Online Appendix for "Location Choices of Multi-plant Oligopolists: Theory and Evidence from the Cement Industry" *by* Chenying Yang

These appendices derive results and provide further details of the paper. Equation numbers refer to those in the main text.

A Model details

For simplicity of derivation, I invert the marginal cost and derive the following equation based on a plant's cost-adjusted productivity

$$\tilde{Z}_{f\ell im} = \frac{Z_{f\ell i}}{w_\ell \tau_{\ell m}}.$$

A.1 Conditional joint distribution of the two lowest cost firms

Conditional on firm f^*'s plant in ℓ^* supplying to a consumer i in market m, i.e., $\tilde{Z}_{1,i(m)} = \tilde{Z}_{f^*\ell^* im}$ and $\tilde{Z}_{2,i(m)} = \max_{f \neq f^*, f \in \mathcal{F}} \{\tilde{Z}_{1,fi(m)}\}$, the joint distribution of the first and second highest firm-level productivity is

$$\begin{aligned} F_{12,m|f^*\ell^*}(z_1, z_2) &= \Pr\left(\tilde{Z}_{1,i(m)} \leq z_1, \tilde{Z}_{2,i(m)} \leq z_2 \mid \tilde{Z}_{1,i(m)} > V\right) \\ &= \Pr\left(\tilde{Z}_{1,i(m)} \leq z_2 \mid \tilde{Z}_{1,i(m)} > V\right) \\ &\quad + \Pr\left(z_2 \leq \tilde{Z}_{1,i(m)} \leq z_1, \tilde{Z}_{2,i(m)} \leq z_2 \mid \tilde{Z}_{1,i(m)} > V\right), \end{aligned} \tag{A-1}$$

where $V \equiv \max\{\tilde{Z}_{2,i(m)}, S\}$ and $S \equiv \max_{\ell \in \mathcal{L}_{f^*}, \ell \neq \ell^*} \{\tilde{Z}_{f^*\ell im}\}$, given $z_1 > z_2$.

The distribution of S is

$$F^S_m(s) = \exp\left(-\left(\Phi_{f^*m} - \phi_{\ell^*m}\right) s^{-\theta}\right),$$

and the distribution of V is

$$F^V_m(\nu) = \exp\left(-\left(\Phi_m - \phi_{\ell^*m}\right) \nu^{-\theta}\right).$$

The first part of equation (A-1) can be simplified as

$$\begin{aligned}
\Pr\left(\tilde{Z}_{1,i(m)} \leq z_2 \mid \tilde{Z}_{1,i(m)} > V\right) &= \frac{\Pr\left(V < \tilde{Z}_{f^*\ell^* im} \leq z_2\right)}{s_{f^*\ell^* m}} \\
&= \frac{\Phi_m}{\phi_{\ell^* m}} \int_0^{z_2} \left[\tilde{F}^{draw}_{\ell^* m}(z_2) - \tilde{F}^{draw}_{\ell^* m}(V)\right] dF^V_m(V) \\
&= \exp\left(-\Phi_m z_2^{-\theta}\right),
\end{aligned} \tag{A-2}$$

where $\tilde{F}^{draw}_{\ell^* m}(z) = \exp\left(-\phi_{\ell^* m} z^{-\theta}\right)$, and the probability of sourcing $s_{f^*\ell^* m} = \phi_{\ell^* m}/\Phi_m$ derived from equation (10) and the setting where plants at the same location are identical regardless of the firm it belongs.

Next, the second part of equation (A-1) is equal to

$$\begin{aligned}
&\frac{\Pr\left(z_2 \leq \tilde{Z}_{f^*\ell^* im} \leq z_1, \tilde{Z}_{2,i(m)} \leq z_2, \tilde{Z}_{f^*\ell^* im} > V\right)}{s_{f^*\ell^* m}} \\
&= \frac{\Pr\left(z_2 \leq \tilde{Z}_{f^*\ell^* im} \leq z_1, \tilde{Z}_{2,i(m)} \leq z_2, \tilde{Z}_{f^*\ell^* im} > S\right)}{s_{f^*\ell^* m}},
\end{aligned}$$

based on the definition of $\tilde{Z}_{2,i(m)}$. The numerator can be further simplified as

$$\begin{aligned}
&\Pr\left(z_2 \leq \tilde{Z}_{f^*\ell^* im} \leq z_1, \tilde{Z}_{2,i(m)} \leq z_2, \tilde{Z}_{f^*\ell^* im} > S\right) \\
&= \Pr\left(z_2 \leq S \leq \tilde{Z}_{f^*\ell^* im} \leq z_1, \tilde{Z}_{2,i(m)} \leq z_2\right) + \Pr\left(S \leq z_2 \leq \tilde{Z}_{f^*\ell^* im} \leq z_1, \tilde{Z}_{2,i(m)} \leq z_2\right) \\
&= \int_{z_2}^{z_1} \left[\tilde{F}^{draw}_{\ell^* m}(z_1) - \tilde{F}^{draw}_{\ell^* m}(S)\right] \prod_{f \neq f^*} \tilde{F}_{1,fm}(z_2) dF^S_m(S) \\
&\quad + \int_0^{z_2} \left[\tilde{F}^{draw}_{\ell^* m}(z_1) - \tilde{F}^{draw}_{\ell^* m}(z_2)\right] \prod_{f \neq f^*} \tilde{F}_{1,fm}(z_2) dF^S_m(S) \\
&= \frac{\phi_{\ell^* m}}{\Phi_{f^* m}} \left(e^{-(\Phi_m - \Phi_{f^* m}) z_2^{-\theta}} e^{-\Phi_{f^* m} z_1^{-\theta}} - e^{-\Phi_m z_2^{-\theta}}\right),
\end{aligned}$$

where $\tilde{F}_{1,fm}(z) = \exp\left(-\Phi_{fm} z^{-\theta}\right)$.

The second part of equation (A-1) is therefore

$$\Pr\left(z_2 \leq \tilde{Z}_{1,i(m)} \leq z_1, \tilde{Z}_{2,i(m)} \leq z_2 \mid \tilde{Z}_{1,i(m)} > V\right) = \frac{\Phi_m}{\Phi_{f^* m}} \left(e^{-(\Phi_m - \Phi_{f^* m}) z_2^{-\theta}} e^{-\Phi_{f^* m} z_1^{-\theta}} - e^{-\Phi_m z_2^{-\theta}}\right). \tag{A-3}$$

By summing equations (A-2) and (A-3), the joint distribution of highest two cost-adjusted productivities conditional on f^* from ℓ^* selling to i in m is

$$F_{12,m|f^*\ell^*}(z_1, z_2) = \frac{\Phi_m}{\Phi_{f^*m}} e^{-(\Phi_m - \Phi_{f^*m}) z_2^{-\theta}} e^{-\Phi_{f^*m} z_1^{-\theta}} - \frac{\Phi_m - \Phi_{f^*m}}{\Phi_{f^*m}} e^{-\Phi_m z_2^{-\theta}}.$$

The associated p.d.f. is

$$f_{12,m}(z_1, z_2) = \Phi_m(\Phi_m - \Phi_{f^*m})\theta^2 z_1^{-\theta-1} z_2^{-\theta-1} e^{-(\Phi_m - \Phi_{f^*m}) z_2^{-\theta}} e^{-\Phi_{f^*m} z_1^{-\theta}}.$$

Note that the conditional joint productivity distribution is invariant to ℓ for the same firm, $F_{12,m|f^*} = F_{12,m|f^*\ell^*}$. The conditional joint cost distribution is given by

$$F^c_{12,m|f^*}(c_1, c_2) = 1 - F_{12,m|f^*}(\infty, c_2^{-1}) - F_{12,m|f^*}(c_1^{-1}, c_1^{-1}) + F_{12,m|f^*}(c_1^{-1}, c_2^{-1}),$$

which results in equation (5).

