

Vulnerabilità del territorio e criminalità organizzata

Andrea Mazzitelli¹

¹ Il presente lavoro è stato realizzato all'interno del progetto Unioncamere-Universitas Mercatorum "SOS Legality – Seized Business and goods from mafia to strategically prevent crimes and promote legality through socio-economic development" (finanziato dalla Commissione Europea)

Sommario

Il presente documento si pone l'obiettivo di analizzare come si configurano i comportamenti criminali di natura mafiosa che tentano di infiltrarsi nell'economia legale. La conoscenza di quanto accade nel proprio territorio è determinante per indirizzare meglio le politiche di prevenzione nella lotta contro la criminalità organizzata, anche quando si tratta di un fenomeno complesso e per certi aspetti poco visibile come quello dell'infiltrazione nell'economia virtuosa del territorio. L'analisi condotta ha portato alla selezione e individuazione di specifici indicatori riguardanti la vulnerabilità delle infrastrutture, delle imprese, delle famiglie e del territorio, utilizzando *open data*. Il principale risultato conseguito è che il processo di diffusione territoriale della criminalità organizzata è trasversale a tutte le provincie italiane, anzi prevarica i confini amministrativi perché è interprovinciale e interregionale (livello mesoeconomico): ciò consente di definire la vera *armatura mafiosa del territorio* e quindi di individuare delle partizioni territoriali funzionali a una migliore interpretazione della distribuzione lungo la nostra Penisola dei gruppi criminali.

Il problema della valutazione quantitativa del grado di vulnerabilità di un'area geografica è estremamente complesso: oltre alle difficoltà di reperimento dei dati esistono problemi di aggregazione e interpretazione dei risultati. La complessità principale risiede nella multidimensionalità del fenomeno e nella scelta di un'opportuna funzione di aggregazione dei dati. Seguendo l'approccio assiomatico dei numeri indice, si è utilizzato l'indice di *Jevons*, che a differenza di altri numeri indice semplici, soddisfa il superamento di specifici test, ovvero rispetta sia le *condizioni essenziali* che le proprietà derivate o desiderate.

Il ricorso a tecniche di statistica spaziale si è reso necessario per affrontare le problematiche connesse all'impiego di dati georeferenziati, collegati a entità geografiche. Ciò consente di valutare per la presenza di interdipendenza o autocorrelazione spaziale e di eterogeneità spaziale. In particolare, si vuole verificare se le caratteristiche di un determinato fenomeno occorso in una provincia non siano spiegate unicamente da determinanti interne alla stessa ma risentano anche di alcune peculiarità proprie di altre provincie, più o meno vicine.

Abstract

Vulnerability of the territory and organized crime

Objectives

This paper aims to analyse how to configure the behaviour of the Mafioso groups that attempt to infiltrate the legal economy. Awareness of what is happening in its territory is crucial in better prevention policies in the fight against organized crime. The Mafia is often entangled with local and national politics and it hinders the economic and social development of several Italian regions and provinces. Getting to a climate of legality and respect for the rule of law is an

indispensable requirement not only for personal security, but for social development, too. The analysis led to the selection of specific indicators regarding the vulnerability of infrastructure, firms, households and territory, using *open data*. The main achievement is that the process geographical coverage of organized crime affects all Italian provinces, crosses the administrative boundaries and is inter-regional (meso-level): this allows to identify some territorial partitions to a better interpretation of the distribution of criminal groups. The issue of the quantitative assessment of the degree of vulnerability of an area is highly complex: in addition to the difficulty of obtaining data there are problems of aggregation and interpretation of findings.

Methods and Results

Following the axiomatic approach index numbers, we used the Jevons index that unlike other numbers simple index fulfils the overcoming of specific tests, which respects both the essential conditions that the properties arisen or want. Typically, there are two different types of spatial analysis methods. The first is often referred to as Exploratory Spatial Data Analysis (ESDA), which is based on global and local *Moran's I statistics*, and is concerned with the description and exploration of spatial data. The second address refers to the spatial modeling. The spatial models help to assess for the presence of interdependence or spatial autocorrelation and spatial heterogeneity. When we model spatial distribution data, we attempt to predict the outcome variable at each location as a function of variables of the focal location and possibly of variables at other locations as well. In other words, effective prediction requires the development of spatial regression models that can find and incorporate the important associative and causative variables available in the data.

Conclusions

In this work we show that the axiomatic Jevons index is the proper method to detect for the presence of organized crime and vulnerability of the territory. A weight matrix is introduced to specify the spatial relations between all pairs of provinces. Finally, we analyze the SAR and CAR models as a tool for prediction of the location choices of organized crime concerning counterfeit goods *made in Italy*.

Keywords: Jevons index, matrix of vulnerability, organized crime, Economic development, spatial autocorrelation, spatial heterogeneity, Moran index, LISA, spatial regression models, *open data*.

JEL Classification: C21, C31, C43, I32, K42, O18, R12.

Introduzione

La criminalità organizzata ha un'elevata capacità di infiltrarsi nel tessuto economico e sociale, riesce ad instaurare relazioni con la società civile, si alimenta con la collusione e la corruzione. Ne risultano intaccati il comportamento civico, la fiducia, l'onestà intellettuale, le reti di relazione, cioè il capitale sociale di un territorio. Si instaura un sistema di intrecci perverso tra società civile e "società mafiosa" che si autoalimenta e di cui è difficile valutare la complessiva portata (Tarantola, 2012). D'altra parte è evidente quando si analizza la presenza della criminalità organizzata che le regioni meridionali registrano un alto numero di omicidi o di altri reati superiore rispetto alla media nazionale. Sono reati particolarmente complessi e rischiosi che richiedono la collaborazione di molti individui appartenenti anche a associazioni mafiose straniere, specialmente dell'Europa dell'Est e dell'Asia Orientale. Questi crimini infliggono alla società danni economici e sociali ben maggiori che nel caso di reati più semplici, come i furti, con il rischio di sfibrare il tessuto di una società; ciò può mettere a repentaglio la democrazia, frenarla dove debba ancora consolidarsi (Draghi, 2011). Pensare che la Mafia sia un fenomeno tipicamente meridionale e che tocca marginalmente le altre regioni italiane è fuorviante: le opportunità connesse al più veloce sviluppo economico e finanziario del Centro Nord inevitabilmente attraggono l'interesse delle cosche, come già evidenziato nel 1994 dalla Commissione parlamentare di inchiesta sul fenomeno della mafia che dimostrava l'esistenza di "una vastissima ramificazione di forme varie di criminalità organizzata di tipo mafioso, praticamente in tutte le regioni d'Italia" (Commissione Antimafia, 1993). Se nel Centro Nord i crimini che destano maggior allarme sociale, ovvero gli omicidi, sono meno numerosi che al Sud è altrettanto vero che altre tipologie di crimine sono emerse negli ultimi anni, soprattutto legati ai danni procurati all'ambiente: si tratta dei reati di stampo mafioso connessi al ciclo del cemento e al ciclo dei rifiuti, in cui spiccano principalmente le province del Centro Nord. Se nelle principali regioni del Nord, *in primis* la Lombardia, la Mafia appare un fenomeno d'importazione, dato che la maggior parte delle denunce di stampo mafioso ha riguardato individui provenienti dalla Sicilia, dalla Calabria, e dalla Campania è altrettanto vero che la stessa criminalità locale appare coinvolta in molti reati tipicamente riconducibili al crimine organizzato di stampo mafioso, come l'usura, il riciclaggio e le estorsioni. Ne emerge una preoccupante saldatura con le mafie tradizionali (Draghi, 2011).

