

Performance delle imprese e dinamiche locali

Stefano De Santis
Stefania Cardinaleschi
Marina Schenkel
Francesco G. Truglia

Introduzione: Background

L' "atmosfera industriale" costruita nel lungo periodo di crescita, si è dissolta per effetto dei processi di delocalizzazione/internazionalizzazione/globalizzazione?.

L'effetto positivo sulla performance delle imprese derivante dall'appartenenza ai distretti è diminuito, sparito, o viceversa continua a manifestarsi?

Introduzione: Obiettivi del lavoro

Verificare se:

- la performance delle imprese è influenzata da caratteristiche territoriali
- i distretti industriali identificano gruppi omogenei di comuni dal punto di vista della performance delle imprese.
- I risultati delle analisi variano a seconda del livello territoriale

Fonti dei dati e ricodifica

La base dati utilizzata è stata creata attraverso il matching di fonti amministrative e statistiche, nel periodo 2004-2014.

:

- I dati di bilancio delle Camere di Commercio (Stato Patrimoniale e Conto Economico)
- Archivio Statistico delle Imprese Attive (ASIA)
- Classificazione delle unità territoriali ISTAT (livello di dettaglio NUTS 5)

Sono state necessarie

- l'armonizzazione di basi dati definite a livello annuale (diversità delle strutture dati nel tempo)
- la gestione delle variazioni territoriali

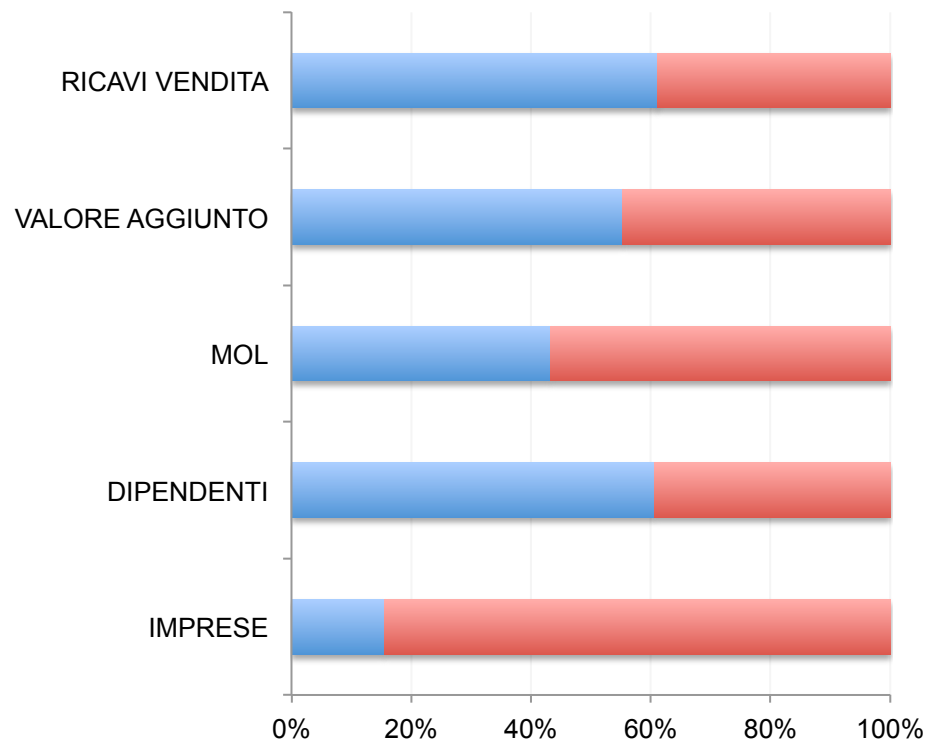
Fonti dei dati e ricodifica

Anni	Bilanci civiltistici	ASIA	FRAME-COE
2004	589.018	4.366.679	
2005	586.297	4.457.575	
2006	610.308	4.484.088	
2007	603.999	4.554.098	
2008	687.787	4.577.277	
2009	651.929	4.526.471	
2010	739.553	4.525.155	
2011	759.649	4.522.313	X
2012	703.894	4.515.691	X
2013	761.254	4.466.518	X
2014	756.962	4.263.878	X

Rappresentatività dei dati

Le fonti disponibili, di natura amministrativo-fiscale, possiedono diversi livelli di copertura e completezza dell'informazione.

Fonti dei dati e ricodifica



Rappresentatività dei dati

Le società soggette al deposito del Bilancio, pur rappresentando il 15,6% del totale delle imprese, costituiscono il segmento più importante dell'economia italiana in termini di fatturato (61,1%), valore aggiunto (55,3%), MOL (43,2%), occupazione (60,7%).

Metodologia di analisi

1. Ricodifica del ROE_{ti} in forma binaria:

$$ROI_{ti} \begin{cases} > 1 \rightarrow y_{ti} = 1 \\ \text{al contrario} \rightarrow y_{ti} = 0 \end{cases}$$

2. Struttura spaziale: matrice di contiguità spaziale \mathbf{W}^k k è l'ordine delle contiguità. L'elemento w_{ij} è:

$$w_{ij} = \begin{cases} 1 & \rightarrow \text{se il comune } i\text{-esimo confina con il comune } j\text{-esimo} \\ 0 & \rightarrow \text{al contrario} \end{cases}$$

3. Indicando J_i l'insieme dei comuni contigui al comune i -esimo e con k l'ordine delle contiguità, v_{11i}^k la variabile indicatrice è:

$$v_{11i}^k = \frac{\sum_{j \in J_i} w_{ij}^k y_{11ij}}{w_j}$$

4. Per i 7 anni considerati (T) la proporzione di $y_{ti}=1$ ($t = 2004, 2005, 2006, 2007, 2008, 2009, 2010$) è:

$$D_i = \frac{\sum_{t=04}^{10} y_{ti}}{T}$$

5. Modello logistico:

$$p(y_{11i} = 1) = \frac{e^{B_0 + B_1 v_{11i}^k + B_2 D_i}}{1 + e^{B_0 + B_1 v_{11i}^k + B_2 D_i}}$$

Metodologia di analisi

Statistiche descrittive

	DROI 2011	
	Frequenza	Percentuale
ROI≤0	1.866	23,1
ROI>0	6.228	76,9
Totale	8.094	100,0

	Statistiche descrittive			
	Minimo	Massimo	Media	Dev.Std.
V^4_{11i}	0	1	0,7719	0,1166
D_j	0	1	0,7722	0,3051

Stima parametri modello logistico

	Variabili nell'equazione					
	B	E.S.	Wald	df	Sig.	Exp(B)
B_0	-2,134	0,204	109,642	1	0,000	0,118
B_1	0,864	0,274	9,942	1	0,002	2,372
B_2	3,808	0,102	1396,111	1	0,000	45,057

Exp(B_0): registra come varia la probabilità di $y_{11i}=1$ rispetto alla probabilità contraria $y_{11i}=0$ (odds) per i comuni senza vicinato e che negli anni 2004-2010 il ROI < 0.

Exp(B_1): misura gli effetti spaziali (spillover) → odds $[p(y_{11i}=1)/p(y_{11i}=0)]$ che nel comune i-esimo in base ai valori assunti dal ROI nei comuni contigui di ordine k .

Exp(B_2): misura l'effetto temporale (persistenza/volatilità) → odds nel comune i-esimo in base ai valori assunti dal ROI negli anni 2004-2010.

Metodologia di analisi

Stima della bontà delle variabili inserite nel modello temporale (V^3_{ij} e D_j).