A.2 Price distribution

The price of consumer i in market m is

$$P_{i(m)} = \min\{\frac{1}{\tilde{Z}_{2,i(m)}}, \frac{\bar{\mu}}{\tilde{Z}_{1,i(m)}}\}.$$

Conditional on firm f^* serving the consumer in the market, the complement of the price c.d.f. is

$$\begin{aligned} 1 - F^p_{m|f^*}(p) = & \underbrace{\Pr\left(p \leq \frac{1}{\tilde{Z}_{2,i(m)}} < \frac{\bar{\mu}}{\tilde{Z}_{1,i(m)}} \mid \tilde{Z}_{1,i(m)} > V\right)}_{\text{T1}} \\ & + \underbrace{\Pr\left(p \leq \frac{\bar{\mu}}{\tilde{Z}_{1,i(m)}} \leq \frac{1}{\tilde{Z}_{2,i(m)}} \mid \tilde{Z}_{1,i(m)} > V\right)}_{\text{T2}}. \end{aligned}$$

By deriving each component, I have the firm term

$$\begin{aligned} \text{T1} &= \int_{p^{-1}}^{\infty} \int_{z_1/\bar{\mu}}^{p^{-1}} f_{12,m} dz_2 dz_1 + \int_0^{p^{-1}} \int_{z_1/\bar{\mu}}^{z_1} f_{12,m} dz_2 dz_1 \\ &= \frac{\Phi_m}{\Phi_{f^*m}} e^{-(\Phi_m - \Phi_{f^*m}) p^{\theta}} - \frac{\Phi_m - \Phi_{f^*m}}{\Phi_{f^*m}} e^{-\Phi_m p^{\theta}} - \frac{\Phi_m}{\Phi_{f^*m} + (\Phi_m - \Phi_{f^*m})\bar{\mu}^{\theta}}, \end{aligned}$$

and the second term

$$
\begin{aligned}
\text{T2} &= \int_0^\infty \int_{\bar{\mu}z_2}^{\bar{\mu}/p} f_{12,m} dz_1 dz_2 \\
&= \frac{\Phi_m}{\Phi_{f^*m}} e^{-\Phi_{f^*m}\bar{\mu}^{-\theta}p^\theta} - \frac{\Phi_m/\Phi_{f^*m}(\Phi_m - \Phi_{f^*m})\bar{\mu}^\theta}{\Phi_{f^*m} + (\Phi_m - \Phi_{f^*m})\bar{\mu}^\theta}.
\end{aligned}
$$

Combining T1 and T2 and subtracting from one, I get the price distribution exactly equals equation (6).

Before calculating the expected price, it is useful to state a Lemma

Lemma 1. *For* $\omega > 0$ *and* $\theta + 1 > 0$*,*

$$
\int_0^\infty z^{-\theta-2} e^{-\omega z^{-\theta}} dz = \omega^{\frac{-1-\theta}{\theta}} \frac{1}{\theta} \Gamma\left(\frac{\theta+1}{\theta}\right)
$$

The expected price is

$$
E\left[P_{m|f^*}\right] = \underbrace{\int_0^\infty \int_{\frac{z_1}{\bar{\mu}}}^{z_1} \left(\frac{1}{z_2}\right) f_{12,m}(z_1, z_2) dz_2 dz_1}_{\text{T1}} + \underbrace{\int_0^\infty \int_0^{\frac{z_1}{\bar{\mu}}} \left(\frac{\bar{\mu}}{z_1}\right) f_{12,m}(z_1, z_2) dz_2 dz_1}_{\text{T2}}.
$$

By changing the order of integration and applying Lemma 1, the first term is simplified to

$$
\text{T1} = \frac{\Phi_m}{\Phi_{f^*m}} \left(\Phi_m - \Phi_{f^*m}\right) \Gamma\left(\frac{\theta+1}{\theta}\right) \left[\left(\Phi_m - (1 - \bar{\mu}^{-\theta})\Phi_{f^*m}\right)^{\frac{-1-\theta}{\theta}} - \Phi_m^{\frac{-1-\theta}{\theta}} \right],
$$

and the second term is

$$
\text{T2} = \bar{\mu}^{-\theta} \Phi_m \Gamma\left(\frac{\theta+1}{\theta}\right) \left(\Phi_m - (1 - \bar{\mu}^{-\theta})\Phi_{f^*m}\right)^{\frac{-1-\theta}{\theta}}.
$$

Combining the two terms, the expected price is derived as in equation (7).

If a firm always prices against the second-lowest cost, its own cost reduction due to building additional plants has no impact on the firm's prices. However, if the monopoly price always prevails, the firm's cost reduction fully pass to its prices. Therefore, intuitively, having more plants weakly decreases the firm's average price. Formally, one can take first-order derivative (FOD) of $E\left[P_{m|f^*}\right]$ with respect to Φ_{f^*m}, and for $\theta > 0$ and $\bar{\mu} > 1$, the FOD is non-positive.

A.3 Markup distribution

Conditional on firm f^* serving the consumer i in market m, for markups below the monopoly level, the distribution is

$$F^{\mu}_{m|f^*}(\mu) = \Pr\left(\frac{\tilde{Z}_{1,i(m)}}{\tilde{Z}_{2,i(m)}} \leq \mu \mid \tilde{Z}_{1,i(m)} > V\right).$$

I first calculate the complement of the c.d.f.,

$$\begin{aligned}
1 - F^{\mu}_{m|f^*}(\mu) &= \Pr\left(\tilde{Z}_{2,i(m)} \leq \mu^{-1}\tilde{Z}_{1,i(m)} \mid \tilde{Z}_{1,i(m)} > V\right) \\
&= \int_0^{\infty}\int_0^{\mu^{-1}z_1} f_{12,m}(z_1, z_2)dz_2dz_1 \\
&= \frac{\Phi_m}{\Phi_{f^*m} + (\Phi_m - \Phi_{f^*m})\mu^{\theta}}.
\end{aligned}$$

The conditional markup distribution is then

$$F^{\mu}_{m|f^*}(\mu) = 1 - \frac{1}{\mu^{\theta} - \frac{\Phi_{f^*m}}{\Phi_m}(\mu^{\theta} - 1)} = 1 - \frac{1}{(1 - s_{f^*m})\mu^{\theta} + s_{f^*m}}.$$

Given the markup $\mu \in (1, \infty)$, it is straightforward that $\lim_{\mu\to 1} F^{\mu}_{m|f^*}(\mu) = 0$ and $\lim_{\mu\to\infty} F^{\mu}_{m|f^*}(\mu) = 1$. With additional plants built by firm f^*, s_{f^*m} increases, and the value of c.d.f. decreases which implies the firm's markup distribution is shifted in a first-order stochastic dominance sense, leading to an increase in the firm's expected markup.

Moreover, the probability of a firm earning monopoly markup increases with its number of plants. Given the second-lowest cost z_2 and firm f^* supplying the consumer,

$$\begin{aligned}
F^{\mu}_{m|f^*,z_2}(\mu) &= \Pr\left(\tilde{Z}_{2,i(m)} \leq \tilde{Z}_{1,i(m)} \leq \mu\tilde{Z}_{2,i(m)} \mid \tilde{Z}_{2,i(m)} = z_2\right) \\
&= \frac{\int_{z_2}^{\mu z_2} f_{12,m}(z_1, z_2)dz_1}{\int_{z_2}^{\infty} f_{12,m}(z_1, z_2)dz_1} = \frac{e^{-\Phi_{f^*m}(\mu z_2)^{-\theta}} - e^{-\Phi_{f^*m}z_2^{-\theta}}}{1 - e^{-\Phi_{f^*m}z_2^{-\theta}}}.
\end{aligned}$$

Therefore, the probability of firm f^* charging a monopoly markup is

$$1 - F^{\mu}_{m|f^*,z_2}(\bar{\mu}) = \frac{1 - e^{-\Phi_{f^*m}(\bar{\mu}z_2)^{-\theta}}}{1 - e^{-\Phi_{f^*m}z_2^{-\theta}}}.$$

Taking the first-order derivative with respect to Φ_{f^*m}, the probability of the firm earning a monopoly markup strictly increases with its producing capability Φ_{f^*m}. This implies that when knowing the price otherwise charged is c_2, the firm is more likely to exploit the maximum markup if it has more plants at favorable locations (hence higher Φ_{fm}), to widen her efficiency gap to the next lowest cost rival.

