_{i,t,n}) = 0$ and $E\left(\epsilon_{i,t,n}^4\right) < \infty$ for all i, t, n.

(v) $E\left(\epsilon_{i,t,n}^2 | n, t\right) = E\left(\epsilon_{j,t,n}^2 | n, t\right)$ for all i, j, t, n.

(vi) The parameter $\gamma \in \Gamma$ where Γ is compact.

In the first two cases, as in AFGK, we also assume that $E(\epsilon_{i,t,n}\epsilon_{j,s,k}) = 0$ for $t \neq s$. This assumption is relaxed in the other two cases. Furthermore, we assume that the fixed effects $\{\alpha_{in}\}$ are not linear combinations of each other in i to guarantee uniqueness of the solutions, which was an implicit assumption in AFGK.

As in AFGK, we analyze consistency using a simplified structure of peer groups with a limited number of firms and periods (expect for Case 4). We do not expect this simplification to affect the general results. In the same spirit, we do not allow firm i's outcome to be affected by other covariates, except for peer effects. That is, we do not include industry-year or local area-year fixed effects as in our main analysis. Again, we do not believe that the general message of this section is affected by this choice of exposition.

Case 1 - General Weights

Outcome $y_{i,t,n}$ of firm i in peer group n at time t is:

$$y_{i,t,n} = \alpha_i + \gamma \sum_{j \in M_{n,t} \neq i} \omega_{i,j,n,t} \alpha_j + \epsilon_{i,t,n}$$

In addition to the six assumptions listed above, we assume that $E\left(\epsilon_{i,t,n}\epsilon_{j,s,k}\right) = 0$ for $t \neq s$. We consider the following limiting case:

(a) Firms are observed for at most two periods.

(b) Each peer group has two firms in each period.

(c) Within each peer group, one firm is observed for two periods and the other firm is observed for one period only.

The optimization problem is

$$\min_{\alpha,\gamma} \frac{1}{N} \sum_{n=1}^{N} ((y_{11n} - \alpha_{1n} - \gamma\omega_{12n}\alpha_{2n})^2 + (y_{12n} - \alpha_{1n} - \gamma\omega_{13n}\alpha_{3n})^2 + (y_{21n} - \alpha_{2n} - \gamma\omega_{21n}\alpha_{1n})^2 + (y_{32n} - \alpha_{3n} - \gamma\omega_{31n}\alpha_{1n})^2)$$

where we omit the time period index in the weight given that two firms meet for only one time period.

Following AFGK, we first concentrate out the αs. Taking the first-order conditions and solving for the firm fixed effects (omitting the index n), we get:

$$\alpha_1 = \frac{\left(\left(1 + (\gamma\omega_{13})^2\right)(1 - \gamma\omega_{21}\gamma\omega_{12})(y_{11} - \gamma\omega_{12}y_{21}) + \left(1 + (\gamma\omega_{12})^2\right)(1 - \gamma\omega_{13}\gamma\omega_{31})(y_{12} - \gamma\omega_{13}y_{32})\right)}{\left((1 - \gamma\omega_{31}\gamma\omega_{13})^2\left(1 + (\gamma\omega_{12})^2\right) + (1 - \gamma\omega_{21}\gamma\omega_{12})^2\left(1 + (\gamma\omega_{13})^2\right)\right)}$$

$$\alpha_2 = \frac{\left(\begin{array}{l} \left((1-\gamma\omega_{31}\gamma\omega_{13})^2\gamma\omega_{12} - \gamma\omega_{21}\left(1+(\gamma\omega_{13})^2\right)(1-\gamma\omega_{21}\gamma\omega_{12})\right)y_{11} \\ \quad + \left((1-\gamma\omega_{31}\gamma\omega_{13})^2 + (1-\gamma\omega_{21}\gamma\omega_{12})\left(1+(\gamma\omega_{13})^2\right)\right)y_{21} \\ \qquad - (\gamma\omega_{21}+\gamma\omega_{12})(1-\gamma\omega_{13}\gamma\omega_{31})\,y_{12} + (\gamma\omega_{21}+\gamma\omega_{12})(1-\gamma\omega_{13}\gamma\omega_{31})\,\gamma\omega_{13}y_{32} \end{array}\right)}{\left((1-\gamma\omega_{31}\gamma\omega_{13})^2\left(1+(\gamma\omega_{12})^2\right) + (1-\gamma\omega_{21}\gamma\omega_{12})^2\left(1+(\gamma\omega_{13})^2\right)\right)}$$

$$\alpha_3 = \frac{\left(\begin{array}{l} \left((1-\gamma\omega_{21}\gamma\omega_{12})^2\gamma\omega_{13} - \gamma\omega_{31}\left(1+(\gamma\omega_{12})^2\right)(1-\gamma\omega_{13}\gamma\omega_{31})\right)y_{12} \\ \quad + \left((1-\gamma\omega_{21}\gamma\omega_{12})^2 + (1-\gamma\omega_{13}\gamma\omega_{31})\left(1+(\gamma\omega_{12})^2\right)\right)y_{32} \\ \qquad - (\gamma\omega_{31}+\gamma\omega_{13})(1-\gamma\omega_{21}\gamma\omega_{12})\,y_{11} + (\gamma\omega_{31}+\gamma\omega_{13})(1-\gamma\omega_{21}\gamma\omega_{12})\,\gamma\omega_{12}y_{21} \end{array}\right)}{\left((1-\gamma\omega_{31}\gamma\omega_{13})^2\left(1+(\gamma\omega_{12})^2\right) + (1-\gamma\omega_{21}\gamma\omega_{12})^2\left(1+(\gamma\omega_{13})^2\right)\right)}$$

Note that the above expressions simplify to the same ones as in AFGK when all weights are equal to 1:

$$\begin{aligned} \alpha_1 &= \frac{(y_{11} - \gamma y_{21}) + (y_{12} - \gamma y_{32})}{2(1-\gamma^2)} \\ \alpha_2 &= \frac{y_{21} - \gamma^3 y_{11} - \gamma y_{12} + \gamma^2 y_{32}}{(1-\gamma^4)} \\ \alpha_3 &= \frac{y_{32} - \gamma y_{11} + \gamma^2 y_{21} - \gamma^3 y_{12}}{(1-\gamma^4)} \end{aligned}$$

After several substitutions, the original minimization problem becomes:

$$\min_{\gamma} \frac{1}{N}\sum_{n=1}^{N} \frac{\left((1-\gamma\omega_{31n}\gamma\omega_{13n})(y_{11n} - \gamma\omega_{12n}y_{21n}) - (1-\gamma\omega_{21n}\gamma\omega_{12n})(y_{12n} - \gamma\omega_{13n}y_{32n})\right)^2}{(1-\gamma\omega_{31n}\gamma\omega_{13n})^2\left(1+(\gamma\omega_{12n})^2\right) + (1-\gamma\omega_{21n}\gamma\omega_{12n})^2\left(1+(\gamma\omega_{13n})^2\right)}$$

which is the same as in AFGK when the weights are equal to 1, i.e.

$$\min_{\gamma} \frac{1}{N}\sum_{n=1}^{N} \frac{(y_{11n} - \gamma y_{21n} - y_{12n} + \gamma y_{32n})^2}{2(1+\gamma^2)}.$$

