



Geografic Concentration and Estabilishment Scale

Thomas J. Holmes and John J. Stevens



Introduzione

La concentrazione geografica delle industrie è un tema pervasivo.

Sappiamo però relativamente poco sulla relazione che intercorre tra la concentrazione delle industrie e la dimensione delle imprese.

Obiettivi dell'analisi

Analizzeremo quindi la relazione tra la concentrazione industriale e la dimensione delle fabbriche.

Questo ci permetterà di capire perché le fabbriche localizzate in aree ad alta concentrazione industriale sono più grandi in media di quelle dello stesso specifico settore.



La questione venne analizzata anche da Ellison & Glaeser nel 1997:

→ anche se le fabbriche sono distribuite casualmente a livello geografico “*like darts on a dartboard*”, emergerà comunque un certo grado di concentrazione. Se casualmente in un luogo si sviluppa una fabbrica molto grande, conseguenzialmente sia la concentrazione industriale, sia la dimensione media delle fabbriche crescerà.

Inizialmente ci si potrebbe aspettare di trovare fabbriche piu piccole in aree ad alta concentrazione industriale, a causa dei fenomeni di “Vertical Disintegration”.



La concentrazione industriale può in effetti portare allo sviluppo di fabbriche piccole di beni intermedi, specializzate quindi in prodotti di nicchia → Per alcuni casi è così (industria dell'abbigliamento a New York), ma sono eccezioni.

Useremo l'indice di concentrazione Ellison-Glaeser invece che lo Standard Industrial Classification Code (il SIC ci da un quadro distorto del vero livello di concentrazione industriale).

Framework

Usiamo lo *standard employment location quotient* per misurare il livello di specializzazione industriale.

L' SELQ cattura le due possibili fonti di specializzazione industriale nella località l : la differenza nel numero di fabbriche pro-capite da quella del livello nazionale ($Q^n_{i,l}$) con le differenze nella dimensione media del settore rispetto alla media nazionale ($Q^s_{i,l}$).

$$Q^x_{i,l} = \frac{x_{i,l}/x_l}{x_i/x}$$

$$Q^x_{i,l} = Q^n_{i,l} \times Q^s_{i,l}$$

$$Q^n_{i,l} = \frac{n_{i,l}/x_l}{n_i/x}$$

$$Q^s_{i,l} = \frac{x_{i,l}/n_{i,l}}{x_i/n_i}$$

Riscrivo $Q^x_{i,l} = Q^n_{i,l} \times Q^s_{i,l}$

in forma logaritmica

$$q^x_{i,l} = q^n_{i,l} + q^s_{i,l}$$

Esistono tre metodi possibili per analizzare tale correlazione tra q^s e q^x :

1. Location-level regression
2. Plant-level regression
3. Excluded LQ

Location-Level Regression

Partendo da $Q_{i,l}^x = q_{i,l}^n + q_{i,l}^s$

Ricaviamo i paramentri β^s e β^n :

$$\beta^s = \frac{\text{cov}(q_{i,l}^s, q_{i,l}^x)}{\text{var}(q_{i,l}^x)}, \quad \beta^n = \frac{\text{cov}(q_{i,l}^n, q_{i,l}^x)}{\text{var}(q_{i,l}^x)}$$

Sappiamo che

$$\beta^s + \beta^n = 1$$

Casi estremi:

La concentrazione è spiegata solo dal numero di imprese

$$\beta^s = 0 \text{ and } \beta^n = 1$$

La concentrazione è spiegata solo da variazioni nella dimensione

$$\beta^s = 1 \text{ and } \beta^n = 0$$

Plant-Level Regression

$$q_e^x \equiv q^x$$

L'employment location quotient è lo stesso per ogni fabbrica 'e' e corrisponde a quello della sua località e della sua industria; ogni osservazione ha il suo size quotient:

$$q_e^s \equiv \ln \left(\frac{x_e}{x_i/n_i} \right)$$

Il coefficiente di regressione è quindi:

$$\beta_e^s = \frac{\text{cov}(q_e^s, q_e^x)}{\text{var}(q_e^x)}.$$

Bias Problem: in una situazione di distribuzione casuale di un numero finito di fabbriche, ovvero con $\beta^s=0$, avremo comunque β^s e β_e^s maggiori di zero.

Excluded LQ

$$\tilde{q}_e^x = \ln \left(\frac{\frac{x_{i,l} - x_e}{x_l - x_e}}{\frac{x_i - x_e}{x - x_e}} \right) \quad \tilde{q}_e^s \equiv \ln \left(\frac{x_e}{\frac{x_i - x_e}{n_i - 1}} \right)$$

Con tali indici ricalcoliamo il *Location Level Regression* escludendo la fabbrica di riferimento dal calcolo.