A.4 Expected revenue

Conditional on firm f^*, its expected revenue of cement sold to destination market m is

$$E\big[R_{m|f^*}\big] = A_m E\big[P_{m|f^*}^{1-\eta}\big],$$

The expectation is taken with respect to the random price realization. The derivation resembles the conditional expected price's derivation in Section A.2, and

$$\begin{aligned} E\big[P_{m|f^*}^{1-\eta}\big] &= \int_0^\infty \int_{\frac{z_1}{\bar{\mu}}}^{z_1} \left(\frac{1}{z_2}\right)^{1-\eta} f_{12,m}(z_1,z_2)dz_2dz_1 + \int_0^\infty \int_0^{\frac{z_1}{\bar{\mu}}} \left(\frac{\bar{\mu}}{z_1}\right)^{1-\eta} f_{12,m}(z_1,z_2)dz_2dz_1. \\ &= \underbrace{\Gamma\left(\frac{\theta+1-\eta}{\theta}\right)}_{\kappa} \underbrace{\frac{\Phi_m}{\Phi_{f^*m}}}_{1/s_{f^*m}} \left(\underbrace{\left(\Phi_m - (1-\bar{\mu}^{-\theta})\Phi_{f^*m}\right)^{-\frac{1-\eta}{\theta}} - \left(\Phi_m - \Phi_{f^*m}\right)\Phi_m^{-\frac{\theta+1-\eta}{\theta}}}_{\bar{R}_{f^*m}} \right), \end{aligned}$$

for $\theta+1-\eta>0$. The unconditional expected revenue is therefore $E\big[R_{f^*m}\big] = \kappa A_m \bar{R}_{f^*m}$.

A.5 Expected cost

Conditional on firm f^*, its expected total variable cost of cement sold to destination market m is

$$\begin{aligned} E\big[C_{m|f^*}\big] &= A_m E\left[\frac{P_{m|f^*}^{1-\eta}}{\mu}\right] \\ &= A_m \left\{ \underbrace{\int_0^\infty \int_{\frac{z_1}{\bar{\mu}}}^{z_1} \left(\frac{1}{z_2}\right)^{1-\eta} \frac{z_2}{z_1} f_{12,m}(z_1,z_2)dz_2dz_1}_{\text{T1}} + \underbrace{\int_0^\infty \int_0^{\frac{z_1}{\bar{\mu}}} \left(\frac{\bar{\mu}}{z_1}\right)^{1-\eta} \frac{1}{\bar{\mu}} f_{12,m}(z_1,z_2)dz_2dz_1}_{\text{T2}} \right\}. \end{aligned}$$

Before further simplification, I provide the following mathematical property,

Lemma 2. *For $\omega>0$ and $2\theta+1-\eta>0$,*

$$\int_0^\infty z^{-2\theta-2+\eta} e^{-\omega z^{-\theta}} dz = \omega^{-\frac{2\theta+1-\eta}{\theta}} \left(\frac{\theta+1-\eta}{\theta^2}\right) \Gamma\left(\frac{\theta+1-\eta}{\theta}\right).$$

Replacing z_1 by μz_2 and changing the order of integration, the first term equals

$$\text{T1} = \Phi_m(\Phi_m - \Phi_{f^*m})(\theta + 1 - \eta)\Gamma\left(\frac{\theta+1-\eta}{\theta}\right)\int_1^{\bar{\mu}} \mu^{-\theta-2}\left(\Phi_m - (1-\mu^{-\theta})\Phi_{f^*m}\right)^{-\frac{2\theta+1-\eta}{\theta}} d\mu.$$

Unfortunately, there is no closed-form expression for the integral. Therefore, I apply the numerical approximation in the empirical section.

Applying Lemma 2, the second term is simplified to

$$\text{T2} = \bar{\mu}^{-\theta-1}\Phi_m\Gamma\left(\frac{\theta+1-\eta}{\theta}\right)\left(\Phi_m - (1-\bar{\mu}^{-\theta})\Phi_{f^*m}\right)^{-\frac{\theta+1-\eta}{\theta}}.$$

Combining the two terms and multiplying by the probability of firm f^* is selected gives the unconditional expected total variable cost $E\left[C_{f^*m}\right] = \kappa A_m \bar{C}_{f^*m}$.

B Model extension: adding core productivity differences at firm level

Suppose a plant's marginal cost of production is affected by its parent firm's endowed core productivity Z_f,

$$C_{f\ell im} = \frac{w_\ell \tau_{\ell m}}{Z_f Z_{\ell i}}, \forall i \in \mathcal{D}_m$$

where $Z_{\ell i}$ are draws from the Fréchet distribution $\exp(-T_\ell z^{-\theta})$. The c.d.f. of the plant's cost-adjusted productivity $\tilde{Z}_{f\ell im} = \frac{Z_{\ell i}}{w_\ell \tau_{\ell m}/Z_f}$ is then

$$\tilde{F}^{draw}_{f\ell m}(z) = \exp(-\phi_{f\ell m} z^{-\theta}),$$

where $\phi_{f\ell m} = Z_f^\theta \phi_{\ell m} = Z_f^\theta T_\ell \left(w_\ell \tau_{\ell m}\right)^{-\theta}$. The distributions of plants' productivities at the same location are shifted by firms' core productivities, although the shape parameter remains the same. Plants owned by an efficient firm are on average more productive than those owned by inefficient firms at the same location. Exploiting the properties of extreme value distribution, the distribution of a firm's highest cost-adjusted productivity in supplying the product to market m is

$$\tilde{F}_{1,fm}(z) = \exp(-\Phi_{fm} z^{-\theta}),$$

where $\Phi_{fm} = \sum_{\ell \in \mathcal{L}_f} \phi_{f\ell m}$. The firm's capability not only depends on plants' spatial setting but also its core productivity. One can complete the model following the same steps but replacing with the new formulation of Φ_{fm}. The model propositions remain to hold.

To estimate this extended version of the model, additional firm-level data is necessary. The gravity model holds at the plant level where $s_{f\ell m} = \frac{\phi_{f\ell m}}{\Phi_m}$ conditional on firm f has a plant at location ℓ, and the estimable form is

$$E\left[\frac{Q_{f\ell m}}{Q_m} \mid \mathbb{I}_{f\ell} = 1\right] = \exp\left[\text{FE}_f + \text{FE}_\ell + \text{FE}_m - \theta \mathbf{X}'_{\ell m}\beta^\tau\right],$$

where $\text{FE}_f = \theta \ln Z_f$ and $\text{FE}_\ell = \ln\left(T_\ell w_\ell^{-\theta}\right)$. Plant-market-level trade flow in volume will be needed to separately identify the location characteristics $T_\ell w_\ell^{-\theta}$ from firm productivities Z_f. Note that the probability of a location exports goods to a market becomes $s_{\ell m} = \frac{\phi_{\ell m} \sum_{f \in \mathcal{F}_\ell} Z_f^\theta}{\Phi_m}$, where $\mathcal{F}_\ell$ indicates the set of firms producing at ℓ. So a regression at bilateral location level is no longer sufficient.

C Estimation details

C.1 Estimation of the trade elasticity

There are three groups of trade flows to consider: across Canada-FAF flow, across US-FAF flow and US-FAF-Canada-FAF flow. For the first group, the cement trade across Canadian FAF zones is directly provided by the Canadian FAF survey. The drawback of using Canadian Freight Analysis Framework is that it is a logistics file built on a carrier survey where the origins and destinations are not necessarily the points of production or final consumption. The US Freight Analysis Framework, on the other hand, is based on the US Commodity Flow Survey (CFS) and collects data on shipments from the point of production to the point of consumption. However, the limitation of using US FAF survey is that the commodities are classified at the 2-digit level of Standard Classification of Transported Goods (SCTG) where cement is a subcategory of nonmetallic mineral products. Other products included in the nonmetallic minerals category are glass, bricks, and ceramic products. To derive US-FAF cement trade, I assume that the cement trade is proportional to nonmetallic mineral trade by the fraction of cement consumed in nonmetallic mineral consumption by destination FAF zone. Because the US Geological Survey only provides cement consumption by state, not by FAF zone, I further assume that the consumption ratio of cement over nonmetallic minerals is the same for every FAF zone within the same state. I also validate that the trade coefficients are not significantly different between cement and nonmetallic minerals using a country-level sample, as shown by the insignificant interaction terms in Table C.1.