Substituting for the true data generating process:

$$y_{i,t,n} = \alpha_i^o + \gamma_0 \sum_{j \in M_{n,t} \neq i} \omega_{i,j,n,t}\alpha_j^o + \epsilon_{i,t,n}^o$$

we obtain

$$\min_{\gamma} \frac{1}{N} \sum_{n=1}^{N} q\left(y_n, \gamma\right)$$

where (omitting the index n):

$$q\left(y_n, \gamma\right) = \frac{\begin{pmatrix} (1-\gamma\omega_{31}\gamma\omega_{13})\left(\alpha_1^o + \gamma_0\omega_{12}a_2^o + \epsilon_{11}^\circ - \gamma\omega_{12}\left(\alpha_2^o + \gamma_0\omega_{21}a_1^o + \epsilon_{21}^o\right)\right) \\ -\left(1-\gamma\omega_{21}\gamma\omega_{12}\right)\left(a_1^o + \gamma_0\omega_{13}a_3^o + \epsilon_{12}^\circ - \gamma\omega_{13}\left(\alpha_3^o + \gamma_0\omega_{31}a_1^o + \epsilon_{32}^o\right)\right)\end{pmatrix}^2}{\left(1-\gamma\omega_{31}\gamma\omega_{13}\right)^2\left(1+\left(\gamma\omega_{12}\right)^2\right) + \left(1-\gamma\omega_{21}\gamma\omega_{12}\right)^2\left(1+\left(\gamma\omega_{13}\right)^2\right)}.$$

Consider the expected value of the function $q(y_n, \gamma)$. Using assumptions (i), (ii), and (iv), the expression simplifies to:

$$\mathbb{E}(q(y,\gamma)) = \sigma_\epsilon^2 + \mathbb{E}\left(\frac{\begin{pmatrix}\left(\left(1-\gamma\omega_{31}\gamma\omega_{13}\right)\left(1-\gamma\omega_{12}\gamma_0\omega_{21}\right) - \left(1-\gamma\omega_{21}\gamma\omega_{12}\right)\left(1-\gamma\omega_{13}\gamma_0\omega_{31}\right)\right)\alpha_1^o \\ + \left(1-\gamma\omega_{31}\gamma\omega_{13}\right)\left(\gamma_0-\gamma\right)\omega_{12}a_2^o - \left(1-\gamma\omega_{21}\gamma\omega_{12}\right)\left(\gamma_0-\gamma\right)\omega_{13}a_3^o\end{pmatrix}^2}{\left(1-\gamma\omega_{31}\gamma\omega_{13}\right)^2\left(1+\left(\gamma\omega_{12}\right)^2\right) + \left(1-\gamma\omega_{21}\gamma\omega_{12}\right)^2\left(1+\left(\gamma\omega_{13}\right)^2\right)}\right).$$

The first term does not depend on γ. Because the denominator of the second term is positive and the numerator is squared, it is equal to zero when $\gamma = \gamma_0$ while strictly positive when $\gamma \neq \gamma_0$. Hence, $\mathbb{E}[q(y,\gamma_0)] < \mathbb{E}[q(y,\gamma)]$ for all $\gamma \in \Gamma$ such that $\gamma \neq \gamma_0$.

As suggested in AFGK, most of the requirements to apply Theorem 12.1 in Wooldridge (2010) are satisfied with the exception of the following: For all $\gamma \in \Gamma$, $|q(\gamma, y)| \leq b(y)$ where b is a non-negative function such that $\mathbb{E}(b(y)) < \infty$ Given that $q(\gamma, y)$ is always positive we can ignore the absolute value. Going back to the definition of $q(\gamma, y)$:

$$q(\gamma, y) = \frac{\left(\left(1-\gamma\omega_{31}\gamma\omega_{13}\right)\left(y_{11}-\gamma\omega_{12}y_{21}\right) - \left(1-\gamma\omega_{21}\gamma\omega_{12}\right)\left(y_{12}-\gamma\omega_{13}y_{32}\right)\right)^2}{\left(1-\gamma\omega_{31}\gamma\omega_{13}\right)^2\left(1+\left(\gamma\omega_{12}\right)^2\right) + \left(1-\gamma\omega_{21}\gamma\omega_{12}\right)^2\left(1+\left(\gamma\omega_{13}\right)^2\right)}$$

and using the triangular inequality, we can show that $q(\gamma, y) < 2\left(y_{11}^2 + 2y_{21}^2 + 2y_{12}^2 + 2y_{32}^2\right)$. The last step is to show that $\mathbb{E}\left(2y_{11n}^2 + 2y_{21n}^2 + 2y_{12n}^2 + 2y_{32n}^2\right) < \infty$. This follows exactly the proof in AFGK and we omit it here. Theorem 12.1 in Wooldridge (2010) can be applied to obtain $\hat{\gamma} \xrightarrow{p} \gamma_0$.

Case 2 - Horse Race

We now consider the consistency of a horse race estimator where both agglomeration spillovers ($\omega_{i,j,n}^a = 1$) and linear-in-mean spillovers ($\omega_{i,j,n}^b = \frac{1}{|M_{n,t}|-1}$, where $M_{n,t}$ denotes the set of firms in peer group n at time t) operate simultaneously. In this case, outcome $y_{i,t,n}$ of firm i in peer group

n at time t is:

$$y_{i,t,n} = \alpha_i + \gamma \sum_{j \in M_{n,t} \neq i} \alpha_{jn} + \rho \sum_{j \in M_{n,t} \neq i} \frac{\alpha_{jn}}{|M_{n,t}| - 1} + \epsilon_{i,t,n}.$$

We consider a situation where each peer group has the following composition:

(a) Firm $0n$ is observed for 3 periods.

(b) Firm $1n$ is observed only in the first period.

(c) Firm $2n$ is observed only in the second period.

(d) Firm $3n$ and $4n$ are observed only in the third period.

The optimization problem is

$$\min_{\alpha,\gamma,\rho} \frac{1}{N} \sum_{i=1}^{N} \left(\epsilon_{01n}^2 + \epsilon_{02n}^2 + \epsilon_{03n}^2 + \epsilon_{11n}^2 + \epsilon_{22n}^2 + \epsilon_{33n}^2 + \epsilon_{43n}^2 \right)$$

where

$$\begin{aligned}
\epsilon_{01n} &= y_{01n} - \alpha_{0n} - (\gamma + \rho)\alpha_{1n} \\
\epsilon_{02n} &= y_{02n} - \alpha_{0n} - (\gamma + \rho)\alpha_{2n} \\
\epsilon_{33n} &= y_{03n} - \alpha_{0n} - \left(\gamma + \frac{\rho}{2}\right)(\alpha_{3n} + \alpha_{4n}) \\
\epsilon_{11n} &= y_{11n} - \alpha_{1n} - (\gamma + \rho)\alpha_{0n} \\
\epsilon_{22n} &= y_{22n} - \alpha_{2n} - (\gamma + \rho)\alpha_{0n} \\
\epsilon_{33n} &= y_{33n} - \alpha_{3n} - \left(\gamma + \frac{\rho}{2}\right)(\alpha_{0n} + \alpha_{4n}) \\
\epsilon_{43n} &= y_{43n} - \alpha_{4n} - \left(\gamma + \frac{\rho}{2}\right)(\alpha_{0n} + \alpha_{3n}).
\end{aligned}$$