Misurare la rilevanza economica delle attività criminali è tuttavia assai complesso, in particolare quando ciò riguarda i costi della criminalità, il fatturato delle organizzazioni criminali e la loro infiltrazione nei mercati legali. Le statistiche ufficiali prodotte dall'ISTAT, dal Ministero dell'Interno, dal Ministero di Giustizia, dalla Banca d'Italia, dall'UIF (Unità d'Informazione Finanziaria) e che forniscono informazioni sulla criminalità derivano da quanto è stato scoperto dall'attività delle Forze dell'Ordine e, quindi, ne rappresentano una sottostima. Ad esempio, quando si vuole calcolare il fatturato dell'economia criminale, non sempre i dati sono attendibili e completi: nel caso delle indagini di vittimizzazione c'è sempre il rischio di omertà da parte degli imprenditori o dei cittadini onesti che vogliono evitare ulteriori ritorsioni e minacce da parte delle associazioni criminali con la conseguenza che la compilazione delle graduatorie delle province più mafiose risulta distorta. Parallelamente, la stima del traffico di droga o di prostituzione solleva numerosi dubbi riguardo alla loro attendibilità. È difficile che una vittima di tali attività possa denunciare il reato. Le denunce dipendono quindi dalla maggiore o minore capacità e attività investigativa delle forze dell'ordine nonché dall'efficacia delle politiche di contrasto.

Il lavoro è così strutturato: nel paragrafo 1 sono descritti gli indicatori di vulnerabilità del territorio; nel paragrafo 2 si introduce l'indice di sintesi, ovvero l'indice di Jevons, secondo l'approccio assiomatico della teoria dei numeri indice, per tener conto della multidimensionalità del fenomeno della vulnerabilità a livello provinciale. Il paragrafo 3 è dedicato alla descrizione degli indici di autocorrelazione spaziale sia globali che locali. I paragrafi 4 e 5 evidenziano i principali risultati conseguiti prima con l'applicazione dell'indice semplice di Jevons e successivamente mediante l'introduzione dei test globali e locali di autocorrelazione spaziale per individuare le aree strutturali sia della vulnerabilità che della criminalità. L'ultima sezione, paragrafi 6 e 7, analizza i modelli regressivi spaziali, vale a dire i modelli SAR e CAR, come strumento di previsione circa le scelte localizzative della criminalità organizzata in tema di contraffazione dei beni del *made in Italy*. In particolare, utilizzando all'interno di tali modelli alcune variabili (predittori) della matrice di vulnerabilità, si verificherà quali interazioni spaziali esistono tra la contraffazione dei beni e la vulnerabilità socio-economica del territorio. I risultati delle applicazioni sono riportati in Appendice.

1. La selezione degli indicatori

Obiettivo del presente lavoro non è quantificare i costi della criminalità organizzata, ovvero il PIL mafioso prodotto illegalmente attraverso il gioco d'azzardo, il traffico di droga, il traffico di esseri umani, la prostituzione, l'usura e il pizzo bensì osservare quali territori siano più appetibili e competitivi per la Mafia (sia nazionale che internazionale), vale a dire più vulnerabili per la società civile. In quest'ottica, la selezione degli indicatori di vulnerabilità a livello provinciale è stata condotta con l'intento di individuare le principali criticità del territorio che impediscono uno sviluppo economico dello stesso in termini di competitività e attrattività. La selezione ha portato all'individuazione di cinque macro-indicatori, ognuno dei quali ulteriormente suddiviso in k componenti, come di seguito elencato²:

- 1. Indicatori di vulnerabilità infrastrutturali:** Indice di dotazione infrastrutture di trasporto; Indice di dotazione infrastrutture servizi alle imprese; Indice di dotazione infrastrutture banda larga; Indice di dotazione infrastrutture culturali.
- 2. Indicatori di criminalità organizzata:** Indice di reati del ciclo del cemento; Indice di reati del ciclo dei rifiuti; Indice di criminalità organizzata ponderato con le pene edittali.
- 3. Indicatori spia di infiltrazione dell'illegalità economica (o della criminalità organizzata):** Indice di contraffazione; Indice di usura ed estorsione; Indice di riciclaggio; Indice di intimidazione;
- 4. Indicatori di vulnerabilità delle imprese:** Indice di sofferenza delle imprese; Indice di propensione all'export; Indice procedure concorsuali su totale imprese; Indice scioglimenti/liquidazioni su totale imprese;

² Per gli indicatori di vulnerabilità infrastrutturale la fonte è Tagliacarne (2012); Per gli Indici di reati del ciclo del cemento e del ciclo dei rifiuti la fonte è Legambiente (2013). Per l'Indice di criminalità organizzata la fonte è ISTAT e per gli Indicatori spia di criminalità organizzata la fonte è ISTAT (2012). Per gli Indici di sofferenza delle imprese e quota impieghi immobili uso produttivo la fonte è Banca d'Italia (2012); per l'Indice propensione all'export la fonte è Unioncamere (2012); per gli Indici procedure concorsuali su totale imprese e scioglimenti/Liquidazioni su totale imprese la fonte è Infocamere (2012); per l'Indice quota previsione di assunzioni di personale *high-skill* la fonte è Unioncamere-Ministero del Lavoro (2012). Per gli Indici tasso di disoccupazione, tasso di disoccupazione giovanile e quota popolazione con titolo universitario su totale popolazione la fonte è ISTAT (2012). Per gli Indici credito al consumo delle famiglie pro-capite/patrimonio pro-capite e sofferenza delle famiglie pro-capite le fonti sono Banca d'Italia (2012) e ISTAT (2012); per l'Indice quota impieghi immobili uso residenziale la fonte è Banca d'Italia (2012); per l'Indice arrivi stranieri su popolazione residente la fonte è ISTAT (2012); per l'Indice quota occupati industria culturale su totale economia la fonte è Unioncamere-Symbola (2012).