Lastly, I need to allocate the cement trade data between Canadian provinces and US states provided by Statistics Canada to that of each FAF zone dyads. A key variable given by the US Commodity Flow Survey is the distance band between origin and destination where there is pos-

Table C.1: Trade estimates for cement and nonmetallic minerals

	Great circle distance	Sea distance	Shipping time
log dist$_{\ell m}$	-2.105^a	-1.255^a	-1.095^a
	(0.090)	(0.051)	(0.068)
log dist$_{\ell m}\times$ industry	-0.032	-0.056	-0.022
	(0.078)	(0.053)	(0.077)
contiguity$_{\ell m}$	1.072^a	1.668^a	1.186^a
	(0.160)	(0.139)	(0.196)
contiguity$_{\ell m}\times$ industry	0.100	0.074	0.082
	(0.184)	(0.171)	(0.222)
language$_{\ell m}$	0.437^a	0.675^a	0.735^a
	(0.143)	(0.133)	(0.143)
language$_{\ell m}\times$ industry	0.083	0.084	0.086
	(0.161)	(0.159)	(0.170)
RTA$_{\ell m}$	0.540^a	0.838^a	0.939^a
	(0.131)	(0.129)	(0.135)
RTA$_{\ell m}\times$ industry	0.237	0.204	0.244
	(0.188)	(0.194)	(0.205)
industry	0.008	0.219	-0.207
	(0.639)	(0.459)	(0.210)
Observations	33842	33842	33842
R^2	0.397	0.398	0.325

The dependent variable is share of export volume. All regressions include origin and destination fixed effects and are performed using PPML. Sample is for 2016 and 144 countries. Trade with own is dropped from the sample since the data are unavailable for the nonmetallic mineral products. Different columns use different measurements of distance. R^2 is the correlation of fitted and true dependent variables. Robust standard errors are in parentheses. Significance levels: ^{c} p<0.1, ^{b} p<0.05, ^{a} p<0.01.

itive cement shipment. Comparing the distance between each US-Canada FAF zone dyads with the distance band and considering the zones with positive cement production, the sample of pairs that are likely to have positive cement trade is reduced substantially. Then, I compute trade in this restricted sample. Trade between each FAF zone pair is derived by apportioning the associated state-province trade by the total export and import of the originating zone and the destination zone. The assumption is that within the same state-province pair, one zone cannot export to a destination more than its nearby zone if its total export is smaller. I acknowledge the restrictiveness of the assumption due to data limitation.

In columns (4) and (5) of Table 1 in the main text, I present the trade elasticity estimates using a country sample. Table C.2 provides alternative specifications and shows that the results are similar using PPML.

Table C.2: Estimation of trade costs at country level

	Great circle distance			Sea distance			Shipping time		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	OLS, $\log Q_{\ell m}$	PPML, $Q_{\ell m}$	PPML, $Q_{\ell m}/Q_m$	OLS, $\log Q_{\ell m}$	PPML, $Q_{\ell m}$	PPML, $Q_{\ell m}/Q_m$	OLS, $\log Q_{\ell m}$	PPML, $Q_{\ell m}$	PPML, $Q_{\ell m}/Q_m$
log (1+ cement tariff$_{\ell m}$)	2.808	-10.980^a	-10.749^a	1.451	-12.635^a	-10.567^a	1.460	-13.648^a	-11.633^a
	(3.716)	(3.248)	(2.736)	(3.712)	(3.475)	(2.590)	(3.787)	(3.441)	(2.711)
log dist$_{\ell m}$	-2.160^a	-1.997^a	-2.083^a	-1.471^a	-1.201^a	-1.359^a	-1.321^a	-1.097^a	-1.067^a
	(0.259)	(0.285)	(0.254)	(0.170)	(0.121)	(0.157)	(0.182)	(0.134)	(0.138)
contiguity$_{\ell m}$	3.916^a	1.685^a	2.255^a	4.005^a	2.286^a	2.740^a	3.609^a	1.693^a	2.617^a
	(0.423)	(0.362)	(0.420)	(0.427)	(0.286)	(0.342)	(0.497)	(0.368)	(0.410)
language$_{\ell m}$	0.296	-0.380	-0.462	0.361	-0.340	-0.449	0.424	-0.377	-0.465
	(0.354)	(0.285)	(0.300)	(0.356)	(0.277)	(0.296)	(0.360)	(0.282)	(0.291)
RTA$_{\ell m}$	0.293	0.972^a	1.224^a	0.801^b	1.231^a	1.559^a	0.829^b	1.261^a	1.738^a
	(0.421)	(0.272)	(0.323)	(0.397)	(0.268)	(0.323)	(0.396)	(0.270)	(0.302)
home$_{\ell m}$	8.306^a	6.323^a	5.893^a	9.823^a	7.895^a	7.456^a	9.547^a	7.394^a	7.749^a
	(0.869)	(0.711)	(0.733)	(0.744)	(0.441)	(0.476)	(0.836)	(0.543)	(0.625)
Observations	1100	20736	20736	1100	20736	20736	1100	20736	20736
R^2	0.719	0.999	0.973	0.715	0.999	0.975	0.709	0.999	0.973

All regressions include origin and destination fixed effects. Regressions use 144 countries' squared sample and for year 2016. R^2 is the correlation of fitted and true dependent variables. Robust standard errors are in parentheses. Significance levels: ^{c} p<0.1, ^{b} p<0.05, ^{a} p<0.01.

C.2 Estimation of demand

Table C.3 presents the first-stage results of the demand estimation using the price survey data. The cost shifters are indeed significantly correlated with cement prices. The F-statistics of the excluded instruments on the endogenous regressor is 21.64, and the Stock-Wright S statistics is 95.59. Both are above the rule-of-thumb threshold of 10. Hence, the tests reject the weak IV concern.

Table C.3: First-stage regression for demand estimation

	log price$_m$
log ($\sum_{\ell \neq m}$natural gas$_\ell / d_{\ell m}$)	0.410^a
	(0.073)
log ($\sum_{\ell \neq m}$electricity$_\ell / d_{\ell m}$)	-0.159
	(0.125)
log ($\sum_{\ell \neq m}$wage$_\ell / d_{\ell m}$)	1.238^a
	(0.146)
log ($\sum_{\ell \neq m}$limestone$_\ell / d_{\ell m}$)	-0.046
	(0.067)
log natural gas$_m$	-0.037^a
	(0.012)
log electricity$_m$	-0.032^c
	(0.017)
log wages$_m$	0.099^a
	(0.031)
log limestone$_m$	0.022^b
	(0.009)
log building permits$_m$	0.025^a
	(0.006)
log population$_m$	-0.038^a
	(0.006)
F test of excluded instruments	21.64
Stock-Wright LM S statistic	95.59
Observations	739

First-stage regression for column (3) in Table 2. Price is from the data based on survey regions and then assigned to the 149 FAF zones. $d_{\ell m}$ is the distance between a location-market pair. The regression includes year fixed effects from 2012 to 2016. Variables other than the number of building permits and population are excluded instruments. Robust standard errors are in parentheses. Significance levels: ^{c} p<0.1, ^{b} p<0.05, ^{a} p<0.01.

C.3 Estimation of fixed costs

To find each firm's optimal plant set given the other firms' plant locations, I adopt the algorithm in Arkolakis and Eckert (2017) as following. Define firm f's marginal profit of including ℓ in a location strategy $\mathcal{L}_f$ as

$$\Delta^{\ell}\Pi_f(\mathcal{L}_f) = \Pi_f(\mathcal{L}_f \cup \ell) - \Pi_f(\mathcal{L}_f \setminus \ell).$$

In the single-player case, starting from $\mathcal{L}_f = \mathcal{L}$, which contains all potential locations, $\ell \in \mathcal{L}_f^1$ if $\Delta^{\ell}\Pi_f(\mathcal{L}) > 0$. Also, at the other extreme, starting from $\mathcal{L}_f = \emptyset$, which contains no entries, $\ell \notin \mathcal{L}_f^1$ if $\Delta^{\ell}\Pi_f(\emptyset) < 0$. The first round of mapping confirms some elements of the location vector. Now I iterate the mapping until a complete equilibrium location set is reached with no possibility of further refinement. When there are indefinite locations, the set of possible vectors is sliced to any two subsets, and then map each of the subsets separately. Slicing and mapping is repeatedly done until a unique optimal location vector $\mathcal{L}_f^*$ emerges.