Simulazione di un modello senza correlazione tra dimensione e concentrazione

- N Fabbriche $\sim$ i.i.d. per ogni industria I (anch'esso un numero)
- Abbiamo $I/2$ che separa le industrie della prima metà dalle industrie della seconda metà.
- 2 tipi di regioni: A e B (50 localizzazioni ciascuna)
- Le industrie della prima metà tendono a localizzarsi in A (con prob $2/I50$ di localizzarsi in A e $1/I50$ prob di localizzarsi in B), le altre in B ($2/I50$ prob di localizzarsi in B, $1/I50$ prob di localizzarsi in A) $\rightarrow$ n° atteso imprese per tutte le industrie è uguale in ogni localizzazione.
- Ogni industria può essere grande o piccola (le grandi hanno il doppio dei dipendenti) con la stessa probabilità (per ognuna 50%)

Risultati simulazione

TABLE 1.—RESULTS FOR A SIMULATED MODEL WITH NO FUNDAMENTAL CONNECTION BETWEEN SIZE AND CONCENTRATION

N	Location Level	Plant Level	
	β^s	β_e^s	$\tilde{\beta}_e^s$
100	0.2389	0.2122	0.0022
500	0.0951	0.0823	−0.0021
1,000	0.0611	0.0555	−0.0018
10,000	0.0071	0.0068	−0.0134

Il data set CBP ci fornisce alcuni dati:

- Four-Digit data code: è un codice di quattro cifre a cui corrisponde il settore.

Esempio:

2300	Apparel & Other Finishd Prods of Fabrics & Similar Matl
2320	Men's & Boys' Furnishgs, Work Clothg, & Allied Garments
2330	Women's, Misses', and Juniors Outerwear
2340	Women's, Misses', Children's & Infants' Undergarments

- Località (Stato e Paese)
- Dimensioni (employment): le fabbriche vengono divise in categorie a seconda del numero di dipendenti (1-4; 5-9; ecc), successivamente ad ogni categoria viene assegnato il numero medio di dipendente che verrà poi assegnato ad ogni industria.



Usiamo due tipi di aggregazione per la nostra analisi: una a livello di contea (livello geograficamente piu dettagliato), l'altra a livello di macro-regione con nove regioni per tutti gli USA (livello poco dettagliato). Usiamo i dati del *Census*. Si tratta di un'analisi preliminare in cui le industrie vengono assegnate al maggior settore di riferimento: abbiamo pochi settori generici cui fanno riferimento tutte le fabbriche.

Per ogni settore, salvo quello manifatturiero in cui sono presenti tutte e nove le regioni, sono segnate le regioni in cui il settore in questione è piu concentrato o meno concentrato.

Analisi preliminare

TABLE 2.—LOCATION QUOTIENTS BY MAJOR SECTOR

Sector	Region	$Q^x_{i,l}$	$Q^n_{i,l}$	$Q^s_{i,l}$
Agriculture	Pacific	1.37	1.12	1.23
	East South Central	0.77	0.77	0.99
Mining	West South Central	3.74	4.19	0.89
	New England	0.13	0.57	0.22
Construction	Mountain	1.25	1.23	1.02
	New England	0.72	1.03	0.70
Manufacturing	East South Central	1.40	1.02	1.37
	East North Central	1.25	1.07	1.17
	New England	1.08	1.13	0.96
	West North Central	1.01	0.95	1.07
	South Atlantic	0.95	0.85	1.12
	Pacific	0.91	1.17	0.78
	West South Central	0.88	0.90	0.98
	Middle Atlantic	0.88	0.99	0.89
	Mountain	0.70	0.94	0.74
Transportation	Mountain	1.11	1.13	0.98
	New England	0.78	0.86	0.91
Wholesale	West North Central	1.10	1.16	0.94
	New England	0.85	0.83	1.02
Retail	Mountain	1.10	1.07	1.03
	Middle Atlantic	0.88	0.99	0.89
Fire	Middle Atlantic	1.29	1.09	1.18
	East South Central	0.71	0.86	0.82
Services	Middle Atlantic	1.10	0.99	1.11
	East South Central	0.86	0.90	0.95

Sono riportati i Q^x , Q^n e Q^s per i maggiori settori industriali delle regioni selezionate.