Indice quota impieghi immobili uso produttivo; Indice quota previsione di assunzione di personale *high skill*

- 5. Indicatori di vulnerabilità delle famiglie:** Indice tasso di disoccupazione; Indice tasso di disoccupazione giovanile; Indice credito al consumo delle famiglie pro-capite su patrimonio pro-capite; Indice di sofferenza delle famiglie pro-capite; Indice quota impieghi immobili uso residenziale; Indice arrivi stranieri su popolazione residente; Indice quota popolazione con titolo universitario su totale popolazione; Indice quota occupati industria culturale su totale economia.

2. L'indice di sintesi

Come già introdotto, il problema della valutazione quantitativa del grado di vulnerabilità (competitività) di un'area geografica è estremamente complesso: oltre alle difficoltà di reperimento dei dati esistono problemi di aggregazione e interpretazione dei risultati³. La complessità principale risiede nella multidimensionalità del fenomeno, la misurazione del quale richiede, inizialmente, il superamento di ostacoli di natura concettuale e definitoria e, successivamente, la scelta, non banale, tra il limitarsi a fornire una misura di natura analitica, rappresentata da un sistema di indicatori semplici, oppure costruire una misura sintetica che, mediante un'opportuna funzione di aggregazione sia capace di raccogliere i molteplici aspetti del fenomeno oggetto di studio (Mazziotta et al., 2012). Tale funzione deve essere in grado di cogliere le variazioni territoriali (oltre che temporali) con implicazioni di natura spaziale. Procedendo in tale direzione, per ogni macro-indicatore si è calcolato il relativo indice di sintesi: l'indice scelto è quello di Jevons (rapporto di medie geometriche semplici). Nelle analisi di concentrazione dei fenomeni socio-economici, la media geometrica è una delle tecniche più usate nella sintesi degli indicatori, in quanto rappresenta una soluzione intermedia tra metodi compensativi, come la media aritmetica, e metodi non-compensativi, come l'analisi multicriteria (OECD,

³ Per completezza di informazione gli indicatori territoriali relativi alle province di più recente costituzione (Monza e della Brianza, Fermo, Barletta-Andria-Trani, Carbonia Iglesias, Medio Campidano, Ogliastra, Olbia) possono sottostimare i fenomeni analizzati semplicemente perché il reato commesso o la quota di vulnerabilità associata alle suddette aree non sempre viene attribuita alle nuove province ma è contemplata, in alcune rilevazioni, all'interno delle vecchie province da cui si sono separate amministrativamente. Per armonizzare i dati sono state eseguite delle operazioni di ripartizioni per quote provinciali.

2008). L'indice di Jevons è stato applicato a un insieme di indicatori di vulnerabilità, rilevati a livello provinciale, in campo economico, sociale, culturale e ambientale. Seguendo l'approccio assiomatico dei numeri indice, l'indice di Jevons, a differenza di quelli di Dutot e di Carli, soddisfa il superamento di specifici test, ovvero rispetta sia le "condizioni essenziali" che le proprietà derivate o desiderate (Eichhorn-Voeller, 1976; Diewert, 1976, 1995; Martini, 1992, 2001). Per definire un numero indice si devono chiarire quali siano le "condizioni di equivalenza" che si intendono rispettare: queste condizioni non devono essere verificate *a posteriori* ma chiarite *a priori*, nella definizione stessa di numero indice. Occorre cioè passare dall'impostazione dei *mechanical tests a posteriori* a un'impostazione *assiomatica* che fissi *a priori* le condizioni da rispettare. Alla luce di questa impostazione, non è lecito, quindi, definire il numero indice come media, senza specificare le condizioni di equivalenza che attribuiscono significato alla nozione stessa di media. L'indice di vulnerabilità provinciale complessivo sarà la media geometrica dei cinque macro-indicatori di Jevons. D'altra parte, l'utilizzo della media geometrica come indice di sintesi non ammette compensazione tra i diversi valori ottenuti, in quanto assume che ciascuna componente della vulnerabilità del territorio non sia sostituibile, o lo sia solo in parte, con le altre componenti. I valori ottenuti consentono di classificare le province in base al loro livello di vulnerabilità (superiore o inferiore alla media) rispetto all'anno di osservazione: lo strumento proposto può costituire un valido ausilio per la misura della vulnerabilità per qualsiasi scala territoriale scelta. La metodologia si sviluppa per step. Per illustrare il calcolo degli indici proposti, si indichi con I_{ijk}^t il valore della *k*-ma componente del (macro) indicatore *j* per la provincia *i* al tempo *t* ($k = 1...m$; $j = 1...5$; $i = 1...n$). Si indichi con I_{rjk}^t il valore base o di riferimento posto uguale alla media nazionale. L'operazione di standardizzazione consente all'indicatore elementare di essere trasformato in numero indice: valori superiori a 100 evidenziano province con un livello dell'indicatore *j* superiore alla media nazionale, mentre valori minori di 100 indicano province con valori inferiori alla media nazionale. Poiché bisogna tener conto del segno e del significato attribuito a ciascun indicatore, la semplice standardizzazione non consente di applicare direttamente l'indice di Jevons: alcuni valori degli indici sopra la media nazionale evidenziano una situazione di competitività; diversamente altri indici con valori superiori a quella della media nazionale segnalano fenomeni di vulnerabilità del territorio. Di conseguenza, si è

effettuata una seconda operazione di standardizzazione sulla provincia *i*-ma che è risultata più vulnerabile. L'indice di vulnerabilità per la provincia *i*-ma relativo al macro (indicatore) *j* può essere definito nel seguente modo:

$$J_{ij}^t = \prod_{k=1}^m (J_{ijk}^t)^{\frac{1}{m}} \quad (1)$$

dove

$$I_{ijk}^t = \begin{cases} \frac{I_{p_i jk}^t}{I_{p_r jk}^t} 100, & \text{se } p_i \text{ è la provincia più vulnerabile rispetto alle altre } r \text{ provincie} \\ \frac{I_{p_r jk}^t}{I_{p_i jk}^t} 100, & \text{se } p_i \text{ è la provincia più vulnerabile rispetto alle altre } r \text{ provincie} \end{cases}$$