Modello: confronta il modello con le variabili V^4_{ij} D_j con il modello nullo (sola intercetta)

Blocco: confronto tra due modelli (modello nullo e modello con l variabili)

Più il chi-quadrato è elevato maggiore è la diminuzione di -2LL dovuta alle variabili aggiunte al modello nullo, maggiore è quindi la significatività di queste variabili

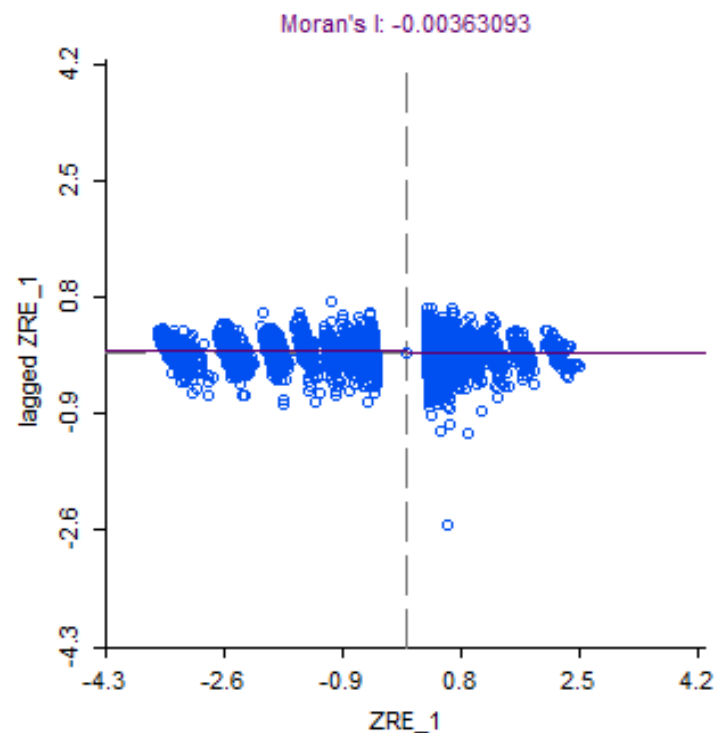
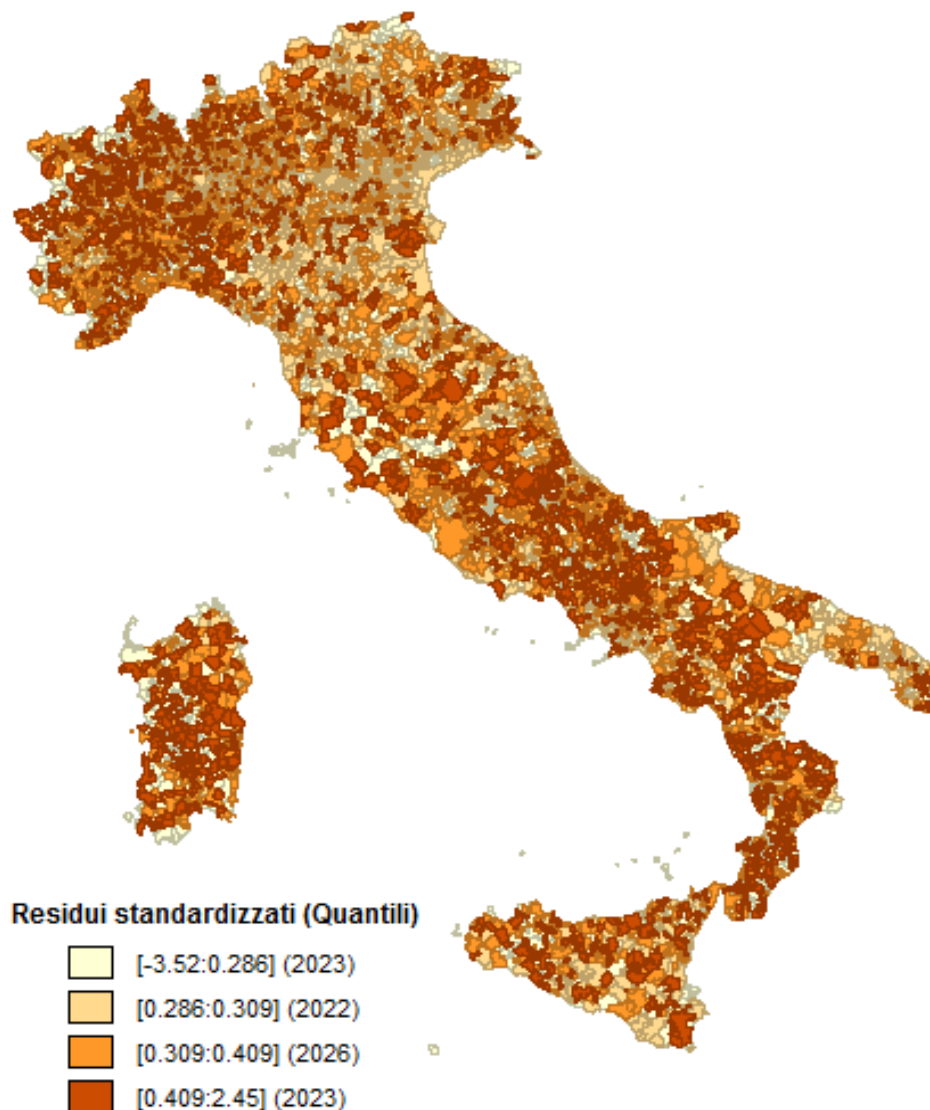
Test omnibus dei coefficienti del modello			
	Chi-quadrato	df	Sig.
Nulla	1.778,7	1,0	0,0
Modello	2.030,9	2,0	0,0

Metodologia di analisi

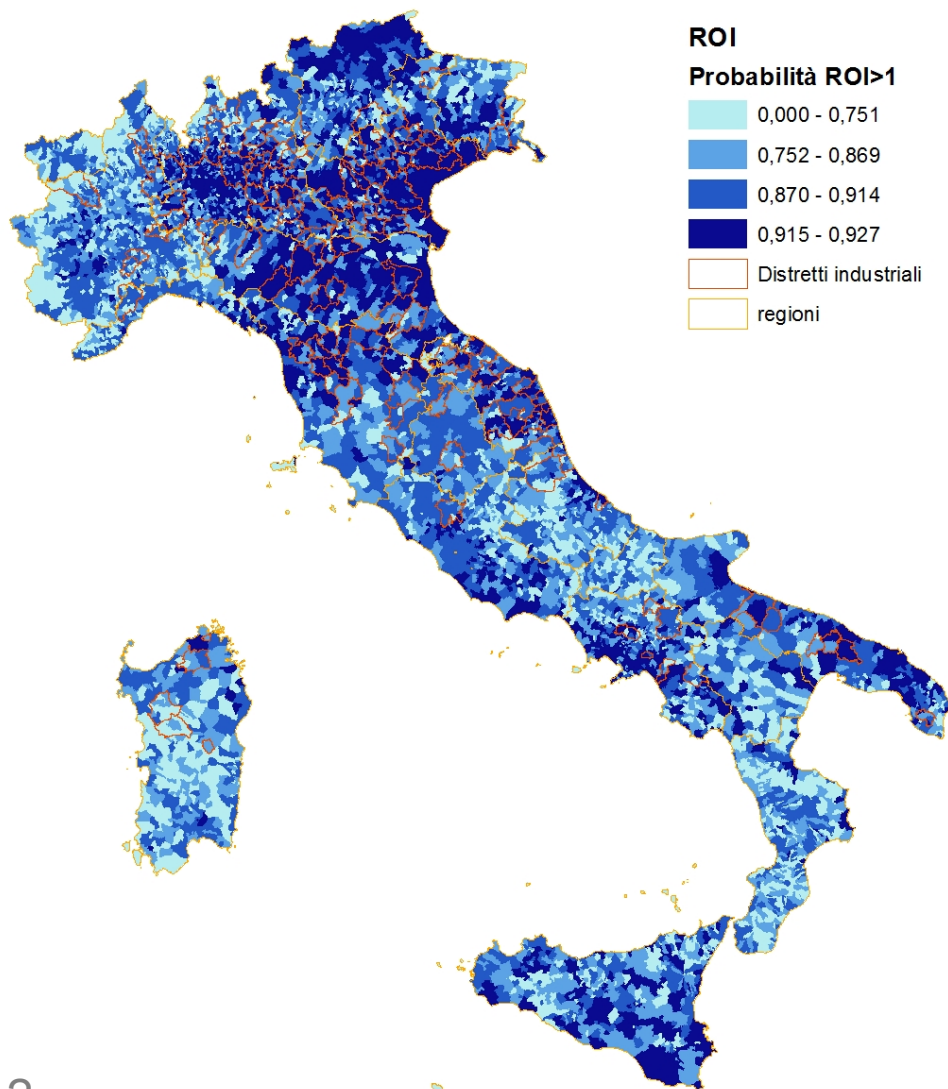
Il modello classifica correttamente circa il 83% dei comuni.
Inoltre, la capacità di discriminare correttamente è maggiore
per i comuni con $ROI \geq 0$.