We first concentrate out α_1, α_2, α_3, and α_4 for all groups. From the first-order conditions, after several substitutions, we are able to rewrite the argument of the summation above as:

$$\begin{aligned}
&\epsilon_{01n}^2 + \epsilon_{02n}^2 + \epsilon_{03n}^2 + \epsilon_{11n}^2 + \epsilon_{12n}^2 + \epsilon_{33n}^2 + \epsilon_{43n}^2 \\
&= \frac{\left(y_{01n} - (\gamma+\rho) y_{11n} - \left(1 - (\gamma+\rho)^2\right)\alpha_{0n}\right)^2 + \left(y_{02n} - (\gamma+\rho) y_{22n} - \left(1 - (\gamma+\rho)^2\right)\alpha_{0n}\right)^2}{1 + (\gamma+\rho)^2} \\
&+ \frac{\left(\left(\gamma + \frac{\rho}{2}\right) y_{33n} + \left(\gamma + \frac{\rho}{2}\right) y_{43n} - \left(1 + \left(\gamma + \frac{\rho}{2}\right)\right) y_{03n} + \left(1 - \left(\gamma + \frac{\rho}{2}\right)\right)\left(1 + 2\left(\gamma + \frac{\rho}{2}\right)\right)\alpha_{0n}\right)^2}{\left(3\left(\gamma + \frac{\rho}{2}\right)^2 + 2\left(\gamma + \frac{\rho}{2}\right) + 1\right)}
\end{aligned}$$

Instead of concentrating out $\alpha_{0\mathrm{n}}$, we work directly with the minimization of the expected criterion function. To do so, we substitute for the true data generating process

$$
\begin{aligned}
y_{01n} &= \alpha^o_{0n} + (\gamma_0 + \rho_0)\,\alpha^o_{1n} + \epsilon^o_{01n}\\
y_{02n} &= \alpha^o_{0n} + (\gamma_0 + \rho_0)\,\alpha^o_{2n} + \epsilon^o_{02n}\\
y_{03n} &= \alpha^o_{0n} + \left(\gamma_0 + \frac{\rho_0}{2}\right)(\alpha^o_{3n} + \alpha^o_{4n}) + \epsilon^o_{03n}\\
y_{11n} &= \alpha^o_{1n} + (\gamma_0 + \rho_0)\,\alpha^o_{0n} + \epsilon^o_{11n}\\
y_{22n} &= \alpha^o_{2n} + (\gamma_0 + \rho_0)\,\alpha^o_{0n} + \epsilon^o_{22n}\\
y_{33n} &= \alpha^o_{3n} + \left(\gamma_0 + \frac{\rho_0}{2}\right)(\alpha^o_{0n} + \alpha^o_{4n}) + \epsilon^o_{33n}\\
y_{43n} &= \alpha^o_{4n} + \left(\gamma_0 + \frac{\rho_0}{2}\right)(\alpha^o_{0n} + \alpha^o_{3n}) + \epsilon^o_{43n}
\end{aligned}
$$

The expected value of that object can then be written as

$$
\begin{aligned}
&\mathbb{E}\left(\epsilon^2_{01n} + \epsilon^2_{02n} + \epsilon^2_{03n} + \epsilon^2_{11n} + \epsilon^2_{22n} + \epsilon^2_{33n} + \epsilon^2_{43n}\right)\\
&= \mathbb{E}\frac{\left(\left((\gamma_0+\rho_0) - (\gamma+\rho)\right)\alpha^o_{1n} + \left(1 - (\gamma+\rho)(\gamma_0+\rho_0)\right)\alpha^o_{0n} - \left(1 - (\gamma+\rho)^2\right)\alpha_{0n}\right)^2}{1 + (\gamma+\rho)^2}\\
&+ \mathbb{E}\frac{\left(\left((\gamma_0+\rho_0) - (\gamma+\rho)\right)\alpha^o_{2n} + \left(1 - (\gamma+\rho)(\gamma_0+\rho_0)\right)\alpha^o_{0n} - \left(1 - (\gamma+\rho)^2\right)\alpha_{0n}\right)^2}{1 + (\gamma+\rho)^2}\\
&+ \mathbb{E}\frac{\left(\begin{array}{l}\left(\left(\gamma+\frac{\rho}{2}\right)\left(1+\left(\gamma_0+\frac{\rho_0}{2}\right)\right) - \left(\gamma_0+\frac{\rho_0}{2}\right)\left(1+\left(\gamma+\frac{\rho}{2}\right)\right)\right)(\alpha^o_{3n}+\alpha^o_{4n})\\ \quad + \left(1+\left(\gamma+\frac{\rho}{2}\right)\right)(\alpha_{0n}-\alpha^o_{0n}) + 2\left(\gamma+\frac{\rho}{2}\right)\left(\left(\gamma_0+\frac{\rho_0}{2}\right)\alpha^o_{0n} - \left(\gamma+\frac{\rho}{2}\right)\alpha_{0n}\right)\end{array}\right)^2}{2\left(\gamma+\frac{\rho}{2}\right)^2 + \left(1+\left(\gamma+\frac{\rho}{2}\right)\right)^2}\\
&+ 3\sigma^2_{\epsilon n}
\end{aligned}
$$

using assumptions (i), (ii), and (v) and assuming $E\left(\epsilon_{i,t,n}\epsilon_{j,s,k}\right) = 0$ for $t \neq s$. Given that the first three terms on the right hand side are non-negative, the minimum is attained at $3\sigma^2_{en}$ when $\alpha_{0n} = \alpha^o_{0n}$, $\gamma = \gamma_0$, and $\rho = \rho_o$. To complete the proof of consistency, as in the previous case, one has to show that the original object can be bounded for all possible γs. We omit this step which can be achieved in a similar fashion to the previous case.

Case 3 - Autocorrelated Errors

As in the first case, we assume that peer groups have a simplified composition:

(a) We observe firms for at most two periods.

(b) Each peer group has two firms in each period.

(c) Within each peer group, one firm is observed for two periods and the other firm is observed for one period only.