L'indice di sintesi di vulnerabilità provinciale (J_i^t) sarà dato dalla seguente formula:

$$J_i^t = \prod_{j=1}^5 (J_{ij}^t)^{\frac{1}{5}} \quad (2)$$

L'indice di sintesi, al pari dei singoli indicatori, è definito per valori non negativi e varia tra 0 (escluso) e 100 (massimo valore che una provincia può assumere in presenza del fenomeno osservato). Valori prossimi allo zero indicano una *quasi*-assenza del fenomeno oggetto di studio. La scelta di utilizzare l'indice di Jevons semplice deriva dal fatto che gli indicatori selezionati sono caratteristici e rappresentativi per tutte le provincie italiane (ogni provincia è vulnerabile dal punto di vista sociale, economico e ambientale; il valore di ciascun indicatore elementare è diverso da zero per ogni provincia). Nel calcolo dell'indice di sintesi, gli indicatori semplici sono stati considerati egualmente importanti ovvero non è stato introdotto alcun sistema di ponderazione. Ciò consente a tutte le provincie di essere comparabili tra loro, dato che il *tableau* degli indicatori elementari, così costruito, è bilanciato. E' possibile definire tali indicatori come i *basic heading mafiosi*: essi rappresentano il più basso livello di aggregazione dei dati per i quali l'informazione è disponibile, attendibile e puntuale, specialmente per le piccole provincie. I valori ottenuti con l'indice di Jevons standard sono uguali a quelli che è possibile ricavare utilizzando l'indice Jevons-GEKS, applicato nel caso in cui gli indicatori elementari non siano caratteristici di tutte le provincie. In tal caso, il calcolo dell'indice di sintesi è limitato al sottoinsieme di provincie i cui indicatori sono caratteristici: nel computo, infatti, non sono considerate quelle provincie in cui non è possibile stimare il valore dell'indicatore elementare, a causa, ad

esempio, della mancanza di dati o perché questi non sono attendibili. Anche in questo caso, non si introduce alcun sistema di pesi per gli indicatori che sono “trattati” in modo equo. In termini analitici, indicando con pc il cluster delle provincie caratteristiche e con M_{pc} l’insieme degli indicatori caratteristici per le provincie osservate, l’indice di sintesi di vulnerabilità per la provincia i -esima relativo al macro (indicatore) j è espresso dalla seguente formula:

$$J_{i_{pc}j_{pc}}^t = \prod_{k_{pc}=1}^{M_{pc}} (J_{ijk}^t)^{\frac{1}{M_{pc}}} \quad (3)$$

Tuttavia, il solo indice di Jevons non garantisce la transitività: per effettuare i confronti multi-spaziali tra le provincie si deve ricorrere alla procedura GEKS (Rao, 2009, 2013).

I risultati così ottenuti vengono, infine, riportati in una matrice di dati, *unità x variabili*, denominata matrice di vulnerabilità. Nelle righe sono collocate le unità di osservazione, ovvero le provincie; nelle colonne si posizionano le variabili delle unità statistiche, vale a dire i diversi indicatori misurati per ogni provincia.

3. Autocorrelazione spaziale

In generale gli indici statistici classici non forniscono una corretta misura dei processi di concentrazione dei fenomeni economici. Tali indici non considerano la vicinanza o l’aggregazione spaziale delle aree ad elevata concentrazione, fornendo valori elevati sia se le aree osservate sono tutte vicine tra loro sia se risultano molto distanti le une dalle altre. L’analisi della concentrazione territoriale di un fenomeno economico può essere arricchita mediante l’introduzione di un indice di autocorrelazione spaziale. L’obiettivo è accertare se sia presente e quanto possa essere intenso il fenomeno dell’autocorrelazione spaziale, ossia dell’influenza della contiguità territoriale sul livello di vulnerabilità delle diverse partizioni del territorio. Si tratta, quindi, di valutare se e come il j -esimo indicatore di vulnerabilità osservato in due aree contigue possa assumere valori simili o dissimili con una certa puntualità. Gli indici di autocorrelazione spaziale possono dividersi in due categorie: alla prima appartengono gli indici globali, come l’indice di Moran, l’indice di Geary o la statistica generale di Getis e Ord, mentre nel secondo gruppo troviamo indicatori locali di interdipendenza spaziale.

Test globale di autocorrelazione - L'indice di Moran, sebbene sottoposto a critiche, è il più utilizzato, tradizionalmente, nelle analisi di correlazione spaziale ed è dato dalla seguente equazione:

$$I = \frac{N}{\sum_{i=1}^N \sum_{j=1}^N w_{ij}(d)} \frac{\sum_{i=1}^N \sum_{j=1}^N w_{ij}(d)(x_i - \bar{x})(x_j - \bar{x})}{\sum_{i=1}^N (x_i - \bar{x})^2} \quad (4)$$

dove x_i è la variabile che descrive il fenomeno oggetto di studio nell'area territoriale i , \bar{x} rappresenta il valore medio campionario, w_{ij} è il singolo elemento della matrice (simmetrica) di contiguità spaziale che assume valore 1 quando l'area j cade in un'area ad una definita distanza d dall'area analizzata i , mentre tutti gli altri elementi sono nulli, inclusi quelli che legano ogni area a se stessa. Precisamente N è la dimensione del campione. Il primo termine rappresenta il rapporto tra il numero di aree geografiche e il numero totale di legami tra aree geografiche stesse; il numeratore del secondo termine indica il prodotto incrociato delle deviazioni dell' i -esima e della j -esima osservazione mentre il denominatore è la devianza della variabile X . Ponendo $z_i = x_i - \bar{x}$ e $W(d) = \sum_i \sum_j w_{ij}(d)$, l'indice I di Moran può essere riscritto in forma compatta:

$$I = \frac{N \sum_i \sum_j w_{ij}(d)(z_{ij})}{W(d) \sum_i z_i^2} \quad (5)$$

Entrando nel dettaglio, w_{ij} denota l'elemento della matrice dei pesi binaria \mathbf{W} , standardizzata sulle righe, e denota l'intensità dell'interazione tra l'area (provincia) i e la vicina area (provincia) j . La matrice spaziale dei pesi \mathbf{W} rappresenta il modo più comune per formalizzare una struttura di prossimità spaziale tra dati areali. Come evidenziato in Espa et al. (2013) non esiste una *natural specification* della matrice \mathbf{W} , per cui una nozione topologica di prossimità spaziale deve essere introdotta, arbitrariamente, dai ricercatori. D'altra parte, la statistica I di Moran è abbastanza sensibile a differenti specificazioni della matrice di contiguità spaziale, soprattutto in termini di risultati robusti. Nel caso di dati areali irregolari, come ad esempio i dati amministrativi, la nozione di vicino può basarsi sul concetto di distanza tra i centroidi delle provincie, usata spesso nelle applicazioni di statistica spaziale o di econometria spaziale. In particolare, si fa riferimento alla definizione di *critical cut-off* tra vicini, ovvero, come già introdotto, due provincie si definiscono vicine se la loro distanza è uguale o minore di una certa distanza fissata (che identifica il *critical cut-off*). Nel presente lavoro, è stata scelta come *cut-off* la distanza minima tra i centroidi delle provincie in modo tale da assicurare che ogni provincia abbia almeno un vicino. La

scelta di un'adeguata matrice W , tra le varie alternative possibili, e di un opportuno *critical cut-off* consentono di controllare per la robustezza dei risultati⁴.

Per valori di N sufficientemente grandi, la I di Moran si distribuisce come una normale standardizzata $Z(I)$ (Cliff et al., 1973). Valori non significativi di $Z(I)$ inducono a rigettare l'ipotesi nulla di assenza di autocorrelazione, mentre valori positivi (negativi) indicano la presenza di autocorrelazione spaziale positiva (negativa). Nello specifico, l'indice può assumere segno negativo rilevando un'autocorrelazione negativa, nel caso in cui le coppie di scarti dalla media delle unità territoriali (aree) contigue, aventi segno diverso, risultino prevalenti rispetto a quelle aventi lo stesso segno e segno positivo nel caso in cui risultino prevalenti le coppie con lo stesso segno. L'indice di Moran varia da -1 a 1, dove 1 indica "perfetta correlazione positiva" e -1 indica "perfetta correlazione negativa".

Dal punto di vista formale indicando con $E(I) = -\frac{1}{n-1}$ il valore atteso, è possibile interpretare l'insieme dei valori assunti dall'indice di autocorrelazione globale, secondo il prospetto di seguito riportato:

1. $I > E(I)$: clusterizzazione ovvero omogeneità (tra aree);
2. $I = E(I)$: indipendenza ovvero casualità;
3. $I < E(I)$: dissimilarità ovvero eterogeneità.

LISA - Nelle applicazioni è utile associare a un indice globale di autocorrelazione anche un indice di autocorrelazione locale, in grado cioè di misurare l'interdipendenza per ognuna delle aree territoriali in esame. È ragionevole supporre che l'intensità dell'autocorrelazione spaziale non sia necessariamente uniforme su tutta l'area d'interesse. Di conseguenza, è plausibile che l'intensità di tale autocorrelazione sia più alta in alcune sub-partizioni e bassa in altre. Gli indici di autocorrelazione globale permettono di verificare la presenza di *pattern* di diffusione spaziale di variabili economiche per diverse sub-partizioni. Tuttavia, i risultati ottenuti nelle analisi spaziali sono fortemente condizionati dalla scelta dell'unità territoriale di riferimento. Analisi aggregate inerenti estensioni geografiche rilevanti possono nascondere *pattern* di diffusione spaziale al loro interno anche molto differenti. Identificare le divergenze esistenti a livello locale dai valori di riferimento provinciali può aiutare a migliorare la definizione dei modelli interpretativi dei processi spaziali, ovvero offrire indicazioni sul

⁴ Per ulteriori approfondimenti sulla scelta del *critical cut-off* rispetto a differenti matrici di prossimità spaziale e per la robustezza dei risultati si rimanda a Espa et al. (2013).

livello territoriale significativo per lo studio dei cambiamenti nei processi spaziali.

Il **LISA** (*Local Indicator of Spatial Association*) permette in modo efficace di associare a ogni unità territoriale una misura del corrispondente livello di associazione spaziale rispetto a un suo intorno. Anche tale indicatore, che si distribuisce come una normale standardizzata, può essere sottoposto a test per verificare la presenza (assenza) di autocorrelazione. Valori positivi segnalano l'esistenza di clusters territoriali con caratteristiche simili; viceversa, valori significativamente negativi evidenziano clusters territoriali tra loro diversificati. In formule l'indice **LISA**⁵ è dato dalla seguente relazione:

$$I_i = \frac{z_i}{\sum_j \frac{z_j}{N}} \sum_j w_{ij}(d) z_j \quad (6)$$

in cui z_i è il valore corrispondente all'area territoriale i della variabile già normalizzata; J rappresenta l'insieme delle aree geografiche vicine a i ; N è il numero di osservazioni che cadono nell'area di contiguità di i ; w_{ij} è il singolo elemento della matrice (simmetrica) di contiguità spaziale, già introdotto precedentemente. Il valore I_i per una data area i rappresenta l'associazione tra l'area i -esima e i valori di j appartenenti alle aree contigue individuate sulla base delle distanze (tra i centroidi delle provincie). La somma degli I_i al netto di un fattore di proporzionalità, restituisce l'indice globale di Moran. Valori positivi dell'indice I_i per una data partizione i denotano la presenza di una concentrazione a livello locale di valori tutti elevati (o tutti ridotti) dell'indicatore di vulnerabilità. D'altra parte, valori positivi elevati possono essere anche causati dal forte impatto di alcuni punti sull'indice di autocorrelazione spaziale globale di Moran. Valori negativi evidenziano, al contrario, l'esistenza di aree eterogenee rispetto alla variabile oggetto di studio. I valori positivi sono quelli per cui $I_i > E [I_i]$, mentre i valori negativi sono quelli per cui $I_i < E [I_i]$; i valori significativi di I_i per individuare le aree omogenee vengono selezionati, sotto l'ipotesi di un distribuzione normale standardizzata, attraverso un test Z a due code (Mazziotta et al., 2005).

⁵ Come il LISA costituisce una versione locale dell'indice di Moran, così esiste una versione locale dell'indice di Geary e dell'indice di Getis e Ord.