Osservato		Previsto		
		DROE11		Percentuale corretta
		ROI \leq 0	ROI $>$ 0	
DROE11	ROI \leq 0	774	1085	41,6
	ROI $>$ 0	309	5895	95,0
Percentuale globale				82,7

Metodologia di analisi: distribuzione dei residui



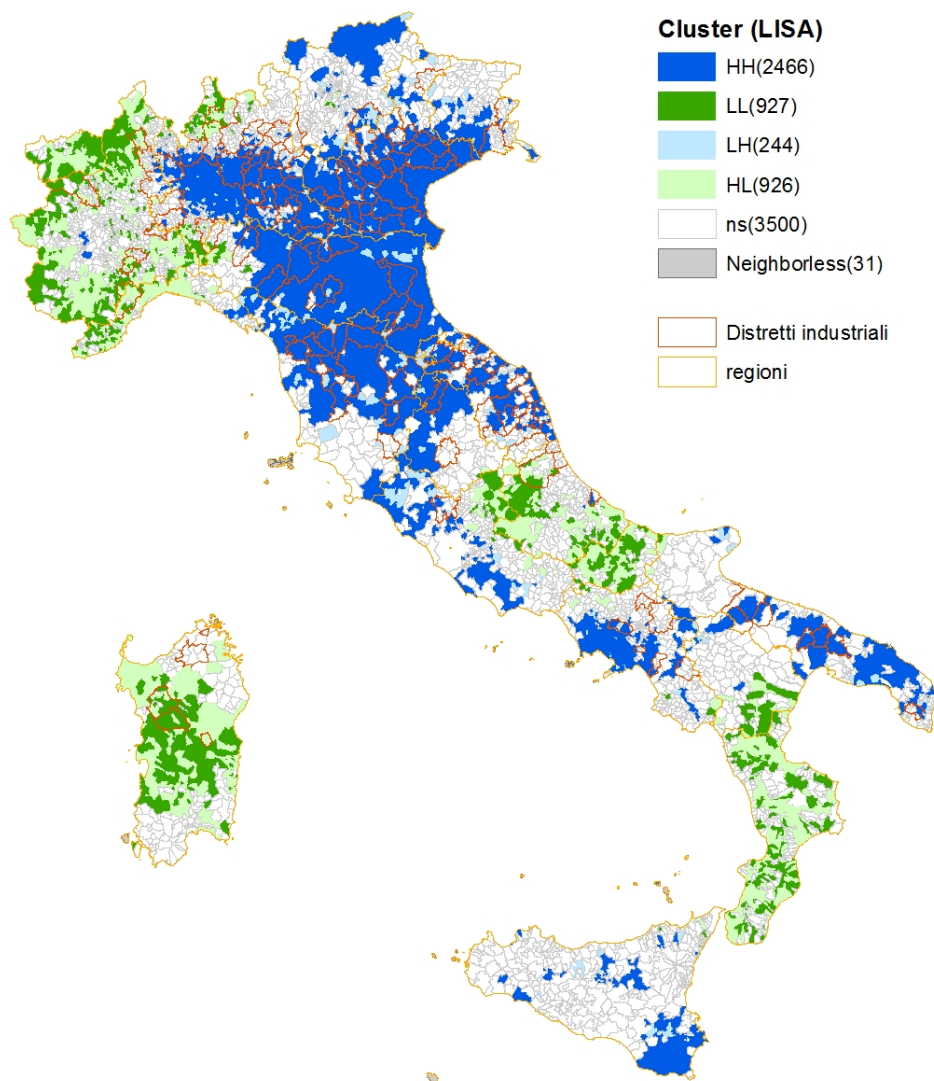
Metodologia di analisi: probabilità



$$I = \frac{n}{\sum_{i=1}^n \sum_{j=1}^n w_{ij}} \frac{\sum_{i=1}^n \sum_{j=1}^n (x_i - M)(x_j - M)w_{ij}}{\sum_{i=1}^n (x_i - M)^2}$$

I= 0,43; p-value<0,001

Metodologia di analisi: cluster



L'indice locale registra il contributo di ciascun comune alla formazione dell'indice globale.

HH. Comuni contigue con alti valori di $\text{prob}(\text{ROI} > 1)$;

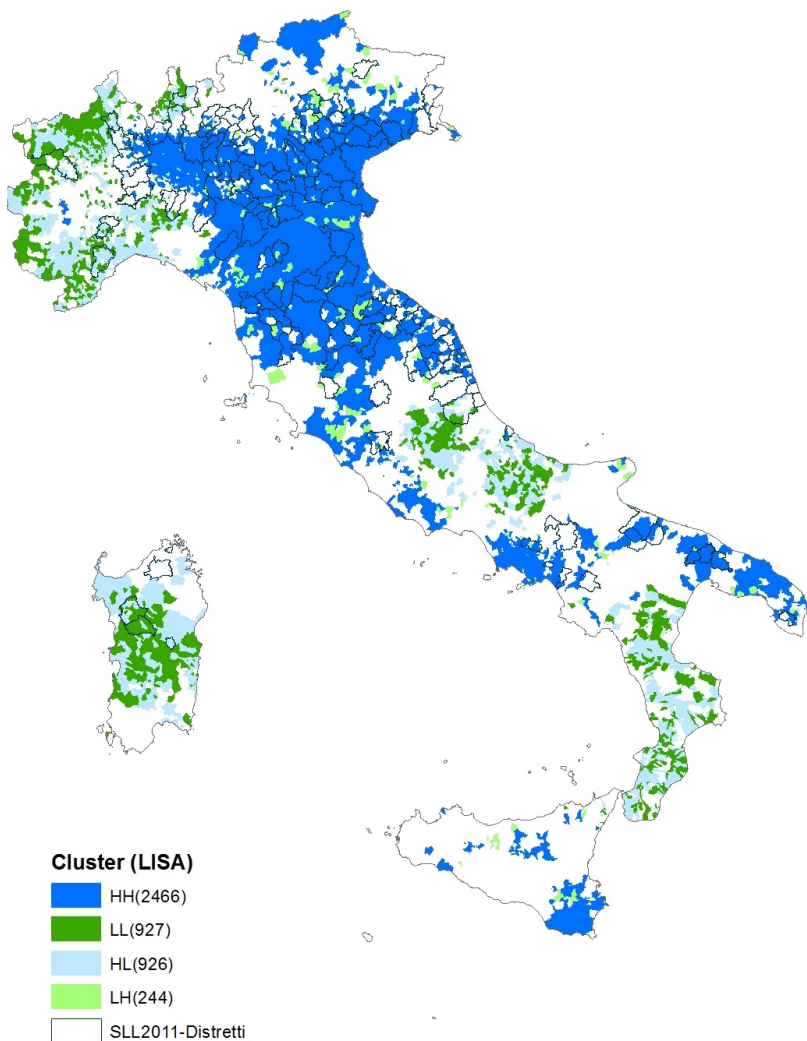
LL. Comuni contigue con valori bassi di $\text{prob}(\text{ROI} > 1)$;

LH. Comuni con valori alti contigue a comuni con valori bassi di $\text{prob}(\text{ROI} > 1)$;