Firms take turns to solve the best location response. The multi-plant firm location game in this paper is also a best-response potential game shown in Section 2.3. Developed in Voorneveld (2000), a best-response potential game, under the condition of a finite game where the number of players is finite and each of them has a finite strategy space, always has pure strategy Nash equilibrium. Moreover, starting from any arbitrary location decision, if players simultaneously deviate to their unique best replies in each period, the process terminates in a Nash equilibrium after finite number of steps. Swenson and Kar (2017) found that the convergence rate is exponential. To gauge the computational cost of solving a multi-player CDC problem, I simulate examples of 6 to 12 locations and two firms. Each example is simulated 1000 times. Table C.4 shows that the maximum number of rounds to find an equilibrium is three. When the potential number of locations is larger and therefore the strategy space is larger, it takes longer to find an equilibrium, but still converges to a solution relatively quickly.

Table C.4: Convergence rate check of best-response potential game

Number of locations	Average time (seconds)	Average number of BR rounds	Max number of BR rounds
6	0.0198	1.0830	3
7	0.0429	1.1010	2
8	0.0494	1.0190	2
9	0.0596	1.1830	3
10	0.0934	1.1230	3
11	0.0963	1.1980	2
12	0.1275	1.1130	2

Next, I would like to briefly discuss different ways of handling multiple equilibria and the reason I choose to impose a certain entry sequence. There are four main approaches in the literature to deal with the multiplicity of equilibria. The first is to model the probabilities of aggregated outcomes that are robust to multiplicity. For example, in the simplest $2 \times 2 \times 1$ game, the number of entrants is unique although the firm identity is undetermined (Bresnahan and Reiss, 1990; Bresnahan and Reiss, 1991; Berry, 1992). However, information on firm heterogeneity is lost. If I used it in this paper, I would not be able to estimate the fixed cost distributions, which are firm-location specific. The second is to embrace the multiplicity and take a bounds approach (Ciliberto and Tamer, 2009; Holmes, 2011; Pakes et al., 2015). The method partially identifies parameters within a set that could be too large to be informative. Lack of point identification also becomes difficult when performing counterfactual exercises. The third approach—the one taken here—is to choose an equilibrium by imposing a certain entry sequence following Jia (2008), Atkeson and Burstein (2008), Eaton et al. (2012), and Edmond et al. (2015) among many others. Although I model the entry game as static, the assumption is convenient to avoid multiple equilibria. In principle, estimates could be sensitive to the equilibrium selected and the predetermined order of entry. Therefore, I provide robustness checks by estimating the model based on equilibria with other ordering specifications. A more recent development of the literature involves specifying a more general equilibrium selection rule that is a function of covariates and observables, as in Bajari et al., 2010. The solution requires computing all equilibria and an equilibrium selection parameter as part of the primitives to be estimated together with the model. Although this approach is more general than imposing a certain entry sequence, the computational burden to calculate all equilibria in an interdependent entry game is prohibitive.

With the solution of the multi-plant firms' location game, the method of simulated moments works as following. For the log-normally distributed fixed costs, I draw a 2×73-dimensional matrix of fixed costs 300 times. I follow Antràs et al. (2017) by using quasi-random numbers from a van der Corput sequence, which has better coverage properties than usual pseudo-random draws. For each draw, firms maximize total expected profits by choosing where to build plants using the algorithm above. I then use the fraction of entry over 300 draws as the simulated entry probability for each firm in every location. There is actually another level of simulation for firm markups. Notice that the expected variable profit function (12) involves numerical integration over the markup. I use a stratified random sampling method in order to obtain good coverage of the higher markup. I define intervals from 1 to $\bar{\mu} = \eta/(\eta - 1)$, [1, 1.1, 1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 1.8, 1.9, 1.95, 1.97, 1.99, $\bar{\mu}$]. I then draw 5 uniform random numbers within these intervals. The draws receive a weight inversely proportional to the length of the interval. The integral part of the profit function is approximated by $\int_1^{\bar{\mu}} f(\mu) \approx \sum_{s=1}^{S} w_s f(\mu_s)$.

The vector of moment functions, $g(\cdot)$, specifies the differences between the observed equilibrium outcomes and those predicted by the model. The following moment condition is assumed to hold at the true parameter value $\delta_0 = \{\beta^F, \sigma^F\}$:

$$E\left[g(\delta_0)\right] = 0.$$

MSM finds an estimate such that

$$\hat{\delta} = \arg\min_{\delta} \frac{1}{|\mathcal{L}|}\left[\sum_{\ell=1}^{|\mathcal{L}|} \hat{g}(\delta)\right]' W \left[\sum_{\ell=1}^{|\mathcal{L}|} \hat{g}(\delta)\right], \tag{C-4}$$

where $\hat{g}(\cdot)$ is the simulated estimate of the moment function and W is a weighting matrix. I use the identity matrix and weight the moments equally as baseline.

The complexity in the presence of having spatial correlation is that the moment functions $g(\cdot)$ are no longer independent across locations. In order for the MSM estimators using a dependent cross-sectional dataset to be consistent, a sufficient condition is that the dependence between locations should fade quickly as the distance increases (Conley, 1999). In the current model setup, competition between plants becomes weaker when locations are further apart due to trade costs. To ensure the speed of dependence decay, I further segregate the 149 FAF zones into eight districts and assume that competition is negligible across them.

The map in Figure C.1 and Table C.5 show the division. The areas shaded in gray in the districts map are FAF zones without cement production. Consumption and production are roughly the same for each district, indicating smaller share of trade with areas outside. Furthermore, Figure C.2 shows the distribution of FAF zones trading within the same district. FAF zones on average export more than 88% of the cement production and import more than 82% of the consumption within the same district. Out of the 73 producing zones, all of them exported at least 50 percent to other FAF zones within the same district and more than three-fourths exported more than 80 percent within the same district. As for the importing cement markets, the distribution is slightly dispersed. But still, three-quarters of the 149 markets imported more than 80 percent from FAF zones located within the same district and more than 90 percent of the markets import at least half of their cement consumption within the district. These trade flow statistics validate my assumption of districts being relatively separated from one another.

An alternative way to restrain the geographic scope of the spillover effect is by assuming dependence only occurs for the set of locations within a certain radius to each location, as in Jia (2008). However, this method does not work for the multi-plant firm model. Existence of overlaps across each location's catchment area causes a violation of the submodularity of profit function, which is essential when solving the equilibrium.

Cluster bootstrap is used to estimate the standard errors. District vectors are re-sampled 100 times with replacement to preserve the dependence among locations. An alternative is to use the

Figure C.1: Districts map

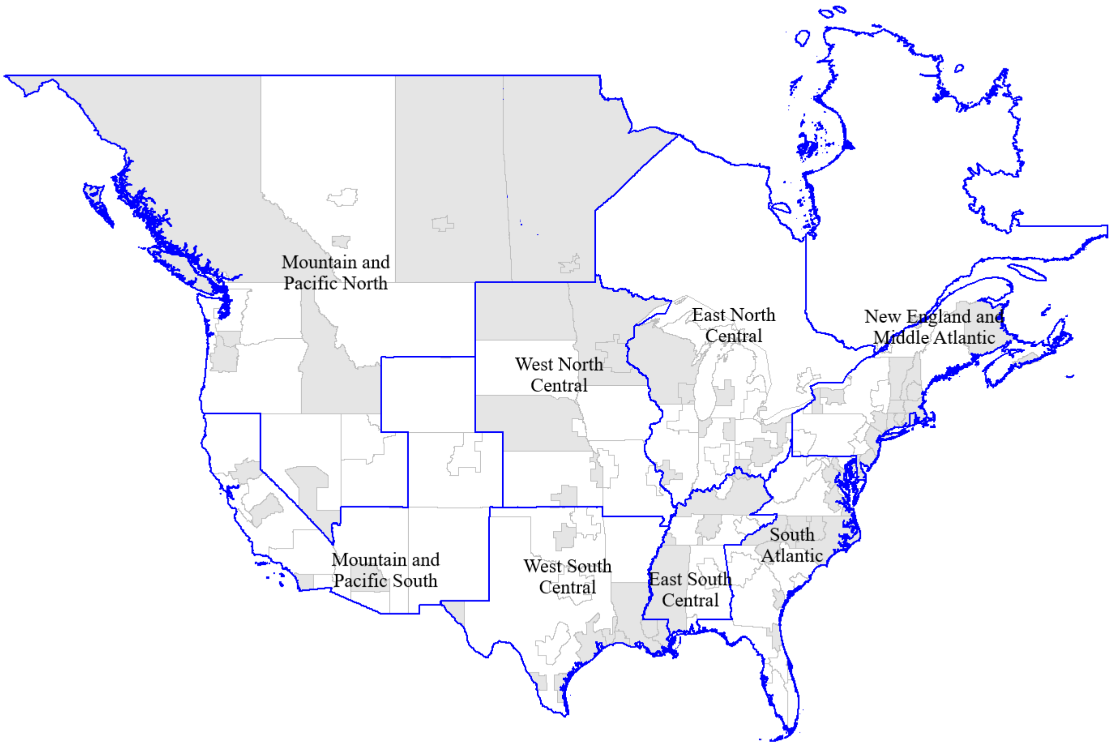


Table C.5: Summary statistics of districts

	Consumption (million ton)	Production (million ton)	Number of Markets	Number of Locations	Number of Plants
Mountain and Pacific North	10.2	10.4	20	10	13
Mountain and Pacific South	13.9	14.2	13	9	16
West North Central	8.8	8.8	13	7	11
West South Central	16.5	16.1	17	7	15
East North Central	15.8	16.5	22	12	19
East South Central	4.3	4.1	11	6	8
New England and Middle Atlantic	10.9	10.5	28	10	18
South Atlantic	16.2	16.1	25	12	17