4. Indicatori di vulnerabilità: principali evidenze

Dall'analisi dell'indice di sintesi di vulnerabilità provinciale (Tab. 1)⁶ emerge come progressivamente la criminalità organizzata stia penetrando nel tessuto della società civile e nelle attività economiche legali. Nessun territorio è esente da possibili infiltrazioni di gruppi mafiosi (sia italiani che stranieri): affermare che in un'area (specialmente quelle del Nord) è presente una bassa vulnerabilità e/o attrattività significa che l'introduzione della criminalità organizzata all'interno del tessuto imprenditoriale e sociale è solo parziale e sta avvenendo gradualmente. L'indice di sintesi evidenzia, inoltre, che le aree a più alta vulnerabilità sono le province della Calabria, segno evidente dell'ascesa prepotente della 'ndrangheta negli ultimi decenni. In particolare, si rileva che le province a maggiore intensità criminale sono Crotona e Vibo Valentia, province che si sono staccate amministrativamente da Catanzaro solo nel 1992. Il risultato conseguito non deve sorprendere, dato che le suddivisioni amministrative non riflettono le vere partizioni economiche (e sociali) del territorio oltre che le dinamiche imprenditoriali, trasversali tra le aree confinanti. Catanzaro, ponendosi tra le due province, ne garantisce la contiguità territoriale. In generale, escludendo la Sardegna e la Sicilia, il cui punto di contatto con la Penisola è costituito dall'asse Messina-Reggio Calabria, la criminalità organizzata è fortemente radicata in Calabria e Campania, le cui province presentano i più alti valori di vulnerabilità globale. Tali aree rappresentano i poli per eccellenza della criminalità: da esse si diramano importanti direttrici per la diffusione della stessa, specialmente nella zona del Basso Sannio, fino al Tavoliere pugliese (Fig. 1). L'analisi grafica, ottenuta con la costruzione dei quartili, denota come la criminalità risalga la Penisola sia lungo la dorsale adriatica (Campobasso, Pescara e Teramo) per penetrare economicamente nelle Marche (Ancona) e in Emilia Romagna (Rimini), sia attraversando l'Appennino centro-meridionale per estendere i propri interessi nell'economia legale nel basso Lazio, in Umbria e bassa Toscana (Fig. 1-2). Analizzando le nuove forme di criminalità legate al ciclo del cemento e dei rifiuti e le aree del *made in Italy*, dove le imprese una volta fiorenti oggi sono in crisi, tale mappa denota come l'imprenditorialità criminale (anche di

⁶ Nella tabella 1 dell'Appendice sono riportate due versioni dell'indice di sintesi di vulnerabilità provinciale: la prima denominata VINDEX in cui si riportano i valori ottenuti direttamente dal calcolo dell'indice di Jevons, che variano tra 0 (assenza di vulnerabilità) e 100 (massimo valore di vulnerabilità che una provincia può assumere); la seconda denominata V'INDEX in cui i valori vengono riscritti da 0 (provincia meno vulnerabile) a 100 (provincia più vulnerabile).

origine straniera) oltre che in Umbria, Marche e Toscana sia arrivata in Lombardia ed abbia percorso tutta la dorsale adriatica fino a Trieste, nuovo polo emergente insieme a Gorizia delle infiltrazioni in campo economico da parte dei gruppi criminali. Il processo di penetrazione al Nord procede a velocità alternata e inizialmente riguarda quei settori e quelle attività, localizzate in province strategiche dal punto di vista della dotazione infrastrutturale e turistica, in cui è più facile riciclare il denaro attraverso investimenti immobiliari. Ciò spiega perché nell'alto Tirreno si segnalano come nuovi possibili centri della criminalità organizzata le province di Livorno, La Spezia e Imperia e le aree a esse limitrofe (in particolare le province di Lucca e Grosseto), tutte dotate, oltre che di importanti infrastrutture portuali, di strutture ricettive e di ristorazione, oggetto di acquisizione e investimento nell'ultimo decennio da parte della criminalità organizzata, soprattutto dell'est Europa. Analogo discorso vale per la dorsale adriatica marchigiana e romagnola fino a giungere al polo di Trieste, cruciale snodo ferroviario e marittimo, fulcro di una nuova *mitteleuropea* di matrice criminale e di scambi terra-mare tra i mercati dell'Europa centro-orientale e dell'Asia. Le cartine geografiche relative agli indicatori (Fig. 3-5) rilevano come in alcune importanti province del Centro- Nord (Roma, Torino, Milano, Venezia e le province liguri) la vulnerabilità sia elevata in termini di indicatori spia dell'illegalità economica (o criminalità organizzata) e in termini di vulnerabilità delle imprese (Roma e Venezia). In generale, anche se il fenomeno della vulnerabilità sia a livello globale che di singolo indicatore è maggiormente presente nelle regioni meridionali, nessuna provincia italiana ne è esente, come già introdotto, segno che tale fenomeno comincia pian piano a strutturarsi in tutta la Penisola: anche per quanto riguarda l'indice di vulnerabilità delle famiglie le due principali metropoli Roma e Milano, pur presentando bassi valori, sono circondate da aree che presentano valori medio-alti.

5. Le aree strutturali della vulnerabilità e della criminalità

I test di autocorrelazione – L'autocorrelazione spaziale può essere dovuta principalmente: 1) a errori di misurazione per osservazioni riferite a unità geografiche contigue; 2) a una vera e propria interazione spaziale. Gli errori di misurazione possono sorgere ogni qualvolta si utilizzano dati per i quali non c'è perfetta corrispondenza tra unità territoriale di analisi ed estensione del fenomeno oggetto di studio. La seconda causa di autocorrelazione, la presenza di reali fenomeni di interazione spaziale, discende dalla particolare localizzazione degli agenti economici. A causa della dipendenza che si crea

tra fenomeni socio-economici in differenti aree geografiche, ciò che si osserva in un punto è determinato, almeno in parte, da ciò che accade in un altro punto del sistema. In entrambe le due tipologie di dipendenza spaziale, la presenza di autocorrelazione spaziale deve essere affrontata per mezzo di appropriate metodologie. Eseguendo un test di autocorrelazione globale sia per l'indice di sintesi di vulnerabilità provinciale che per i singoli indicatori (Tabella 2), si evince che l'autocorrelazione spaziale è presente denotando valori elevati, essendo i corrispondenti valori di $Z(I)$ sempre positivi e significativamente diversi da zero (sia nel caso di distribuzione normale che *randomizzata*).