HL. Comuni con valori bassi contigue a comuni con valori alti di $\text{prob}(\text{ROI} > 1)$;

$$I_i = \frac{x_i - M}{S_x^2} \sum_{j=1}^n (x_j - M) w_{ij}$$

Analisi dei risultati: cluster e distretti



CLUSTER	comuni distrettuali	comuni	comuni distrettuali su totale comuni (%)	Distribuzione distretti (%)
HH	1.129	2.458	45,9%	55,5%
HL	108	917	11,8%	5,3%
LH	89	225	39,6%	4,4%
LL	41	616	6,7%	2,0%
ns	666	3.248	20,5%	32,8%
TOTALE	2.033	7.464		100,0%

Analisi multivariata: *Indicatori di bilancio*

Indicatori di redditività

ROI = Reddito Operativo / Capitale Investito

ROI = Reddito Operativo / Totale Attivo

ROE = Utile d'esercizio / Capitale proprio

Redditività Lorda = MOL / Fatturato (Valore Aggiunto)

ROD = Oneri finanziari / Capitale di credito

Indicatori patrimoniali

Patrimonio Netto / Totale Fonti di Finanziamento

Elasticità patrimoniale = Attivo Corrente / Attivo Fisso

Leverage = Capitale di Credito / Capitale Proprio

Indicatori di solvibilità

PFN = Debiti Finanziari (a Breve e Medio-Lungo Termine) – (Disponibilità Immediate + Differite)

Indicatore di sostenibilità = PFN / EBITDA

Indicatori di produttività

Produttività media del lavoro = Valore Aggiunto / Numero addetti

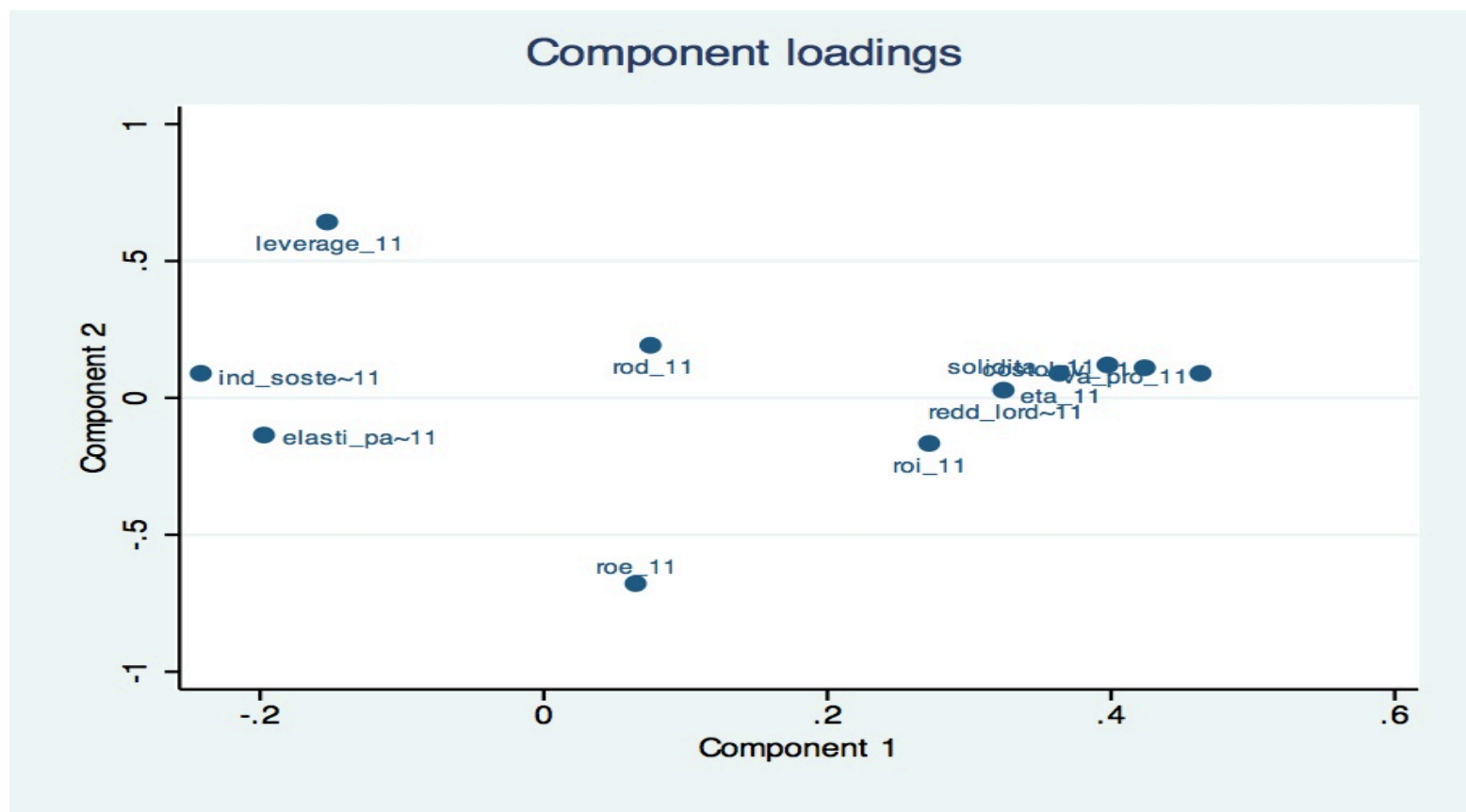
Costo medio del lavoro = Costo del Personale / Numero Dipendenti

Analisi multivariata: *Indicatori di bilancio*

	VA pro	MOL/ VA	roe	roi	rod	leverag e	solidità patr.le	elasticità patr.le	indice sost.tà	Età impres a	costo lavoro pro
Valore Agg.to pro-capite	1,00										
MOL/VA	0,60	1,00									
Roe	0,01	0,01	1,00								
Roi	0,41	0,45	0,16	1,00							
Rod	0,02	0,06	-0,06	-0,08	1,00						
leverage	-0,11	-0,04	-0,85	-0,13	0,01	1,00					
solidità patrimoniale	0,54	0,30	-0,06	0,25	0,17	-0,15	1,00				
elasticità patrimoniale	-0,22	-0,15	0,00	0,08	-0,37	0,12	-0,38	1,00			
indice sostenibilità	-0,26	-0,32	-0,03	-0,43	0,07	0,07	-0,33	-0,11	1,00		
Età impresa	0,50	0,15	0,02	0,16	0,10	-0,16	0,46	-0,24	-0,14	1,00	
costo lavoro pro-capite	0,77	0,27	0,00	0,17	0,07	-0,14	0,52	-0,22	-0,20	0,66	1,00

Media coefficiente di correlazione: 0,12

Analisi multivariata: *Analisi in Componenti principali*



Analisi spaziale: *Spatial-Lag Model*

Obbiettivo: individuare effetti spaziale e temporale negli anni 2004-2010 sulla PRODUTTIVITA' 2011?? Dei SLL, individuata dalla prima dimensioni fattoriale.

Risultato: il modello che meglio si adatta ai dati empirici è

Spatial-Lag Model (Anselin, 1980,1988,1994,2002) :

$$y = \rho W_1 y + X\beta + \varepsilon \quad [1]$$

$$\varepsilon = \lambda W_2 \varepsilon + u \quad [2]$$

dove:

y è il vettore $N \times 1$ delle osservazioni della variabile dipendente (N è il numero delle unità geografiche);

W_1 W_2 sono le matrici di contiguità $N \times N$ associate alle variabili ritardate e di disturbo;

X è una matrice $N \times k$ di osservazioni delle variabili esplicative;

ρ e λ sono i due coefficienti di autoregressione (scalari) associati rispettivamente alla variabile ritardata Wy e al fattore di disturbo ε ;

β , sono vettori $k \times 1$ dei parametri di regressione e autoregressione associati rispettivamente alle variabili esplicative X ;

u è il termine di errore ed ha distribuzione normale con media 0 e matrice diagonale di covarianza Ω .