Outcome $y_{i,t,n}$ of firm i in peer group n at time t is:

$$y_{i,t,n} = \alpha_i + \gamma \sum_{j \in M_{n,t} \neq i} \alpha_j + \epsilon_{itn}$$

where we assume $\epsilon_{i,t,n} = \rho \epsilon_{i,t-1,n} + u_{i,t,n}$. The optimization problem is

$$\min_{\alpha,\gamma} \frac{1}{N} \sum_{n=1}^{N} ((y_{11n} - \alpha_{1n} - \gamma \alpha_{2n})^2 + (y_{12n} - \alpha_{1n} - \gamma \alpha_{3n})^2$$
$$+ (y_{21n} - \alpha_{2n} - \gamma \alpha_{1n})^2 + (y_{32n} - \alpha_{3n} - \gamma \alpha_{1n})^2).$$

As in the first case we concentrate out the αs. From the first-order conditions, we find that:

$$\alpha_{1n} = \frac{(y_{11n} - \gamma y_{21n}) + (y_{12n} - \gamma y_{32n})}{2(1-\gamma^2)}$$
$$\alpha_{2n} = \frac{y_{21n} - \gamma^3 y_{11n} - \gamma y_{12n} + \gamma^2 y_{32n}}{(1-\gamma^4)}$$
$$\alpha_{3n} = \frac{y_{32n} - \gamma y_{11n} + \gamma^2 y_{21n} - \gamma^3 y_{12n}}{(1-\gamma^4)}$$

Substituting, we get:

$$\min_{\gamma} \frac{1}{N} \sum_{n=1}^{N} \frac{(y_{11n} - \gamma y_{21n} - y_{12n} + \gamma y_{32n})^2}{2(1+\gamma^2)}$$

Substituting y with the true data generating process yields

$$\min_{\gamma} \frac{1}{N} \sum_{n=1}^{N} \frac{((\gamma_0 - \gamma)\alpha^o_{2n} - (\gamma_0 - \gamma) a^o_{3n} + \epsilon^o_{11} - \epsilon^o_{12n} + \gamma \epsilon^o_{32n} - \gamma \epsilon^o_{21n})^2}{2(1+\gamma^2)}.$$

Consider the expected value of the argument within the summation

$$\mathbb{E}\left(\frac{\left(\left(\gamma_0-\gamma\right)\alpha_2^o-\left(\gamma_0-\gamma\right)\alpha_3^o+\epsilon_{11}^o-\epsilon_{12}^o+\gamma\epsilon_{32}^o-\gamma\epsilon_{21}^o\right)^2}{2\left(1+\gamma^2\right)}\right)$$
$$=\mathbb{E}\left(\frac{\left(\gamma_0-\gamma\right)^2\left(\alpha_{20n}-\alpha_{30n}\right)^2}{2\left(1+\gamma^2\right)}\right)$$
$$+\mathbb{E}\left(\frac{\left(\epsilon_{11n}-\epsilon_{12n}+\gamma\epsilon_{32n}-\gamma\epsilon_{21n}\right)^2}{2\left(1+\gamma^2\right)}\right)$$
$$+\mathbb{E}\left(\frac{2\left(\gamma_0-\gamma\right)\left(\alpha_{20n}-\alpha_{30n}\right)\left(\epsilon_{11n}-\epsilon_{12n}+\gamma\epsilon_{32n}-\gamma\epsilon_{21n}\right)}{2\left(1+\gamma^2\right)}\right).$$

All three terms on the right hand side are non-negative. The third term is equal to zero because of assumption (ii). Assuming that $\sigma_{\alpha\alpha}$ is the covariance between α_{20n} and α_{30n} and that the variance of α is the same across different αs, the whole expression becomes:

$$\frac{\left(\gamma_0-\gamma\right)^2}{\left(1+\gamma^2\right)}\left(\sigma_\alpha^2-\sigma_{\alpha\alpha}\right)+\sigma_\epsilon^2-\frac{\rho\sigma_\epsilon^2}{\left(1+\gamma^2\right)}$$

using assumptions (i), (iv), and (v). If $\rho=0$, this expression is minimized at $\gamma=\gamma_0$. If $\rho\neq 0$, minimizing the above leads to the following expression:

$$\left(\gamma_0-\gamma\right)\left(1+\gamma\gamma_0\right)=\frac{\gamma\rho\sigma_\epsilon^2}{\left(\sigma_\alpha^2-\sigma_{\alpha\alpha}\right)}$$

Assuming that γ_0 and ρ are positive, γ cannot be equal to γ_0 because the left hand side would be 0 while the right hand side would be strictly positive. The optimal solution to the limiting objective function is not γ_0, and hence the estimator of γ would be asymptotically biased.

Case 4 - $T\to\infty$

Here, we show that the bias in Case 3 disappears if we allow for T to diverge as well. As in Case 3, we assume (in addition to the standard six assumptions) that

$$\epsilon_{i,t,n}=\rho\epsilon_{i,t-1,n}+u_{i,t,n}.$$

We consider the following simplified structure of peer groups:

(a) Each peer group has two firms in each period.

(b) Within each group, one firm is observed for T periods and the other firm is observed for one period only.

Outcome $y_{i,t,n}$ of firm i in peer group n at time t is:

$$y_{i,t,n} = \alpha_i + \gamma \sum_{j \in M_{n,t} \neq i} \alpha_j + \epsilon_{i,t,n}$$

which is consistent with either linear-in-means or aggregate spillovers. To simplify notation, the staying firm is characterized by the indices $(0, t)$ and the one-period firms by (t, t) so the time t identifies those firms. Dropping the peer group index n, the optimization problem is:

$$\min_{a,\gamma} \frac{1}{N} \sum_{n=1}^{N} \frac{1}{T} \left(\sum_{t=1}^{T} (y_{0t} - \alpha_0 - \gamma\alpha_t)^2 + (y_{tt} - \alpha_t - \gamma\alpha_0)^2 \right)$$

First we concentrate out the fixed effects for the one-period firms. From the first-order conditions, we have:

$$\alpha_t = \frac{\gamma (y_{0t} - \alpha_0) + (y_{tt} - \gamma\alpha_0)}{1 + \gamma^2}$$

Substituting back into the problem:

$$\min_{a,\gamma} \frac{1}{N} \sum_{n=1}^{N} \frac{1}{T} \left(\sum_{t=1}^{T} \frac{((y_{0t} - \alpha_0) - \gamma (y_{tt} - \gamma\alpha_0))^2}{1 + \gamma^2} \right).$$

The first-order condition for α_0 leads to

$$\alpha_0 = \frac{\sum_{t=1}^{T} (y_{0t} - \gamma y_{tt})}{T (1 - \gamma^2)}.$$

Substituting back in yields

$$\min_{\gamma} \frac{1}{N} \sum_{n=1}^{N} \frac{1}{T} \left(\sum_{t=1}^{T} \frac{\left(\left(y_{0t} - \frac{\sum_{\tau=1}^{T} (y_{0\tau} - \gamma y_{\tau\tau})}{T(1-\gamma^2)} \right) - \gamma \left(y_{tt} - \gamma \frac{\sum_{\tau=1}^{T} (y_{0\tau} - \gamma y_{\tau\tau})}{T(1-\gamma^2)} \right) \right)^2}{1 + \gamma^2} \right).$$

Next, we substitute in the true data generating process:

$$y_{0t} = \alpha_0^o + \gamma_0 \alpha_t^o + \epsilon_{0t}^o$$

$$y_{tt} = \alpha_t^o + \gamma_0 \alpha_0^o + \epsilon_{tt}^o$$

to obtain

$$\min_{\gamma} q_{N,T}(\gamma, y) \equiv$$
$$\min_{\gamma} \frac{1}{N}\sum_{n=1}^{N}\frac{1}{T}\left(\sum_{t=1}^{T}\frac{\left((\gamma_0-\gamma)\left(\alpha_t^o-\sum_{\tau=1}^{T}\frac{\alpha_\tau^o}{T}\right)+\epsilon_{0t}^o-\gamma\epsilon_{tt}^o-\sum_{\tau=1}^{T}\frac{\epsilon_{0\tau}^o-\gamma\epsilon_{\tau\tau}^o}{T}\right)^2}{1+\gamma^2}\right).$$

We consider the performance of the estimator under a sequential asymptotic framework where we first let $N \to \infty$, and then let $T \to \infty$ (i.e. $(N,T) \overset{seq}{\to} \infty$).