Il Moran scatterplot - In riferimento alla statistica I di Moran è possibile associare il relativo grafico, denominato *Moran scatterplot*, che fornisce tra l'altro utili informazioni complementari e integrative, con particolare riferimento agli *outliers*. Sull'asse delle ascisse si riporta il valore dell'indice J di vulnerabilità normalizzato e su quello delle ordinate il ritardo spaziale di detto indice anch'esso normalizzato (WJ).

Seguendo l'impostazione di diversi autori (Atzeni et al., 2004; Wong et al., 2005, Bivand et al., 2013) e disegnando il *Moran scatterplot* per l'indice di sintesi di vulnerabilità provinciale (Grafico 1), la I di Moran è rappresentata dal coefficiente angolare della relazione lineare tra le due variabili riportate sugli assi del suddetto grafico. Se i punti sono dispersi fra i quattro quadranti questo indicherà assenza di correlazione, dato che il coefficiente angolare è pari zero. Se invece esiste una qualche forma di relazione, il *Moran scatterplot* può essere utilizzato per distinguere diverse tipologie di correlazione spaziale. Nello specifico, se i punti sono situati nei quadranti Nord-Est e Sud-Ovest si avrà correlazione positiva, distinguendo tra i casi (quadrante Nord-Est) dove la relazione si stabilisce per valori alti sia dell'indice J di vulnerabilità provinciale (la provincia i) che di WJ (le province limitrofe a i) e quelli dove (quadrante Sud-Ovest) i valori sono entrambi bassi. Nel primo caso si parla di relazione Alto-Alto, nel secondo di Basso-Basso. Alternativamente, se la concentrazione dei punti è maggiore negli altri due quadranti la correlazione è negativa: si parlerà di relazione Basso-Alto per il quadrante Nord-Ovest e di relazione Alto-Basso per il quadrante Sud-Est.

I risultati del *Moran scatterplot* per l'indice di sintesi di vulnerabilità provinciale (I di Moran = 0,66) possono essere riportati su una mappa (Figura 7) in modo da distinguere geograficamente le aree con le diverse tipologie di correlazione (Alto-Alto, Basso-Basso, Alto-Basso, Basso-Alto).

E' possibile, di conseguenza, verificare se le regioni accomunate da un certo tipo di correlazione sono fra loro contigue raggruppandosi in un *cluster*. Il *Moran scatterplot* consente, inoltre, di evidenziare, come sopra descritto, i possibili casi limite (*outliers*): la rappresentazione su una cartina geografica aiuta a individuare esattamente le province anomale e a valutare se la loro posizione geografica ha una relazione con essa. Ciò è tanto più verosimile se gli *outliers* sono tra loro vicini oppure identificano aree di confine, dove le esternalità si estendono a un livello mesoeconomico, oltre i confini provinciali, e non vengono colte all'interno di un'analisi statistica tradizionale con dati amministrativi (regionali, provinciali, comunali).

La figura 7 mostra un'Italia divisa in due: il Nord e alcune province del Centro (Emilia Romagna, la costa adriatica delle Marche e la dorsale tirrenica dell'alta Toscana) caratterizzate da una relazione del tipo Basso-Basso); il Sud e alcune aree del Centro fino all'Appennino umbro in cui prevalgono relazioni del tipo Alto-Alto. Le aree analizzate contribuiscono a definire la correlazione spaziale complessiva. Da queste si distinguono le province caratterizzate da una correlazione negativa che sono situate prevalentemente nelle zone di confine tra le aree in precedenza definite. Nel dettaglio, tali province sono localizzate sulla costa sia tirrenica che adriatica e nell'Appennino toscano. Le mappe relative ai singoli indicatori (Fig. 8-12) mostrano anch'esse una prevalenza di relazioni Basso-Basso al Nord e Alto-Alto al Sud, soprattutto per gli indici di criminalità organizzata del territorio e di vulnerabilità delle famiglie i cui *Moran scatterplot* (Fig. 9 e 12) sono simili a quello dell'indice di sintesi globale, ma denotano, al tempo stesso, la presenza di diffuse aree di eterogeneità spaziale, ovvero di correlazione negativa. In realtà, piuttosto che ipotizzare la rimozione di alcune osservazioni anomale ovvero delle province responsabili di un'apparente discontinuità del processo di diffusione spaziale, si assiste, al contrario, a una graduale propagazione dei fenomeni di vulnerabilità associati ai singoli indicatori J della matrice dei dati.

LISA - Spesso è utile associare a un indicatore di autocorrelazione globale e alla sua rappresentazione grafica anche un indicatore di autocorrelazione locale, in grado cioè di misurare l'interdipendenza per ognuna delle aree analizzate. Il LISA consente in modo efficace di associare ad ogni unità territoriale una misura del livello di associazione spaziale rispetto al suo intorno. Nuovamente l'ipotesi nulla è l'assenza di autocorrelazione spaziale per cui se il test, che si distribuisce come una normale standardizzata, presenta valori significativamente positivi si otterrà un *cluster* di regioni con caratteristiche simili. Viceversa, valori significativamente negativi