Identificazione e stima del modello

	Test	Value	prob
Lagrange Multiplier (lag)	1	54,36097	0,00000
Robust LM (lag)	1	52,02336	0,00000
Lagrange Multiplier (error)	1	1,89285	0,16888
Robust LM (error)	1	10,35524	0,00129
Lagrange Multiplier (SARMA)	2	64,71621	0,00000

La corretta specificazione del modello è testimoniata dalla relazione tra le statistiche Word (Wd), Likelihood Ratio (LR) e Lagrange Multiplier (LM) che assumono rispettivamente i seguenti valori:

$$W=55,2049 > LR=52,45 > LM=52,02$$

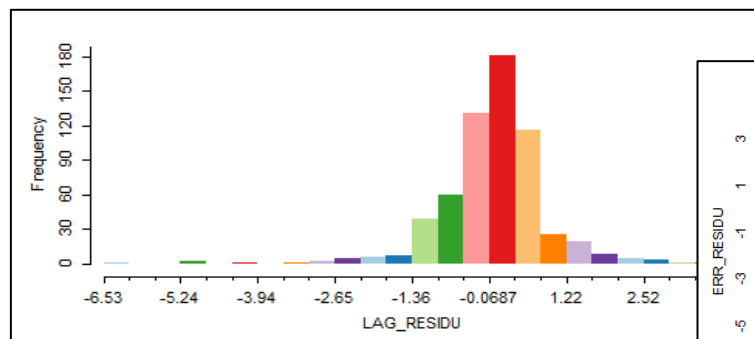
Variable	Coefficient	Std.Error	z-value	Probability
WA1_11	0,2307	0,0310	7,4307	0,0000
CONSTANT	0,0018	0,0364	0,0503	0,9599
A1_10	0,4405	0,0388	11,3408	0,0000
A1_09	0,1039	0,0424	2,4500	0,0143
A1_08	0,2838	0,0342	8,3044	0,0000
Bontà di adattamento			OLS	Spatial-Model
likelihood			- 831,509	- 805,283
Akaike info criterion dello Spatial-Lag			1671,02	1620,57

La componente spaziale che registra l'effetto *spillover* è significativa, così pure la Produttività 2010, 2009 e 2008.

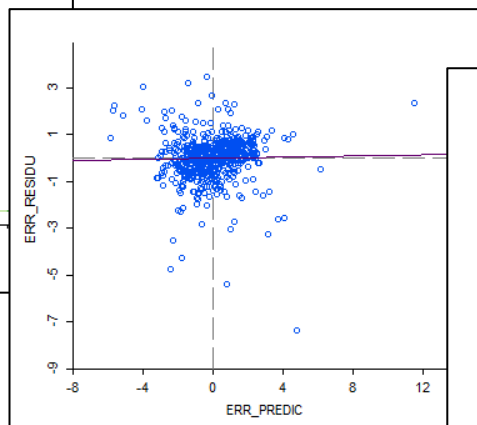
Statistiche bontà adattamento

Diagnostiche

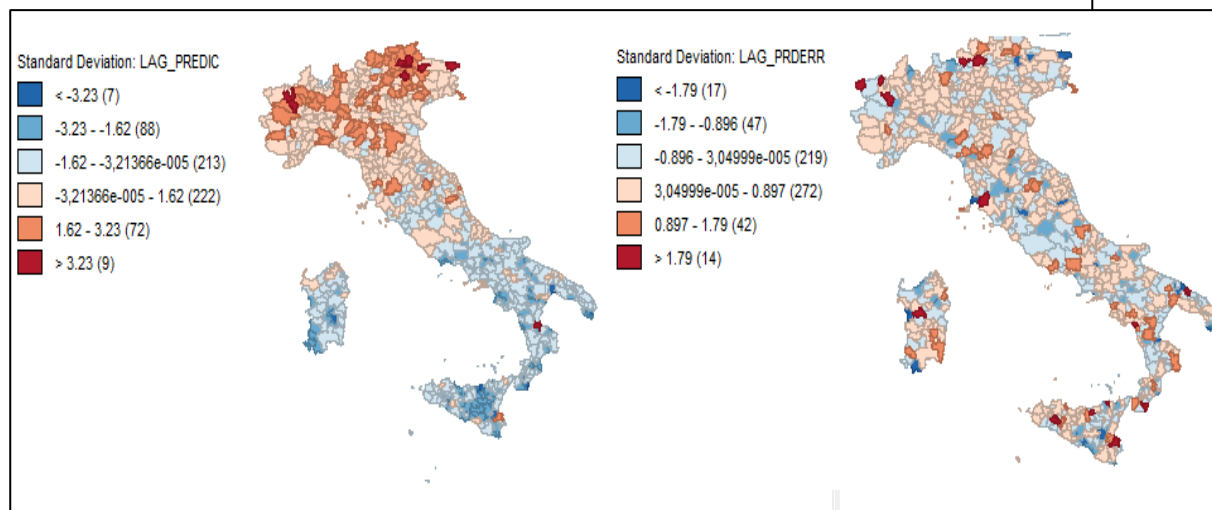
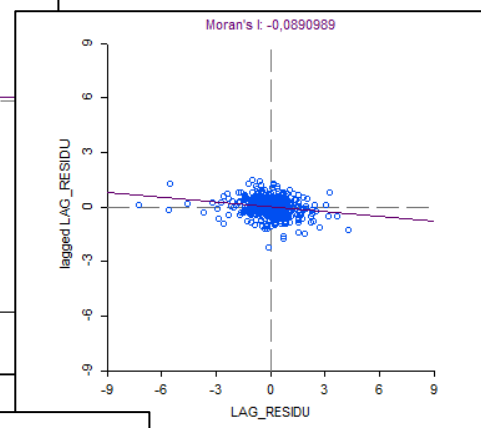
Residui



Residui*Predetti



Autocorrelazione spaziale residui



I risultati dell'analisi spaziale svolta sulla prima Componente individuata ("Produttività") confermano quelli dell'analisi del ROI su base comunale