Define the limiting objective function as

$$q(\gamma, y) \equiv \lim_{T\to\infty} \underset{N\to\infty}{plim}\, q_{N,T}(\gamma, y)$$

and we first consider $plim_{N\to\infty}\, q_{N,T}(\gamma, y)$:

$$\mathbb{E}\left[\frac{1}{T}\left(\sum_{t=1}^{T}\frac{\left((\gamma_0-\gamma)\left(\alpha_t^o-\sum_{\tau=1}^{T}\frac{\alpha_\tau^o}{T}\right)+\epsilon_{0t}^o-\gamma\epsilon_{tt}^o-\sum_{\tau=1}^{T}\frac{\epsilon_{0\tau}^o-\gamma\epsilon_{\tau\tau}^o}{T}\right)^2}{1+\gamma^2}\right)\right]$$
$$=\frac{1}{T}\frac{1}{1+\gamma^2}\sum_{t=1}^{T}\frac{(\gamma_0-\gamma)^2}{1+\gamma^2}\mathbb{E}\left(\alpha_t^o-\sum_{\tau=1}^{T}\frac{\alpha_\tau^o}{T}\right)^2+\frac{1}{T}\frac{1}{1+\gamma^2}\sum_{t=1}^{T}\mathbb{E}\left(\epsilon_{0t}^o-\gamma\epsilon_{tt}^o-\sum_{\tau=1}^{T}\frac{\epsilon_{0\tau}^o-\gamma\epsilon_{\tau\tau}^o}{T}\right)^2,$$

where the cross-products are 0 by assumption (ii). Notice that both terms are squared and non-negative. The first term is minimized at 0 when $\gamma = \gamma_0$. The expectation in the second term can be expressed as

$$\begin{aligned}
&\mathbb{E}\left[\left(\epsilon^o_{0t} - \gamma\epsilon^o_{tt} - \sum_{\tau=1}^{T}\frac{\epsilon^o_{0\tau} - \gamma\epsilon^o_{\tau\tau}}{T}\right)\right]^2 \\
&= \mathbb{E}\left[\left(\left(1-\frac{1}{T}\right)\epsilon_{0t} - \gamma\left(1-\frac{1}{T}\right)\epsilon_{tt} - \sum_{\tau=1\neq t}^{T}\frac{\epsilon_{0\tau}-\gamma\epsilon_{\tau\tau}}{T}\right)^2\right] \\
&= \mathbb{E}\Bigg[\left(1-\frac{1}{T}\right)^2\epsilon_{0t}^2 + \gamma^2\left(1-\frac{1}{T}\right)^2\epsilon_{tt}^2 + \frac{1}{T^2}\Bigg(\sum_{\tau=1\neq t}^{T}\epsilon_{0\tau}-\gamma\epsilon_{\tau\tau}\Bigg)^2 - 2\gamma\left(1-\frac{1}{T}\right)^2\epsilon_{0t}\epsilon_{tt} \\
&+ 2\gamma\left(1-\frac{1}{T}\right)\epsilon_{tt}\sum_{\tau=1\neq t}^{T}\frac{\epsilon_{0\tau}-\gamma\epsilon_{\tau\tau}}{T} - 2\left(1-\frac{1}{T}\right)\epsilon_{0t}\sum_{\tau=1\neq t}^{T}\frac{\epsilon_{0\tau}-\gamma\epsilon_{\tau\tau}}{T}\Bigg] \\
&= (1+\gamma^2)\left(1-\frac{1}{T}\right)^2\sigma_\epsilon^2 + \frac{2}{T^2}\Bigg(\sum_{\tau=1\neq t}^{T}\sigma_\epsilon^2 + 2\sum_{\tau,\tau'=1\neq t,\tau\neq\tau'}^{T}\mathbb{E}(\epsilon_{0\tau}\epsilon_{0\tau'})\Bigg) \\
&+ \frac{2\gamma^2}{T^2}\Bigg(\sum_{\tau=1\neq t}^{T}\sigma_\epsilon^2 + 2\sum_{\tau,\tau'=1\neq t,\tau\neq\tau'}^{T}\mathbb{E}(\epsilon_{\tau\tau}\epsilon_{\tau\tau'})\Bigg) - \frac{2}{T}\left(1-\frac{1}{T}\right)\sum_{\tau=1\neq t}^{T}\mathbb{E}(\epsilon_{0t}\epsilon_{0\tau})
\end{aligned}$$

Note that

$$\left|\sum_{\tau=1\neq t}^{T}\mathbb{E}(\epsilon_{0t}\epsilon_{0\tau})\right| \leq \left|\sum_{\tau,\tau'=1\neq t,\tau\neq\tau'}^{T}\mathbb{E}(\epsilon_{0\tau}\epsilon_{0\tau'})\right| \leq 2\sum_{k=1}^{T-1}|\rho^k\sigma_\epsilon^2| \leq 2\sigma_\epsilon^2\frac{1-|\rho|^T}{1-|\rho|} = O(1),$$

even as $T \to \infty$ since $|\rho| < 1$. The same can be said for the covariances of $\epsilon_{\tau\tau}$ since they are equivalent to that of $\epsilon_{0\tau}$.

We can thus show that the whole second term can be written as

$$\left(\frac{T-1}{T}\right)^2\sigma_\epsilon^2 + \frac{2}{T^2}\Bigg(\sum_{\tau=1\neq t}^{T}\sigma_\epsilon^2 + 2\sum_{\tau,\tau'=1\neq t,\tau\neq\tau'}^{T}\mathbb{E}(\epsilon_{0\tau}\epsilon_{0\tau'})\Bigg) - \frac{2}{T^2(1+\gamma^2)}\left(1-\frac{1}{T}\right)\sum_{t=1}^{T}\sum_{\tau=1\neq t}^{T}\mathbb{E}(\epsilon_{0t}\epsilon_{0\tau}).$$

Since the final term above is a function of γ, it is not guaranteed that setting $\gamma = \gamma_0$ minimizes the objective function and hence may be biased as shown in Case 3. However, when we take the limit as $T \to \infty$, this term is $O(1/T)$ and hence goes to 0. In addition, the second term is $O(1/T)$ and also approaches 0, and thus the whole term reduces to σ_ϵ^2 in the limit.

This means that, in the limit as $T \to \infty$, we only have to consider the first term of $p\lim_{N\to\infty} q_{N,T}(\gamma, y)$ in optimizing γ, and hence the solution is $\gamma = \gamma_0$.