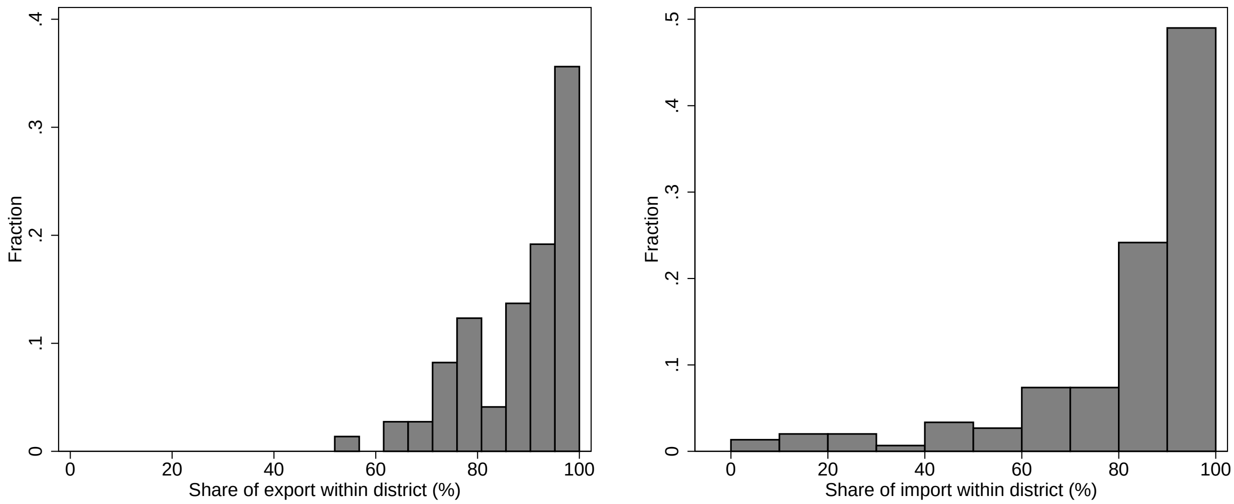
asymptotic normality of the MSM estimators. With spatial dependence, the asymptotic covariance matrix of moment function according to Conley (1999) and Conley and Ligon (2002) should be

$$V_0 = \sum_{\ell' \in R_\ell} E\big[g(\delta_0; \mathbf{X}_f, \hat{\boldsymbol{\phi}}, \hat{\mathbf{A}}, \hat{\theta}, \hat{\eta}) g(\delta_0; \mathbf{X}_f, \hat{\boldsymbol{\phi}}, \hat{\mathbf{A}}, \hat{\theta}, \hat{\eta})'\big],$$

and its sample analogue is

$$\hat{V} = \frac{1}{|\mathcal{L}|} \sum_{\ell} \sum_{\ell' \in R_\ell} \big[\hat{g}(\delta)\hat{g}(\delta)'\big],$$

Figure C.2: Trade within the same district



(a) Exports

(b) Imports

where R_ℓ is the set of locations belonging to the same district as location ℓ. Note that the variance-covariance estimator is not always positive semidefinite. I follow Jia (2008) and use a numerical device to weight the moment by 0.5 for all the neighbors.

Adjusted for spatial correlation, the asymptotic distribution is

$$\sqrt{|\mathcal{L}|}(\hat{\delta}-\delta_0) \xrightarrow{d} N(\mathbf{0},(1+S^{-1})(G_0'W_0G_0)^{-1}G_0'W_0V_0W_0G_0(G_0'W_0G_0)^{-1}),$$

where the gradient matrix $G_0 = E\left[\nabla_{\delta'} g(\delta_0)\right]$ and S is the number of simulations of the fixed cost draws. In practice, I take 600 simulation draws from a van der Corput sequence for good coverage.

Associated with the covariance matrix, one can also use the optimal weighting matrix, $W_0 = V_0^{-1}$ instead of an identity matrix. Using a consistent estimator of the optimal weighting matrix, the MSM estimates are asymptotically efficient, with the asymptotic variance being

$$\text{Avar}(\hat{\delta}) = (1+S^{-1})(G_0'V_0^{-1}G_0)^{-1}/|\mathcal{L}|.$$

Table C.6 displays results using a combination of identity weighting matrix, optimal weighting matrix, clustered bootstrap standard errors, and asymptotic standard errors. Across different methods, the estimates are consistent and close, and those using optimal weighting matrix exhibit slightly greater precision.

A final remark is that the discrete choice decisions makes the objective function non-smooth and the firm's problem not globally convex. The shortcoming is that I cannot guarantee that my solution is the global optimum of the problem. To address this issue, I tried the particle swarm

Table C.6: Robustness check: estimation of fixed costs

	Favor LafargeHolcim			Favor Cemex			Local advantage for two firms		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
β^F_{cons}	-6.631	-6.631	-6.643	-6.126	-6.126	-6.038	-5.617	-5.617	-5.616
	(1.616)	(1.048)	(0.209)	(1.688)	(1.268)	(0.138)	(1.559)	(0.621)	(0.165)
$\beta^F_{\text{CEX-USA}}$	-0.406	-0.406	-0.313	-0.363	-0.363	-0.303	-0.280	-0.280	-0.234
	(0.373)	(1.707)	(0.180)	(0.382)	(0.661)	(0.145)	(0.372)	(0.318)	(0.158)
$\beta^F_{\text{LFH-CAN}}$	-3.734	-3.734	-3.698	-3.475	-3.475	-3.430	-3.480	-3.480	-3.587
	(1.867)	(0.724)	(1.702)	(2.318)	(1.046)	(0.255)	(1.992)	(1.133)	(1.616)
β^F_{dist}	1.795	1.795	1.803	1.698	1.698	1.700	1.634	1.634	1.648
	(0.220)	(0.130)	(0.018)	(0.245)	(0.073)	(0.021)	(0.221)	(0.080)	(0.025)
σ^F	2.790	2.790	2.568	2.581	2.581	2.437	2.694	2.694	2.591
	(0.481)	(0.472)	(0.159)	(0.504)	(1.342)	(0.105)	(0.503)	(0.411)	(0.104)

Columns (1), (4) and (7) are baseline estimates in Table 3 using identity weighting matrix and bootstrapped standard errors. Columns (2), (5), and (8) are estimates using identity weighting matrix and asymptotic standard errors. Columns (3), (6) and (9) are 2-step estimates using optimal weighting matrix and asymptotic standard errors. The first step is performed using identity weight on moments, followed by computing the optimal weight using the first-step estimates to be fed in the second-step estimation.

optimization algorithm to search through 100 starting points. All sets of starting points resulted in close outcomes.