In generale osserviamo già preliminarmente che esiste una correlazione positiva tra la dimensione delle imprese e la loro concentrazione spaziale.

Analisi

Sector	Location Level	Plant Level	
	β^s	β_e^s	$\tilde{\beta}_e^s$
All	0.322 (0.00576)	0.183 (0.00153)	0.161 (0.00152)
Agriculture	0.313 (0.0903)	0.141 (0.0119)	0.128 (0.0119)
Mining	0.193 (0.0162)	0.0304 (0.00613)	0.0240 (0.00608)
Construction	0.410 (0.0429)	0.225 (0.00668)	0.213 (0.00668)
Manufacturing	0.436 (0.00674)	0.334 (0.00377)	0.274 (0.00370)
Transportation	0.316 (0.0314)	0.183 (0.00792)	0.151 (0.00785)
Wholesale	0.135 (0.0154)	0.0907 (0.00478)	0.0747 (0.00477)
Retail	0.224 (0.0245)	0.228 (0.00446)	0.217 (0.00446)
FIRE	0.398 (0.0308)	0.265 (0.00488)	0.253 (0.00487)
Services	0.359 (0.0153)	0.137 (0.00312)	0.121 (0.00311)

- La regressione è stata eseguita a livello di 9 macro-regioni
- Riferimento a 9 macro-settori
- Osserviamo che $\beta^s > \beta_e^s > \tilde{\beta}_e^s$
- Tuttavia, la riduzione del parametro “excluded LQ” non è ampia come potevamo aspettarci (si abbassa ma non tende a zero)



**Solo una piccola parte
dell'analisi location-level è
spiegabile come una situazione
“dartboard”**

Regressione per contee

Sector	Location Level	Plant Level	
	β^s	β_e^s	$\tilde{\beta}_e^s$
All	0.365 (0.000816)	0.190 (0.000603)	0.073 (0.000572)
Agriculture	0.212 (0.00593)	0.083 (0.00367)	0.006 (0.00360)
Mining	0.232 (0.00474)	0.141 (0.00407)	0.087 (0.00397)
Construction	0.402 (0.00394)	0.221 (0.00235)	0.063 (0.00222)
Manufacturing	0.415 (0.00145)	0.274 (0.00165)	0.126 (0.00159)
Transportation	0.373 (0.00406)	0.208 (0.00281)	0.070 (0.00265)
Wholesale	0.338 (0.00196)	0.182 (0.00177)	0.048 (0.00167)
Retail	0.262 (0.00221)	0.200 (0.00169)	0.049 (0.00161)
FIRE	0.518 (0.00314)	0.240 (0.00191)	0.152 (0.00180)
Services	0.463 (0.00156)	0.212 (0.00115)	0.071 (0.00107)

Regressione per le dieci più grandi Msa

Sector	Location Level	Plant Level	
	β^s	β_e^s	$\tilde{\beta}_e^s$
All	0.402 (0.00580)	0.197 (0.00229)	0.149 (0.00225)
Agriculture	0.833 (0.102)	0.457 (0.0241)	0.394 (0.0234)
Mining	0.261 (0.0198)	0.0427 (0.0151)	0.0205 (0.0148)
Construction	0.542 (0.0428)	0.304 (0.0104)	0.271 (0.0104)
Manufacturing	0.443 (0.00604)	0.288 (0.00539)	0.181 (0.00513)
Transportation	0.452 (0.0282)	0.217 (0.0130)	0.136 (0.0126)
Wholesale	0.273 (0.0156)	0.0900 (0.00719)	0.0491 (0.00715)
Retail	0.479 (0.0279)	0.447 (0.00765)	0.409 (0.00762)
FIRE	0.365 (0.0378)	0.172 (0.00690)	0.144 (0.00685)
Services	0.409 (0.0128)	0.155 (0.00398)	0.127 (0.00397)

Considerazioni

- In queste due tabelle osserviamo β^s e β_e^s maggiori



- **Una maggiore disgregazione geografica implica un maggiore impatto per ciascuno stabilimento**

Settore manifatturiero

- Dimostra una forte relazione tra la dimensione e il livello di specializzazione
- In questa tabella vediamo 20 micro-categorie di industrie manifatturiere
- Per ognuna è assegnate le regioni con il livello di $Q(x)$: più alto, più basso, e più vicino a 1
- Per quasi tutte le regioni, è uguale l'ordinamento di $Q(x)$ e di $Q(s)$.
- Da notare il comportamento delle micro-fabbriche nelle regioni con scarsa concentrazione industriale