To have convergence for our estimator, we require uniform convergence of the objective function $q_{N,T}(\gamma, y)$ to $q(\gamma, y)$. As in the previous case (and as in AFGK), we only need to show that for all

$\gamma \in \Gamma$,

$$\left| \frac{1}{T} \sum_{t=1}^{T} \frac{\left(\left(y_{0t} - \frac{\sum_{\tau=1}^{T} (y_{0\tau} - \gamma y_{\tau\tau})}{T(1-\gamma^2)} \right) - \gamma \left(y_{tt} - \gamma \frac{\sum_{\tau=1}^{T} (y_{0\tau} - \gamma y_{\tau\tau})}{T(1-\gamma^2)} \right) \right)^2}{1+\gamma^2} \right| \leq b(\gamma)$$

where b is a non-negative function such that $E\left(b\left(\gamma\right)\right) < \infty$. This can be shown using repeatedly the triangular inequality.

Appendix D. Details about Connectivity Weights

This section provides details about the connectivity weights used in the empirical analysis.

Input-output weights allow for examination of the extent to which spillovers operate through the flow of goods. Stronger input-output linkages may facilitate knowledge transfer about production practices or demand conditions. We build input-output weights using the Basic Price version of the 2010 4-digit NAICS Statistics Canada input-output table (Statistics Canada, 2015). As in Ellison, Glaeser and Kerr (2010), underlying continuous weights are the maximum of upstream and downstream input and output shares:

$$w_{ij}^{\text{IOC}} = \max[\text{Input}_{k(i),k(j)}, \text{Input}_{k(j),k(i)}, \text{Output}_{k(i),k(j)}, \text{Output}_{k(j),k(i)}].$$

We also construct separate weights using each component of w_{ij}^{IOC}. These produce similar results.

Occupational similarity weights allow for examination the extent to which knowledge transfer that is specific to particular occupations is an important driver of firm spillovers.[2] We view results using these weights as informative about the extent to which industries with more similar occupational mixes have more productive knowledge flows. Closer occupational similarity with peers could mean that workers learn more about how to effectively perform their core occupational tasks, where such knowledge transfer may happen through chance encounters (Atkin, Chen and Popov, 2022). We build occupational similarity measures using the 2002 US National Industry Occupation Employment Matrix, which is built using data from the Occupational Employment Statistics survey conducted by the Bureau of Labor Statistics (Bureau of Labor Statistics, 2008). For each industry, it gives the share of employees in each four-digit occupation. Similar to Ellison, Glaeser and Kerr (2010), we define occupational similarity weights as:

$$w_{ij}^{\text{OCCSIM}} = \max[\text{Corr}(\text{Occ. Share}_{k(i)}, \text{Occ. Share}_{k(j)}), 0].$$

Worker flows weights similarly capture the extent to which workers in firm i's industry are likely to have either direct experience working in peers' industries or to use a similar set of skills in their jobs. Seeing a high rate of worker flows from peers' industries is an indicator of closer connections in one or both of these dimensions. We build information on the prevalence of inter-industry worker flows by using the employer-employee match component of our data set. Using all employees in

[2] Ellison, Glaeser and Kerr (2010) interpret greater coagglomeration of firms in occupationally similar industries in local labor markets as reflecting labor market pooling. Their interpretation is likely to be less relevant at the small spatial scale of spillovers that we examine in this paper.

Canada earning at least CAD 5,000 that had different employers in 2001 and 2002, we calculate the share of worker flows from firms in each industry k' that go to each other industry k, adjusting for the share of industry k' in total employment. In particular,

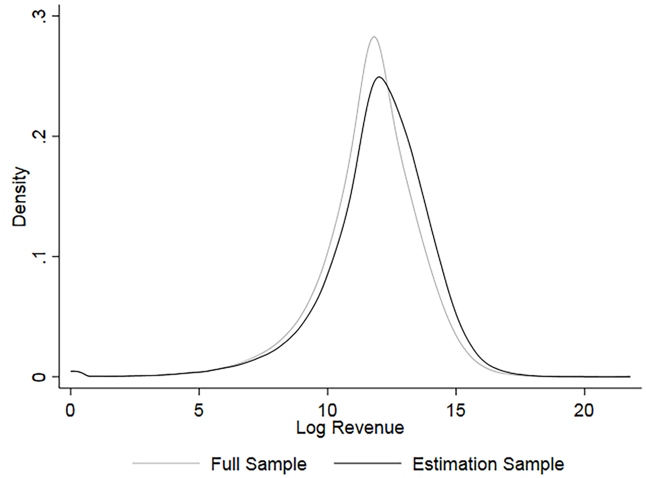
$$w_{ij}^{\text{WFLOW}} = \frac{\text{fraction of industry job changers to industry } k(i) \text{ that are from } k'(j)}{\text{fraction of total job changers from industry } k'(j)}.$$

The denominator accounts for the fact that random choices out of industries with greater worker shares and/or mobility rates would mechanically generate greater flows to all other industries. Therefore, w_{ij}^{WFLOW} measures the extent to which worker flows from industry $k'(j)$ to industry $k(i)$ are greater or less than expected relative to random destination industry choices, taking transitions out of industry $k'(j)$ as given.

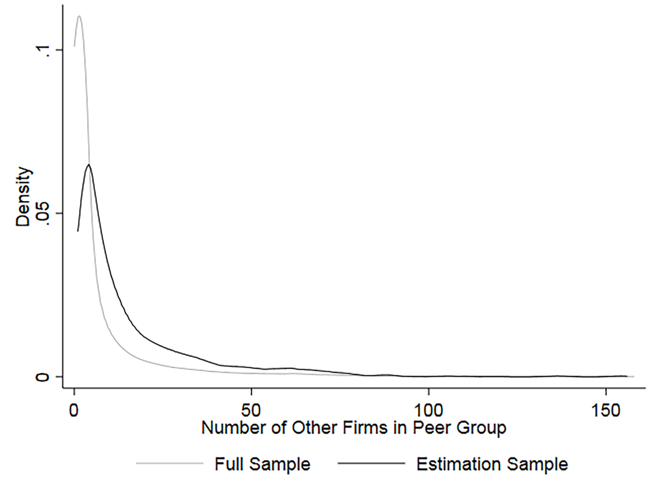
Finally, similar to Greenstone, Hornbeck and Moretti (2010), we also test whether firms in the same 2-digit industry generate differential spillovers to those in other 2-digit industries. In this case, $w_{ij}^{\text{SAME}} = 1$ if $k(i) = k(j)$ at the 2-digit NAICS level and 0 otherwise. Rather than using terciles, we implement this weight in the empirical work by examining impacts of having a higher fraction of peers in the same industry.