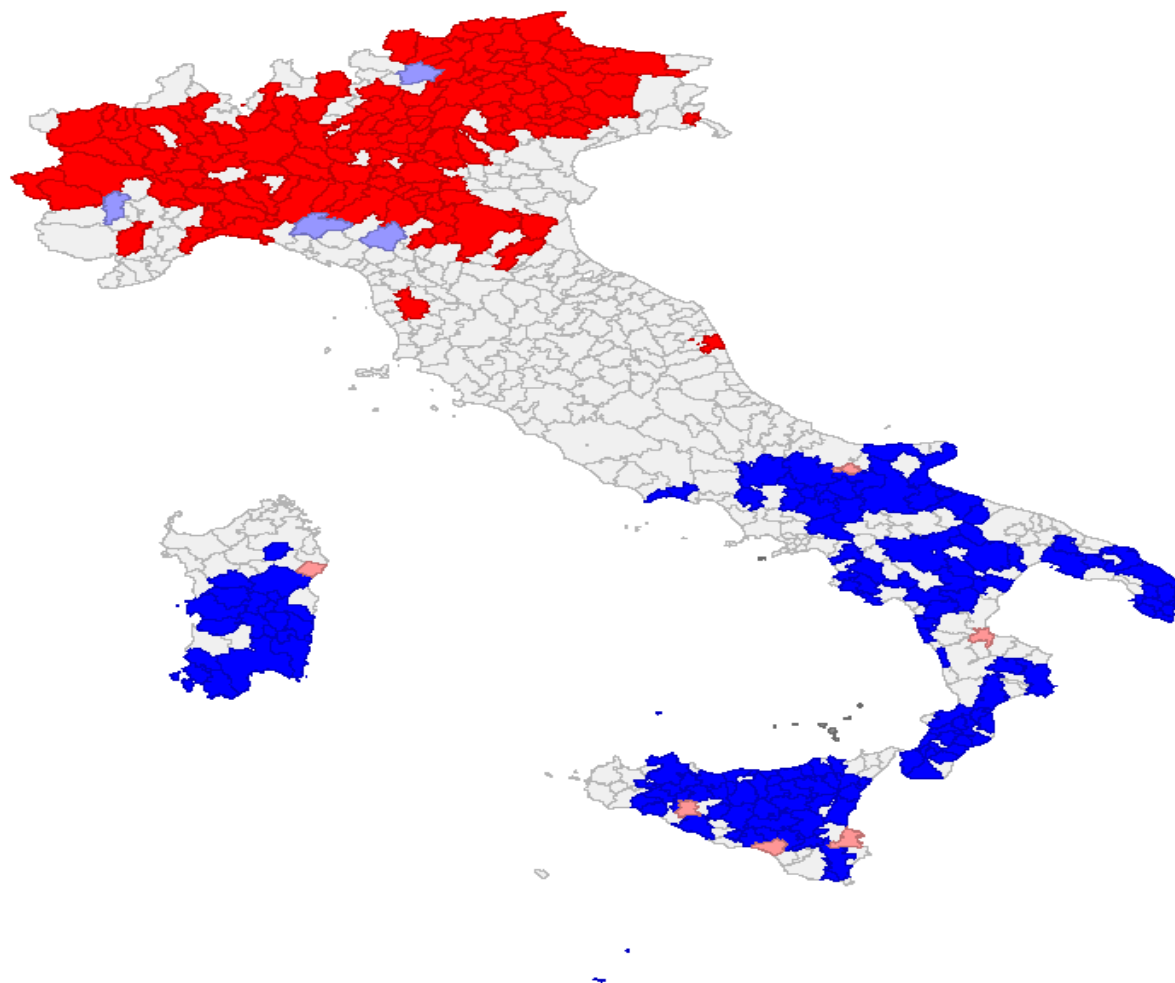
gli effetti spaziali risultano maggiormente significativi.

Gli effetti spaziali e temporali non risultano significativi sulla seconda componente (Profittabilità complessiva) sulla quale influiscono le mutevoli condizioni finanziarie esogene.

Analisi dei risultati: valori predetti

LISA Cluster Map: WXLOG,

- Not Significant (328)
- High-High (135)
- Low-Low (136)
- Low-High (4)
- High-Low (6)
- Neighborless (2)

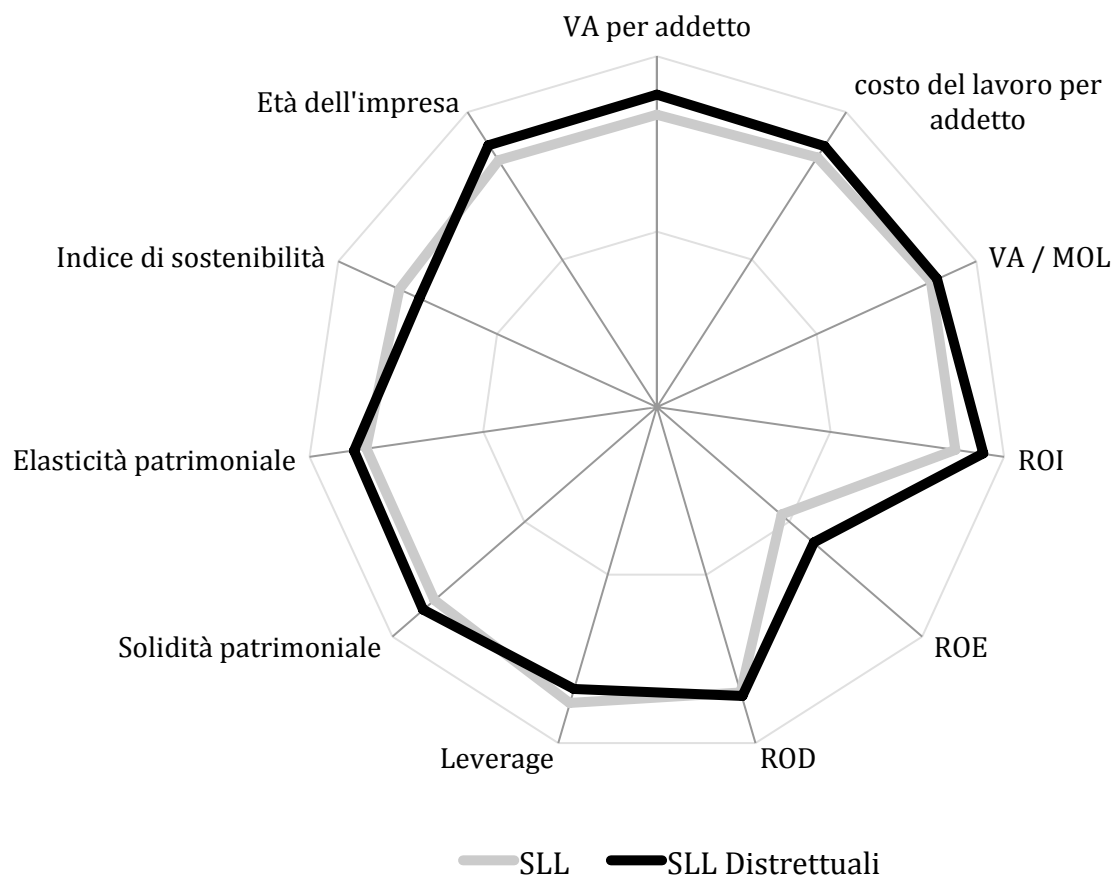


Analisi dei risultati: dati descrittivi

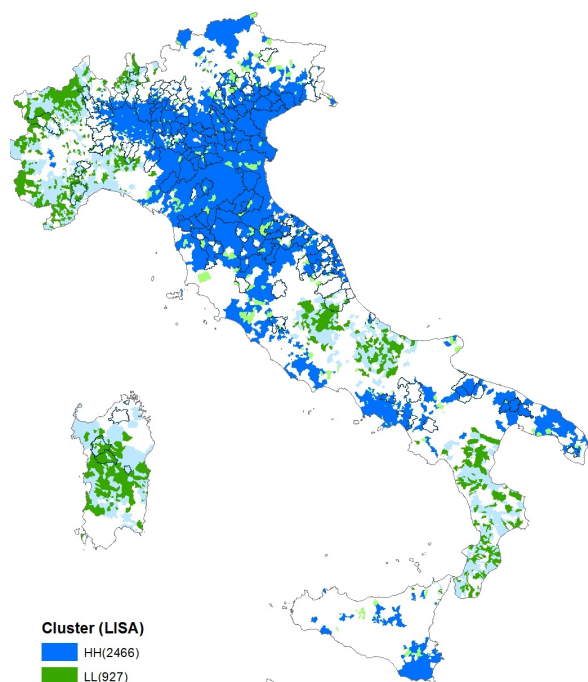
moran	Numero SLL	VA per addetto	costo del lavoro per addetto	VA / MOL	ROI	ROE	ROD	Leverage	Solidità patrimoniale	Elasticità patrimoniale	Indice di sostenibilità	Età dell'impresa
high-high	128	€ 55.966	€ 37.894	35,49%	2,98%	1,59%	1,93%	3,26	32,1%	1,37	4,50	14,80
no significative	330	€ 42.037	€ 31.311	31,16%	2,09%	4,28%	2,01%	3,77	26,0%	1,50	7,04	11,69
low-low	141	€ 33.216	€ 25.495	28,87%	1,80%	-6,92%	1,83%	5,25	21,8%	1,85	8,13	9,77
Totale	599	€ 42.937	€ 31.349	31,55%	2,21%	1,07%	1,95%	4,01	26,3%	1,56	6,75	11,90

	moran	Numero SLL	VA per addetto	costo del lavoro per addetto	VA / MOL	ROI	ROE	ROD	Leverage	Solidità patrimoniale	Elasticità patrimoniale	Indice di sostenibilità	Età dell'impresa
Distretti Industriali	high-high	47	€ 53.795	€ 38.261	33,29%	2,62%	1,43%	1,85%	3,23	31,0%	1,57	4,25	14,2
	no significative	87	€ 45.868	€ 33.572	32,93%	2,57%	1,10%	2,03%	3,58	28,0%	1,57	4,84	12,66
	low-low	7	€ 32.493	€ 25.114	30,53%	2,39%	1,13%	1,66%	4,88	20,5%	1,94	6,10	9,7
NO Distretti Industriali	high-high	81	€ 53.835	€ 37.609	34,61%	2,96%	2,74%	1,91%	3,12	32,0%	1,27	4,11	14,3
	no significative	243	€ 39.220	€ 30.275	31,24%	2,11%	-0,35%	1,96%	4,02	24,8%	1,41	6,06	10,98
	low-low	134	€ 32.945	€ 25.550	29,43%	1,96%	0,13%	1,82%	4,57	21,9%	1,64	6,23	9,7