C.4 Monetary transformation of estimates

In the first step of the estimation, location fixed effects are estimated up to a scalar, i.e., $\widetilde{\text{FE}}_\ell = \lambda \ln\left(N_\ell T_\ell w_\ell^{-\theta}\right)$ where the normalization constant is denoted as λ. Hence, the estimated market sourcing potential is $\widetilde{\Phi}_m = \sum_\ell \exp(\widetilde{\text{FE}}_\ell)\tau_{\ell m}^{-\theta} = e^{\lambda}\Phi_m$. Substituting to equation (13), the estimated local price index with normalization would be

$$\widetilde{P}_m = \Gamma\left(\frac{\theta+1}{\theta}\right)\widetilde{\Phi}_m^{-1/\theta} \times \left[(1-N) + \sum_{f\in\mathcal{F}}\left(1-(1-\bar{\mu}^{-\theta})s_{fm}\right)^{-1/\theta}\right] = e^{-\lambda/\theta}P_m,$$

where the normalization parameter enters the price index multiplicatively through Φ_m. Correspondingly, firm's costs are also scaled by $e^{-\lambda/\theta}$. In order to map the estimates to their dollar values, I run the observed cement price on the model prediction and obtain a slope of 140.575, which would then be used to scale up all the estimated costs.

Each firm's cost of supplying cement is computed by taking average of equation (4) across all the markets adjusted for the scaling. Their average fixed costs are computed based on the mean of distribution (19) for locations where firms choose to build plant and also multiplied by 140.575. Since the sample is for one year only, the fixed costs are further transformed to net present values using 6 percent interest rate for discounting.

C.5 Additional check for model fitness

Table C.7 provides additional checks of the model prediction to the trade data. The predicted bilateral share of imports is able to explain 64.4 percent of the data variation. To check to what extent the prediction is affected by the gravity errors, I regress the final prediction after solving for the endogenous plant locations on the gravity-predicted import share. The fit improves by around 20 percent. Restricting the sample to intra-district trade further increases the fit by another 6.7 percent. Since the import share is indirectly targeted through the first-step gravity regression, I further compare the trade volume as shown in the last column of Table C.7. The degree of fit does not fall.

Table C.7: Model fit of trade flows

	Bilateral share of import	Gravity-predicted share of import	Gravity-predicted share of import within district	Bilateral import volume
Model prediction	0.767	0.797	0.990	0.631
	(0.005)	(0.003)	(0.008)	(0.004)
Observations	10877	10877	1437	10877
R^2	0.644	0.850	0.917	0.645

All regressions include a constant. Column (1) is regressed on the actual bilateral share of imports. Column (2) is regressed on the gravity-predicted import share after teasing out the gravity errors. Column (3) restricts the sample to intra-district trade. Column (4) compares to the trade volume instead of share.

D Counterfactual details

Using Table D.8 and the fact that producing one tonne of cement requires an energy of 4.432 million BTU, I compute the average cost of fuel to produce a tonne of cement in 2016 before the carbon levy $= (42\% \times 2.366 + 22\% \times 5.003 + 13\% \times 1.722 + 4\% \times 12.223) \times 4.432 = \$12.44$/tonne cement. The pre-tax unit cost of fuel is close to \$13.82, as found by Miller et al. (2017) using 2010 data. After the carbon levy, rates for each fuel subject to the levy are set based on the Canadian Federal Carbon Pricing Backstop Technical Paper, such that they are equivalent to \$50 per tonne of CO_2 by 2022. Assuming that there is no substitution of fuel to other carbon-saving sources after the policy, the levy on fuel by 2022 becomes $= (42\% \times (158.99/27.77) + 22\% \times (0.0979/0.035)+$

$13\% \times (0.1919/0.04)$ $+4\% \times (0.1593/0.036)) \times 4.432 =$ \$16.93/tonne cement, and hence the cost of fuel in 2022 will be $16.93 + 12.44 =$ \$29.37/tonne cement.

Table D.8: Fuel costs and energy content

	Energy Source Breakdown (%)	Energy Content	Price, 2016 (\$/mBTU)	Levy, 2022
Coal (coke)	42	27.77 mBTU/t	2.366	\$158.99/t
Natural gas	22	0.035 mBTU/m^3	5.003	\$0.0979/m^3
Petroleum coke	13	0.04 mBTU/L	1.722	\$0.1919/L
Heavy fuel oil	4	0.036 mBTU/L	12.223	\$0.1593/L

Based on the Portland Cement Association's US and Canadian Portland Cement Labor-Energy Input Survey, the amount of energy required to produce one tonne of cement is **4.432** million BTU. The remaining 11% energy is provided by electricity and 7% by other sources, which are carbon tax free and excluded from the computation of fuel cost. computing cost of fuels.
Source: Energy Consumption Benchmark Guide: Cement Clinker Production, Energy Fact Book 2019-2020 (Natural Resources Canada), Technical Paper on the Federal Carbon Pricing Backstop, US Energy Information Administration energy conversion calculators.

E SP approximation details

The expected variable profit of a plant at ℓ is proportional to its parent firm's profit depending on the share of consumers it supplied due to the identical price distribution (6) across plants owned by the same firm. With the Fréchet distributed productivities, the share of consumers sourcing from plant ℓ over all its firm's consumers in m is $s_{f\ell m} = \frac{\phi_{\ell m}}{\Phi_{fm}}$, making the expected variable profit $E\big[\pi_{f\ell}\big] = \sum_m s_{f\ell m} E\big[\pi_{fm}\big]$. I construct it using the same first two-step estimates from Sections 4.1 and 4.2 to disentangle the effects solely stemming from assuming separate entry of plants. Table E.9 reports the binary Probit regression results in the single-plant approximation.

When translating the SP estimates to monetary terms, I also use 140.575 as the scalar to match actual prices and discount rate 6% to be consistent when comparing to the multi-plant estimates. Table E.10 shows the changes to firms and market aggregates when \$50 carbon tax on fuel is imposed to a model using SP approximated costs. Indeed, given smaller fixed costs, the SP approximation predicts a 10 percentage point decrease in the number of Canadian plants relative to the MP predicted result, implying an over-prediction of 75 percent more closures of the top two cement firms' plants due to biased fixed cost estimates. Panel B finds that an over-prediction of plant relocation leads to an increase in the carbon leakage rate from 26 percent to almost 29 percent. The large difference in the prediction of plant relocation is partially offset by intensive margin adjustment among remaining plants, resulting in moderate but still nontrivial difference in carbon leakage rate. In terms of welfare, the biased estimates lead to a further reduction of 11 million for Canada. Policymakers who use the naive separate-entry approach to estimate multi-plant firms' interdependent location decisions would exaggerate the amount of production and carbon leakage.

Table E.9: Estimation of entry without interdependency

	Probit
constant	0.123
	(1.489)
log distance to $\text{HQ}_{f\ell}$	-0.413^{b}
	(0.202)
log variable $\text{profits}_{f\ell}$	0.563^{a}
	(0.174)
LFH-CAN	0.884^{c}
	(0.482)
CEX-USA	0.166
	(0.281)
Observations	146
R^2	0.161

Robust standard errors are in parentheses. Significance levels: ^{c} $p<0.1$, ^{b} $p<0.05$, ^{a} $p<0.01$.

Table E.10: Aggregate effects of $50 carbon levy on fuel: MP vs. SP

Panel A: Impacts on market outcomes

	%ΔNumber of plants			%ΔPrice	%ΔConsum	%ΔProd	%ΔTrade	
	LFH	CEX	Combined				Canada	US
(a) MP:								
Canada	-10.98	-29.58	-12.75	27.85	-37.97	-66.14	-54.25	-94.67
US	0.57	1.16	0.81	0.69	-2.63	2.33	224.83	1.02
(b) SP:								
Canada	-19.16	-34.69	-22.35	28.68	-37.95	-66.84	-55.39	-94.91
US	1.06	1.55	1.28	0.56	-2.38	2.52	234.88	1.14

Panel B: Impacts on welfare and emissions

	Δ CS	Δ PS	Δ TaxRev	Δ Emissions	Leakage rate
(a) MP:					
Canada	-310.50	-68.04	77.40	-6.05	26.32
US	-35.54	10.70	-	1.60	-
(b) SP:					
Canada	-324.59	-64.87	77.25	-6.22	28.95
US	-30.02	9.54	-	1.80	-

Columns in Panel A are denoted as percentage change relative to baseline using MP or SP fixed cost parameters. Columns in Panel B are changes in levels relative to baseline. Consumer surplus, producer surplus, and government revenue are denoted in million US dollars. Emissions are denoted in million of tonnes. The leakage rate is represented as a percentage.