Appendix E. Supplemental Figures and Tables

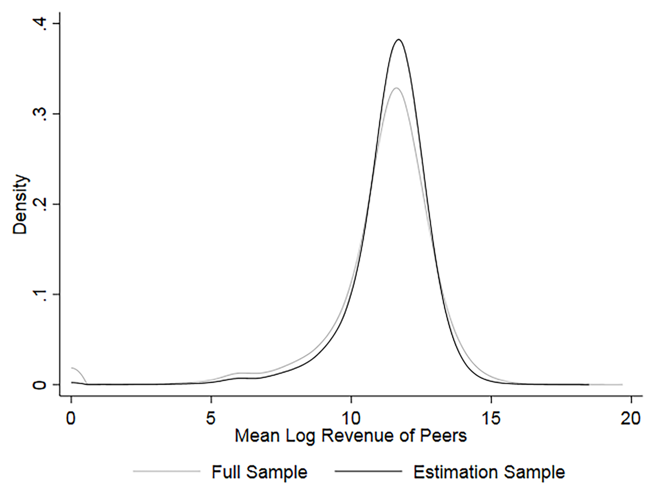
FIGURE E1. – Descriptive Graphs



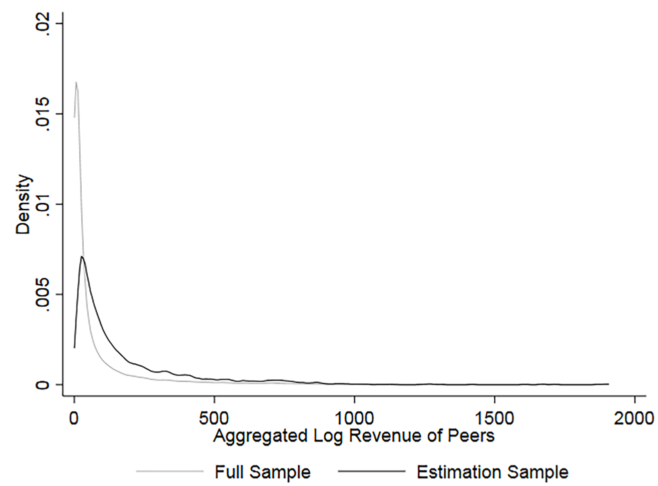
(a) Distribution of Firm Log Revenue



(b) Distribution of Peer Group Size



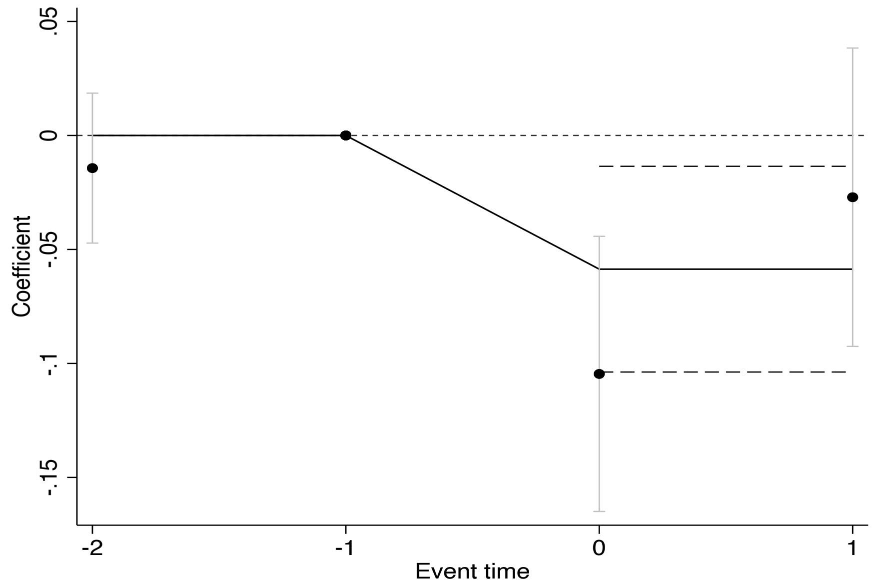
(c) Distribution of Mean Log Revenue in Peer Groups



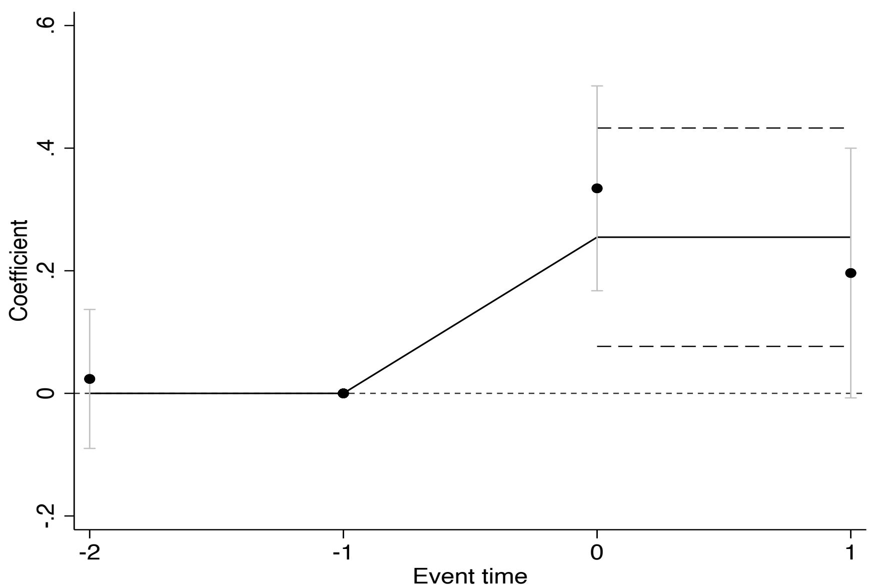
(d) Distribution of Aggregate Log Revenue in Peer Groups

Notes: Plots are for all single-location high-skilled services firms in Montreal, Toronto, and Vancouver census metropolitan areas. The estimation sample excludes firms in peer group areas with one or more member postal code with an area that is greater than $\pi 75^2$sq meters (0.018 sq km) and peer group areas with fewer than two high-skilled services firms in any year 2001-2012.

Figure E2. – Changes in Log Revenue Induced by Firm Entry



(a) Negative shock: firm entry causes a change in average peer quality below 10th percentile



(b) Positive shock: firm entry causes a change in average peer quality above 90th percentile

Notes: Figures show coefficients and confidence intervals from event-study regressions. The dependent variable is firm log revenue residualized for estimated fixed effects in Table 2 column (1) Panel C. Dots are coefficients on yearly event-time dummies, normalizing the coefficient on event-time -1 to 0. Horizontal solid lines are coefficients on biyearly event-time dummies, normalizing the coefficient on pooled event-time -1 and -2 to 0. Whiskers and dashed lines show associated 90% confidence intervals calculated using standard errors clustered at the event level. Estimation samples only include incumbent firms in peer groups in which all incumbents experience a change in average peer quality that is below the 10^{th} percentile (Panel A) or above the 90^{th} percentile (Panel B) of the change in average peer quality distribution across all firm-year observations in the primary estimation sample. Only such events that are induced by the arrival of new firms in a peer group location with no other changes in peer composition up to two years prior and one year after the event are included. Panel A has 93 events, 220 incumbents, and 2200 observations. Panel B has 25 events, 60 incumbents, and 560 observations.