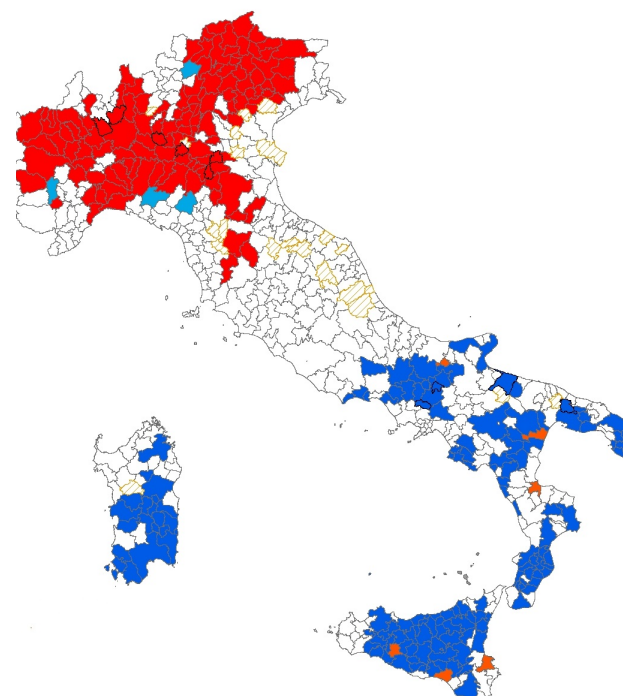
Analisi dei risultati: differenze fra non classificati



Analisi dei risultati: cluster e distretti



Cluster (LISA)
HH(2466)
LL(927)
HL(926)
LH(244)
SLL2011-Distretti



LISA
HH
LL
LH
HL
Neighborless
No Sign. No Distretti
No Sign. Distretti
Sign. Distretti

Analisi dei risultati: incrocio con i dati censuari

CLUSTER	exp_fatt	addetti	mkt	innproc	innprod	innorg	innmark
High-High	12,18	46,63	Italia	0,23	0,28	0,28	0,24
Nosignal-ID	14,67	34,17	UE	0,23	0,27	0,26	0,23
Nosignal-NoID	5,61	38,93	Italia	0,18	0,20	0,24	0,19
Low-Low	2,86	32,49	Regione	0,16	0,16	0,23	0,17
Total	9,80	41,42	Italia	0,21	0,25	0,26	0,22

Analisi dei risultati: investimenti e clusters

	INVESTIMENTI su TOTALE ATTIVO	INVESTIMENTI su Patrimonio	INVESTIMENTI su RISULTATO ANTE IMPOSTE	INVESTIMENTI su UTILE NETTO
High-High	1,76	4,76	81,32	1,98
Nosignal-ID	1,9	5,78	96,8	3,21
Nosignal-NoID	1,44	4,75	120,8	10,4
Low-Low	-30,15	-113,58	-5299,23	70,07
Totale	0,32	0,94	17,95	0,56

Considerazioni

- Gli effetti temporali predominano su quelli spaziali. Tuttavia gli effetti spaziali sono significativi, soprattutto per le aree con migliore/peggiore performance
- I distretti riproducono i cluster identificati nelle due analisi, e le aree distrettuali rappresentano circa il 45% dei cluster di maggior successo.
- L'appartenenza a un distretto non discrimina fra SLL High-High e Low-Low, ma discrimina invece fra gli SLL in cui non sono rilevanti gli effetti spaziali (non clusterizzati)
- Gli effetti sulla performance delle imprese delle variabili strategiche (rilevate dal Censimento industriale) variano a seconda dei cluster, come pure la quota degli investimenti sul totale dell'attivo e degli utili

Sviluppi in corso

Costruzione di basi dati da Big Data, fungibili rispetto a molteplici scopi.

- Dati amministrativi
- Produzione di dati da banche dati/servizi web
- Informazioni da indagini statistiche (statistical matching / stime per domini non pianificati)

Geolocalizzazione: Standardizzazione degli indirizzi

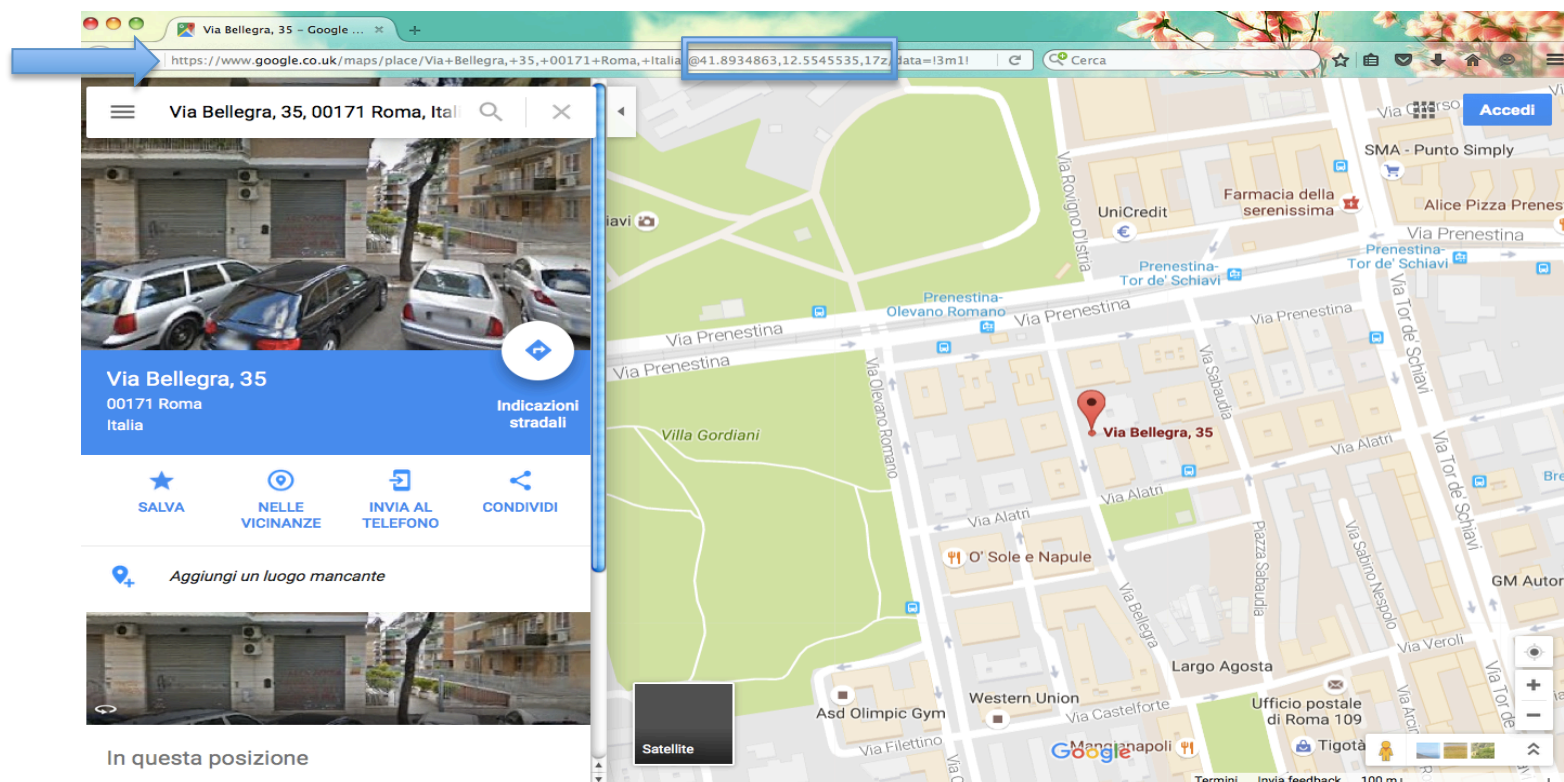
DUG - Denominazione Urbanistica Generica, ossia il qualificatore del toponimo (es. Via, Piazza, Largo, ecc.);

Il toponimo, ossia l'indirizzo (es: [via] dei Querceti);

il civico relativo all'impresa nel toponimo.