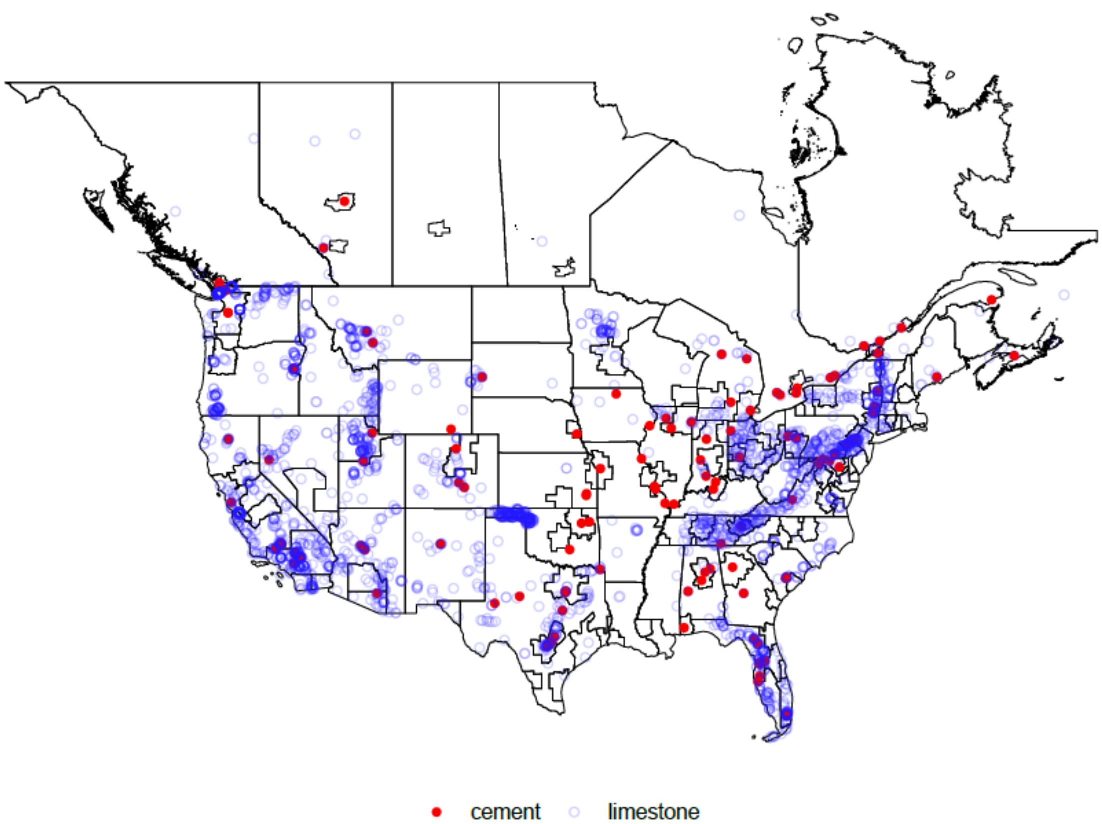
F Data appendix

F.1 Locations of limestone deposits and cement plants

Figure F.3 maps the distribution of cement plants versus limestone resources. The information is obtained from the US Geological Survey. There are 2909 limestone quarries in the US and 40 in Canada. Most of the FAF zones studied in my sample have at least one limestone quarry available. Obvious exceptions are Saskatchewan and North Dakota, where there are no limestone quarries or cement plants. The locations where access to limestone is limited are outside the potential set of locations to establish cement plants in my study.

Another issue is that large cement firms such as LafargeHolcim and Cemex typically use limestone mined from their own quarries, and process and transport it to their cement plants right after extraction. The vertical integration of limestone quarries and cement plants is not a focus of this paper. Since the cement plants are usually only a few kilometers away from the limestone quarries, the location choice of cement plants studied here can be regarded as a decision for an integrated set of facilities, including mining activities and further processing.

Figure F.3: Cement and limestone resource location distribution



F.2 Market structure details

In Table F.12, I report joint distributions for 26 cement firms by the number of plants owned and the number of production locations entered. Panel A presents the distribution of the number of firms; panel B shows the distribution of the number of plants owned; and panel C reports the distribution of market share measured by capacity. From panel A, one can see that 34.6 percent of firms are single-plant owners producing at one location. They account for 7.4 percent of cement plants and 6.5 percent of the market. In contrast, 11.5 percent of firms that own 11 or more plants across locations control around 40.5 percent of plants and 41.6 percent of the market. Large cement firms produce in more locations, own more plants, and have greater production capacity at each location. Nevertheless, the group of smaller cement manufacturers is also too big to ignore.

Table F.12: Distribution by number of plants and FAF zones

Panel A: Percentage of firms

	Number of FAF zones				
Number of plants	1	2-4	5-10	11+	**Total**
1	34.6	0.0	0.0	0.0	**34.6**
2-4	0.0	30.8	0.0	0.0	**30.8**
5-10	0.0	3.8	19.2	0.0	**23.1**
11+	0.0	0.0	0.0	11.5	**11.5**
Total	**34.6**	**34.6**	**19.2**	**11.5**	**100.0**

Panel B: Percentage of plants

	Number of FAF zones				
Number of plants	1	2-4	5-10	11+	**Total**
1	7.4	0	0	0	**7.4**
2-4	0	19	0	0	**19**
5-10	0	4.1	28.9	0	**33.1**
11+	0	0	0	40.5	**40.5**
Total	**7.4**	**23.1**	**28.9**	**40.5**	**100**

Panel C: Market share

	Number of FAF zones				
Number of plants	1	2-4	5-10	11+	**Total**
1	6.5	0	0	0	**6.5**
2-4	0	21.5	0	0	**21.5**
5-10	0	3.6	26.7	0	**30.4**
11+	0	0	0	41.6	**41.6**
Total	**6.5**	**25.1**	**26.7**	**41.6**	**100**

Notes: Without actual data on plants' sales, market share is proxied by the percentage of production capacity over the total installed capacity across all plants, assuming capacity is proportional to sales by a constant.

References

Antràs, P., Fort, T. C., and Tintelnot, F. (2017). The margins of global sourcing: Theory and evidence from us firms. *American Economic Review*, 107(9):2514–64.

Arkolakis, C. and Eckert, F. (2017). Combinatorial discrete choice. *Available at SSRN 3455353*.

Atkeson, A. and Burstein, A. (2008). Pricing-to-market, trade costs, and international relative prices. *American Economic Review*, 98(5):1998–2031.

Bajari, P., Hong, H., and Ryan, S. P. (2010). Identification and estimation of a discrete game of complete information. *Econometrica*, 78(5):1529–1568.

Berry, S. T. (1992). Estimation of a model of entry in the airline industry. *Econometrica: Journal of the Econometric Society*, pages 889–917.

Bresnahan, T. F. and Reiss, P. C. (1990). Entry in monopoly market. *The Review of Economic Studies*, 57(4):531–553.

Bresnahan, T. F. and Reiss, P. C. (1991). Entry and competition in concentrated markets. *Journal of Political Economy*, 99(5):977–1009.

Ciliberto, F. and Tamer, E. (2009). Market structure and multiple equilibria in airline markets. *Econometrica*, 77(6):1791–1828.

Conley, T. G. (1999). Gmm estimation with cross sectional dependence. *Journal of econometrics*, 92(1):1–45.

Conley, T. G. and Ligon, E. (2002). Economic distance and cross-country spillovers. *Journal of Economic Growth*, 7(2):157–187.

Eaton, J., Kortum, S. S., and Sotelo, S. (2012). International trade: Linking micro and macro. Technical report, National bureau of economic research.

Edmond, C., Midrigan, V., and Xu, D. Y. (2015). Competition, markups, and the gains from international trade. *American Economic Review*, 105(10):3183–3221.

Holmes, T. J. (2011). The diffusion of wal-mart and economies of density. *Econometrica*, 79(1):253–302.

Jia, P. (2008). What happens when Wal-Mart comes to town: An empirical analysis of the discount retailing industry. *Econometrica*, 76(6):1263–1316.

Miller, N. H., Osborne, M., and Sheu, G. (2017). Pass-through in a concentrated industry: empirical evidence and regulatory implications. *The RAND Journal of Economics*, 48(1):69–93.

Pakes, A., Porter, J., Ho, K., and Ishii, J. (2015). Moment inequalities and their application. *Econometrica*, 83(1):315–334.

Swenson, B. and Kar, S. (2017). On the exponential rate of convergence of fictitious play in potential games. In *2017 55th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, pages 275–279. IEEE.

Voorneveld, M. (2000). Best-response potential games. *Economics letters*, 66(3):289–295.