Table E1. – Coefficient Stability Around Large Events

Event time	$t-2$	$t-1$	t	$t+1$	$t+2$
	(1)	(2)	(3)	(4)	(5)
Change in average peer quality < 10th percentile					
Avg. Peer Firm F.E.	0.0214	0.0250	0.0215	0.0367	0.0107
	(0.0063)	(0.0057)	(0.0060)	(0.0068)	(0.0063)
$\bar{X}$	0.06	0.17	-0.56	-0.39	-0.30
Number of Obs.	16,100	19,600	19,500	16,300	12,700
Change in average peer quality < 25th percentile					
Avg. Peer Firm F.E.	0.0262	0.0228	0.0167	0.0236	0.0158
	(0.0048)	(0.0042)	(0.0043)	(0.0055)	(0.0050)
$\bar{X}$	0.05	0.13	-0.30	-0.19	-0.14
Number of Obs.	36,700	44,800	44,500	36,900	28,200
Change in average peer quality > 75th percentile					
Avg. Peer Firm F.E.	0.0243	0.0241	0.0162	0.0225	0.0196
	(0.0047)	(0.0042)	(0.0037)	(0.0044)	(0.0048)
$\bar{X}$	-0.30	-0.33	0.16	0.16	0.14
Number of Obs.	40,400	49,200	49,100	40,300	30,200
Change in average peer quality > 90th percentile					
Avg. Peer Firm F.E.	0.0206	0.0272	0.0085	0.0170	0.0168
	(0.0075)	(0.0069)	(0.0075)	(0.0063)	(0.0071)
$\bar{X}$	-0.54	-0.63	0.19	0.18	0.13
Number of Obs.	16,900	20,600	20,500	16,900	12,700

Notes: Each entry is from a separate regression analogous to that in Table 2 column (1) Panel C but using post-estimation data and different sub-samples. All firms exposed to changes in average peer quality of the amount indicated in each panel are assigned an event year. Regressions of firm log revenue residualized for estimated fixed effects in Table 2 column (1) Panel C on average peer quality, aggregate peer quality, and number of peers are run separately by event time. $\bar{X}$ is the average of average peer quality in each estimation sample. Standard errors are clustered by peer group area.

Table E2. – More Information About Peer Groups

	Peer Group Area Radius			
	75m	150m	200m	250m
Panel A: Average and SD Across Firm-Years				
# of Peers	15.95	28.26	36.34	45.13
	(19.55)	(42.15)	(53.43)	(64.72)
Area (sq. km)	0.043	0.100	0.155	0.216
	(0.115)	(0.216)	(0.340)	(0.403)
Panel B: Average and SD Across Firms				
# of Peer Groups Experienced	1.45	1.44	1.44	1.43
	(0.72)	(0.71)	(0.71)	(0.70)

Notes: Averages and standard deviations (in parentheses) are for single-location high-skilled services firms in the primary estimation sample. The sample excludes firms in peer group areas with one or more member postal code with an area that is greater than $\pi 75^2$sq meters (0.018 sq km) and peer group areas with fewer than two high-skilled services firms in any year 2001-2012. The sample only includes firms in the Montreal, Toronto, or Vancouver census metropolitan areas. Statistics in Panel A are calculated using all firm-year observations. Statistics in Panel B are calculated using one observation per firm.

TABLE E3. – Aggregate Impacts of Counterfactual Firm Allocation Across Peer Groups, Het. Treatment

Randomization Type	Fixed Group Size		Equal Group Size	
Nature of Spillovers Considered	LIM + HET	+ AGG	LIM + HET	+ AGG
	(1)	(2)	(3)	(4)
Estimates w/ Area × Year F.E.,	-0.0034	-0.0029	-0.0037	-0.0071
Randomized Within Areas	(0.0006)	(0.0008)	(0.0006)	(0.0004)
Estimates w/o Area × Year F.E.,	-0.0129	-0.0170	-0.0134	-0.0238
Randomized Within Areas	(0.0007)	(0.0016)	(0.0006)	(0.0011)
Estimates w/o Area × Year F.E.,	-0.0094	-0.0231	-0.0093	-0.0271
Randomized Across All Locations	(0.0013)	(0.0015)	(0.0013)	(0.0015)

Notes: Table presents the means and standard deviations of changes in aggregate revenue that would ensue under 100 simulations of various scenarios in which sorting of firms across peer groups is eliminated. Results in the two columns under the header "Fixed Group Size" are generated holding peer group size fixed and those under the header "Equal Group Size" are generated given full randomization of firms across peer groups. In each column headed by LIM + HET, counterfactual firm revenue absent sorting is calculated adjusting for the linear-in-means component of the spillover as well as the fraction of peers in the top tercile of the local 500-meter radius area's firm quality distribution, using coefficients from Table 6, column 6. In each column headed by +AGG, the same two terms plus the agglomeration term are included in the calculation, again using coefficients from Table 6, column 6. The first row uses fixed effects estimates from Table 2 column (1) Panel C and imposes demeaning and randomization across peer groups within 500-meter radius areas. The second row uses fixed effects estimates from Table 2 column (2) Panel C instead with the same demeaning and randomization procedures. The third row uses fixed effects estimates from Table 2 column (2) Panel C but demeans and randomizes across all peer groups.

*

REFERENCES

Arcidiacono, Peter, Gigi Foster, Natalie Goodpaster, and Josh Kinsler. 2012. "Estimating Spillovers Using Panel Data, with an Application to the Classroom." *Quantitative Economics*, 3(3): 421–470.

Atkin, David, M. Keith Chen, and Anton Popov. 2022. "The Returns to Face-to-Face Interactions: Knowledge Spillovers in Silicon Valley." National Bureau of Economic Research Working Paper 30147.

Bank of Canada. 2019. "Chartered Bank Administered Interest Rates - Prime Business." URL: https://www.banqueducanada.ca/wp-content/uploads/2010/09/selected_historical_v122495.pdf. Accessed: January 31, 2019.

Bureau of Labor Statistics. 2008. "Occupational Employment Statistics (OES) Survey, 2002." National industry-specific estimates. URL: https://www.bls.gov/oes/tables.htm#2002. Accessed: June 6, 2018.

De Loecker, Jan. 2011. "Product Differentiation, Multiproduct Firms, and Estimating the Impact of Trade Liberalization on Productivity." *Econometrica*, 79(5): 1407–1451.

De Loecker, Jan, and Jan Eeckhout. 2018. "Global Market Power." CEPR Discussion Paper 13009.

Ellison, Glenn, Edward L. Glaeser, and William R. Kerr. 2010. "What Causes Industry Agglomeration? Evidence from Coagglomeration Patterns." *American Economic Review*, 100(3): 1195–1213.

Greenstone, Michael, Richard Hornbeck, and Enrico Moretti. 2010. "Identifying Agglomeration Spillovers: Evidence from Winners and Losers of Large Plant Openings." *Journal of Political Economy*, 118(3): 536–598.

Statistics Canada. 2015. "Symmetric Input-Output Tables, 2010." Catalogue Number 15-207-X. https://www150.statcan.gc.ca/n1/en/catalogue/15-207-X2013001. Accessed: July 6, 2018.

Statistics Canada. 2019*a*. " Flows and stocks of fixed non-residential capital, by industry and type of asset, Canada, provinces and territories (x 1,000,000)." Table 36-10-0096-01. DOI: https://doi.org/10.25318/3610009601-eng. Accessed: January 9, 2019.

Statistics Canada. 2019*b*. "Consumer Price Index, annual average, not seasonally adjusted." Table 18-10-0005-01. DOI: https://doi.org/10.25318/1810000501-eng. Accessed: February 1, 2019.

Teranet-National Bank. 2019. "House Price Index." 1990-2018. https://housepriceindex.ca/. Accessed: February 1, 2019.

Wooldridge, Jeffrey M. 2010. *Econometric Analysis of Cross Section and Panel Data.* The MIT Press.