Four-Digit Industry	Region	$Q_{i,l}^x$	$Q_{i,l}^s$
2011 Meat-packing plants	West North Central	5.73	2.28
	East South Central	1.01	0.78
	New England	0.04	0.14
2111 Cigarettes	South Atlantic	4.90	1.17
	Middle Atlantic	0.01	0.02
	West North Central	0.001	0.001
2211 Broadwoven fabric mills, cotton	South Atlantic	4.47	1.96
	West South Central	0.41	0.45
	East North Central	0.02	0.05
2311 Men's and boys' suits and coats	East South Central	2.56	1.55
	West South Central	0.85	1.06
	Mountain	0.11	0.31
2411 Logging	East South Central	2.64	0.97
	New England	0.99	1.02
	Middle Atlantic	0.17	0.64
2511 Wood household furniture	South Atlantic	2.74	2.10
	East North Central	0.68	0.96
	West South Central	0.34	0.58
2611 Pulp mills	East South Central	4.55	1.07
	New England	0.90	1.17
	West North Central	0.01	0.04
2711 Newspapers	New England	1.24	1.27
	South Atlantic	1.00	1.20
	East South Central	0.78	0.70
2812 Alkalis and chlorine	West South Central	3.61	1.60
	Pacific	0.76	0.78
	West North Central	0.10	0.14
2911 Petroleum refining	West South Central	4.49	1.36
	East North Central	0.88	1.22
	New England	0.005	0.03

Four-Digit Industry	Region	$Q_{i,l}^x$	$Q_{i,l}^s$
3011 Tires and inner tubes	East South Central	5.42	1.91
	West North Central	0.93	0.79
	Mountain	0.005	0.01
3111 Leather tanning and finishing	New England	2.28	0.81
	South Atlantic	0.92	1.83
	Mountain	0.04	0.05
3211 Flat glass	East South Central	4.08	2.28
	South Atlantic	0.98	0.87
	New England	0.06	0.20
3312 Blast furnaces and steel mills	East North Central	2.71	1.74
	East South Central	1.07	0.76
	New England	0.04	0.11
3411 Metal cans	East North Central	1.43	1.17
	South Atlantic	1.00	1.10
	New England	0.18	0.39
3511 Turbines and turbine generator sets	New England	3.40	0.92
	Pacific	0.88	1.19
	Mountain	0.001	0.00
3612 Transformers, except electronic	East South Central	2.70	2.17
	West North Central	0.90	0.87
	New England	0.20	0.27
3711 Motor vehicles and car bodies	East North Central	3.39	1.84
	West North Central	1.06	0.93
	New England	0.01	0.02
3812 Search and navigation equipment	Pacific	1.89	1.11
	West North Central	1.01	0.60
	East South Central	0.15	0.45
3911 Jewelry, precious metal	New England	3.32	1.47
	West North Central	0.76	1.67
	East South Central	0.10	0.58

Caso 2335

Il settore della moda femminile (four-digit code 2335) risulta un'eccezione per quanto visto finora. Già Lichtenberg nel 1960 aveva notato una relazione inversa tra concentrazione e dimensione delle fabbriche operanti nel settore, infatti:

1. Il settore era già allora altamente concentrato nella Middle Atlantic Region ($Q_x=2,70$), mentre per la stessa regione il size index era solo di 0,9 (ovvero, le fabbriche presenti in questa regione tendevano ad essere più piccole rispetto alla media nazionale).
2. A livello urbano (NYC) il settore risulta addirittura più concentrato con fabbriche ancora più piccole.



**Questo è uno dei rari casi in cui
la concentrazione di un settore industriale,
porta alla nascita di piccole imprese
che si occupano di prodotti di nicchia.**