DUG	Toponimo	Civico	Den. Comune	CAP	Sigla Prov.
Via	Bellegra	35	Roma	00171	RM

Geolocalizzazione: referenziazione degli indirizzi



Geolocalizzazione: referenziazione degli indirizzi

Esempio *response* json

```
{
  "results" : [
    {
      "address_components" : [
        {
          "long_name" : "22",
          "short_name" : "22",
          "types" : [ "street_number" ]

          "long_name" : "Viterbo",
          "short_name" : "Viterbo",
          "types" : [ "locality", "political" ]

          "long_name" : "Provincia di Viterbo",
          "short_name" : "VT",
          "types" : [ "administrative_area_level_2"

          "long_name" : "Italia",
          "short_name" : "IT",
          "types" : [ "country", "political" ]

          "location" : {
            "lat" : 42.4259435,
            "lng" : 12.0925167

            "long_name" : "Via Luigi Galvani",
            "short_name" : "Via Luigi Galvani",
            "types" : [ "route" ]

            "long_name" : "Viterbo",
            "short_name" : "Viterbo",
            "types" : [ "administrative_area_level_3", "political" ]

            "long_name" : "Lazio",
            "short_name" : "Lazio",
            "types" : [ "administrative_area_level_1", "political" ]

            "long_name" : "01100",
            "short_name" : "01100",
            "types" : [ "postal_code" ]

            "partial_match" : true,
            "place_id" : "types" : [ "street_address" ]
          }
        }
      ]
    }
  ]
  "status" : "OK"
```

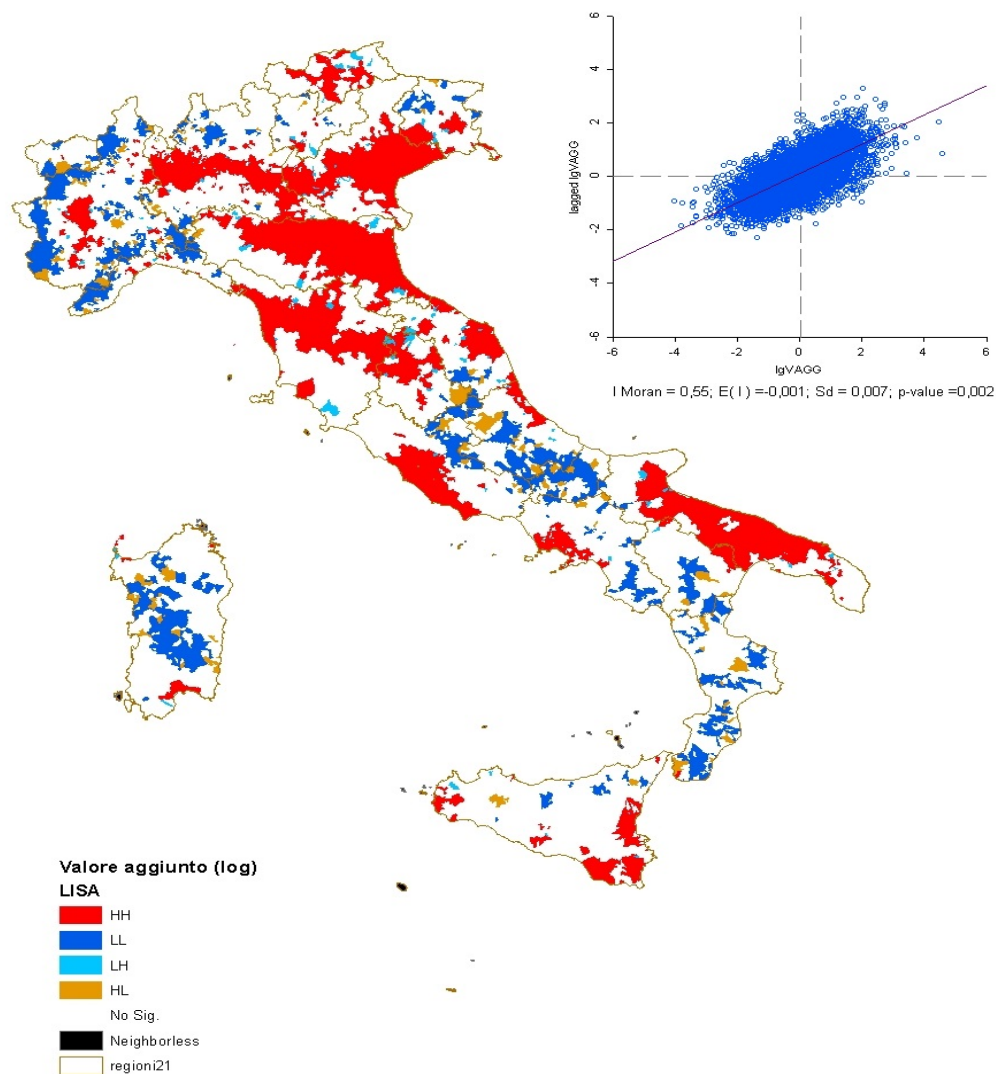

Geolocalizzazione: routing

The screenshot displays the Google Maps routing interface. The sidebar on the left lists three routes from Via Bellegra, 35 to Via Giuseppe Antonio Andriulli:

- tramite Viale della Serenissima**: 7 min, 1,7 km (highlighted with a blue box and a blue arrow pointing to it).
6 min senza traffico
[DETTAGLI](#)
- tramite Via Prenestina**: 8 min, 2,3 km
7 min senza traffico
- 18:01–18:14**: 13 min
542 544 548

The map on the right shows the route from Via Bellegra, 35 to Via Giuseppe Antonio Andriulli. The route is highlighted in blue. The map includes labels for various streets, landmarks, and public transport stops. A scale bar at the bottom right indicates 200 m.

Geolocalizzazione: definizione dei cluster



Geolocalizzazione: analisi dei risultati

Buona precisione della geolocalizzazione (casi Missing/indirizzo non geolocalizzabile **0,03%**; casi dubbi /localizzazione al di fuori del territorio della regione teorica di appartenenza **0,5%**; casi approssimati/al baricentro del comune **1,1%**)

I risultati della clusterizzazione confermano quelli delle analisi precedenti.

Conclusioni

Dalle analisi finora svolte emerge che:

- l'allungamento delle catene globali del valore non ha ancora eliminato il lascito positivo sulla produttività e la redditività delle imprese esercitato dalla formazione e dal consolidamento dei distretti industriali
- sia che si tratti di imprese appartenenti all'industria caratteristica dei distretti, sia che si tratti di imprese localizzate nel distretto, ma non appartenenti a tale industria
- tali effetti assumono più rilevanza quando si analizza l'efficienza delle imprese al netto dell'influsso delle condizioni finanziarie.

GRAZIE DELL'ATTENZIONE!