Sector	Average $\tilde{Q}_e^s$ for $\tilde{Q}_e^x$ in:						
	Quintile 1	2	3	4	5	Bottom 5%	Top 5%
Census Regions							
All	0.91	0.99	1.00	1.02	1.09	0.83	1.22
Agriculture	0.94	1.03	1.02	0.93	1.14	1.12	0.59
Mining	0.76	1.04	1.07	1.38	0.81	0.62	0.66
Construction	0.89	0.98	1.02	1.10	1.01	0.84	1.19
Manufacturing	0.76	0.91	1.01	1.11	1.25	0.67	1.35
Transportation	0.93	1.10	0.97	1.11	0.91	0.93	1.22
Wholesale	0.95	0.97	1.04	1.00	1.04	0.96	1.06
Retail	0.93	1.00	1.03	1.02	1.03	0.91	1.10
FIRE	0.86	1.00	0.99	1.03	1.13	0.81	1.29
Services	0.95	0.96	0.99	1.01	1.10	0.85	1.20
Counties							
All	0.86	1.00	1.04	1.03	1.12	0.77	1.30
Agriculture	0.91	0.99	0.98	0.99	1.16	0.81	1.39
Mining	0.71	0.83	0.73	0.87	2.10	0.62	3.34
Construction	0.90	1.01	1.03	1.04	1.05	0.85	1.11
Manufacturing	0.83	0.87	0.96	1.04	1.31	0.78	1.52
Transportation	0.83	1.06	1.02	1.10	1.14	0.81	1.20
Wholesale	0.88	1.00	1.06	1.06	1.09	0.83	1.08
Retail	0.93	1.03	1.07	1.04	0.96	0.84	0.95
FIRE	0.68	0.91	1.09	1.11	1.29	0.53	1.54
Services	0.86	0.99	1.02	1.03	1.15	0.77	1.39

Per ogni settore abbiamo diviso la popolazione di industrie in quintili in base al valore di $\tilde{Q}_e^x$, successivamente abbiamo calcolato il valore medio di $\tilde{Q}_e^s$.
I valori ottenuti confermano una relazione positiva (soprattutto per manifattura e servizi).

Implicazioni per la misura della concentrazione

- La classificazione four-digit non misura l'attività delle fabbriche (equipara piccole attività che offrono servizi nel settore con grandi stabilimenti industriali)
- Abbiamo calcolato l'indice di concentrazione usato da Ellison & Glaeser per il settore manifatturiero
- Analizziamo l'indice in questione:

$$\gamma \equiv \frac{\sum_{i=1}^M (s_i - x_i)^2 - \left(1 - \sum_i x_i^2\right) \sum_k z_k^2}{\left(1 - \sum_i x_i^2\right) \left(1 - \sum_k z_k^2\right)} = \frac{\sum_{i=1}^M (s_i - x_i)^2 - \left(1 - \sum_i x_i^2\right) H}{\left(1 - \sum_i x_i^2\right) (1 - H)}$$

In cui:

S(i) è la percentuale di lavoratori impiegati (nel settore considerato) per la regione i;

X(i) percentuale di occupati totale nella regione i;

Z(k) è la dimensione della fabbrica k rispetto al settore considerato;

H è l' Herfindahl Index

- Possiamo riscrivere l'Indice utilizzando l'indice di Gini:

$$\gamma_l^{EG} = \frac{G_l - \left(1 - \sum_{i=1}^M x_i^2\right) H_l}{\left(1 - \sum_{i=1}^M x_i^2\right) (1 - H_l)}$$

Vediamo i risultati della stima calcolando l'indice EG per le fabbriche del settore manifatturiero, per valori crescenti del numero dei dipendenti.

TABLE 8.—ELLISON-GLAESER INDICES FOR FOUR-DIGIT MANUFACTURING

Areas	Statistic	EG Index by Plant Employment				
		0+	20+	50+	100+	250+
Census regions	Mean	0.087	0.091	0.094	0.102	0.120
	25%	0.013	0.014	0.014	0.014	0.012
	Median	0.043	0.044	0.048	0.053	0.066
	75%	0.100	0.108	0.124	0.134	0.174
States	Mean	0.045	0.047	0.048	0.051	0.058
	25%	0.008	0.008	0.009	0.009	0.005
	Median	0.022	0.025	0.025	0.027	0.031
	75%	0.053	0.056	0.061	0.066	0.085
Counties	Mean	0.011	0.011	0.011	0.012	0.012
	25%	0.002	0.002	0.002	0.002	0.001
	Median	0.005	0.005	0.005	0.005	0.004
	75%	0.010	0.010	0.011	0.012	0.012
No. of Industries		459	459	458	457	428
Share of emp. (%)		100	93.4	84.8	74.9	57.4

Considerazioni

- Vi è una chiara correlazione tra il livello dell'indice EG a livello regionale e statale: eliminando progressivamente gli stabilimenti più piccoli, l'indice aumenta. Questo conferma che i fattori “dartboards” aumentano al diminuire delle fabbriche.
- Notiamo che mentre il 25° percentile non ha un andamento definito via via che aumentiamo le dimensioni degli impianti considerati, il 75° percentile invece cresce notevolmente, passando da 0,100 in 0+ a 0,174 in 250+.



La distribuzione dei grossi impianti manifatturieri è più geograficamente concentrata di quanto non avessimo immaginato.