indicheranno un *cluster* di regioni diversificate. Per ogni unità territoriale è possibile indicare il tipo di correlazione (negativa o positiva) e il suo livello di significatività che serve confermare o meno l'attendibilità dei valori assunti dal LISA (si utilizza un livello di confidenza al 95%). Ciò consente di verificare se i processi di clusterizzazione sono casuali o strutturali. Inoltre, tramite un test di autocorrelazione locale possono essere escluse quelle correlazioni che non raggiungono il livello di significatività presenti, invece, nella mappa del *Moran scatterplot*. L'analisi grafica del LISA insieme ai valori degli *z-scores* ha evidenziato, per l'indice di sintesi globale, tre agglomerazioni (Fig. 13-14): la prima si colloca al Nord, partendo dalla provincia di Varese, e si estende lungo la Pianura Padana fino al Nord-Est nella Pianura Veneta mostrando le provincie con i valori più bassi dell'indice di vulnerabilità complessiva (relazione di tipo Basso-Basso). La seconda clusterizzazione, dove i valori della vulnerabilità sono tra i più elevati, è localizzata al Sud e precisamente in Calabria, Basilicata e nelle provincie dell'Appennino meridionale (relazione di tipo Alto-Alto). Il terzo raggruppamento riguarda il Nord-Est della Sardegna, altra area tra le più vulnerabili (relazione di tipo Alto-Alto). Il test di autocorrelazione locale per i singoli indicatori di vulnerabilità con le relative rappresentazioni grafiche (Fig. 15-24) evidenzia come le provincie della Calabria siano in generale le più vulnerabili per ogni caratteristica esaminata, in particolare per l'indice di criminalità organizzata che, si ricorda, comprende anche i reati del ciclo dei rifiuti e del cemento. Nel *Moran scatterplot* tali provincie mostrano una relazione del tipo Alto-Alto ad esclusione della provincia di Catanzaro che evidenzia una relazione, al contrario, del tipo Basso-Alto (e con *z-scores* non significativi), a conferma di quanto già sottolineato a proposito della differenza tra partizioni funzionali del territorio e dati amministrativi. La seconda aggregazione rilevante per questa tipologia di indice di vulnerabilità è rappresentata dall'area campana con esclusione della provincia di Caserta che nel *Moran scatterplot* presenta al pari delle altre provincie una relazione del tipo Alto-Alto. Ciò non implica che Caserta non sia tra le provincie più vulnerabili (si veda il test di autocorrelazione globale) per la presenza di criminalità organizzata associata anche a nuove forme di imprenditorialità mafiose in campo ambientale ed edilizio ma il LISA segnala solo specifiche aree che denotano i valori maggiori (o minori) dell'indice esaminato. Anche per l'indice di vulnerabilità delle imprese, la Calabria evidenzia le provincie con i valori più elevati (relazioni di tipo Alto-Alto). Infine, per l'indice di vulnerabilità

delle famiglie si osserva la formazione di tre diversi cluster al Sud, dove si osservano relazioni di tipo Alto-Alto: il primo relativo alla Sicilia con esclusione delle provincie di Trapani, Agrigento e Messina; il secondo formato dalle provincie di Cosenza, Crotone e Catanzaro; il terzo comprendente le provincie di Napoli, Avellino, Benevento e Campobasso. A questi raggruppamenti si aggiunge la provincia di Taranto. Parallelamente, si rilevano due distinte agglomerazioni al Centro-Nord, caratterizzate da relazioni di tipo Basso-Basso: la prima, localizzata in Toscana, è formata dalle provincie di Siena, Firenze, Pisa; la seconda è situata nel Triveneto e precisamente nelle provincie di Bolzano, Trento, Belluno e nell'area di Padova e Venezia. Tuttavia, le provincie meridionali non sono le più vulnerabili per qualsiasi indicatore analizzato. Osservando l'indice spia di infiltrazione dell'illegalità economica (o della criminalità organizzata) si evince che le provincie di Milano e Genova presentano valori degli *z-scores* significativi ma negativi. Ciò non deve sorprendere in quanto le due aree urbane del Nord nella graduatoria dell'indice semplice di Jevons denotano valori più elevati rispetto a quelli delle provincie limitrofe e nel *Moran scatterplot* sono caratterizzate da relazioni del tipo Alto-Basso. Le altre due provincie che segnalano, al contrario, valori positivi e significativi degli *z-scores* sono Imperia e Latina contraddistinte da relazioni del tipo Alto-Alto.

6. I modelli spaziali autoregressivi

Uno dei problemi che spesso si incontra nell'analisi dei processi di interazione spaziale è dovuto alla mancanza di uniformità nello spazio circa la localizzazione degli agenti economici: la mancanza di una stabilità strutturale nello spazio e la poca omogeneità delle unità territoriali pone questioni metodologiche di non poca rilevanza. In un'analisi multivariata, nel contestualizzare le peculiari caratteristiche di ogni unità territoriale, si deve affrontare il problema dell'eterogeneità. Dal punto di vista della modellizzazione questo implica che le forme funzionali e i parametri variano in ragione della loro localizzazione geografica e non sono omogenei in tutto il *data set* (Cliff et al., 1969). Poiché questa eterogeneità può essere messa in relazione direttamente con l'allocazione spaziale, viene comunemente designata con il termine di *eterogeneità spaziale*. Nella misura in cui questi aspetti dell'eterogeneità si riflettono in errori di misurazione (variabili mancanti, misspecificazione funzionale etc.) possono avere come conseguenza l'eteroschedasticità (Atzeni et al., 2004, Haining, 1990). Per tenere conto, quindi, della dipendenza spaziale delle aree vicine

tra loro si ricorre all'utilizzo di modelli spaziali autoregressivi. La dipendenza spaziale può essere strutturata in differenti modi utilizzando appropriati modelli statistici (Cressie, 1993). In letteratura, per semplicità, si assume che le osservazioni siano indipendenti e identicamente distribuite ma ciò spesso non trova riscontro nelle analisi empiriche quando si lavora con dati spaziali. Le osservazioni non risultano indipendenti poiché esistono correlazioni tra aree vicine. In questa sezione si provvederà a fornire una breve *review* dei principali metodi utilizzati nell'analisi dei processi spaziali quando si utilizzano dati areali, cercando di risolvere anche i problemi connessi all'eteroschedasticità (Waller et al., 2004, pp. 362-380; Schanberberger et al., 2005, pp. 335-338).

I modelli SAR – Il modello autoregressivo simultaneo (SAR) è un modello di regressione che tiene conto dei valori delle aree vicine tra loro. Ciò significa che i termini di errore sono costruiti in modo tale da dipendere gli uni dagli altri come di seguito riportato:

$$e_i = \sum_{i=1}^m b_{ij} e_i + \varepsilon_i \quad (7)$$

in cui e rappresenta il termine di errore in un modello multivariato normale, b_{ii} è posto uguale a zero e il termine ε_i denota gli errori residui che si assumono indipendenti e distribuiti secondo una distribuzione normale di media zero e matrice diagonale di varianze e covarianze $\Sigma_\varepsilon = \text{diag}[\sigma_1^2 \dots \dots \sigma_n^2]$. I termini b_{ij} rappresentano i parametri di dipendenza spaziale dal momento che misurano il contributo delle altre osservazioni (vale a dire $e_j, j \neq i$) alla variazione di e_i . Se i termini b_{ij} fossero uguali a zero non ci sarebbe autoregressione e il modello si ridurrebbe a un tradizionale modello di regressione lineare con errori indipendenti. Esprimendo il termine di errore tramite la seguente espressione $e = B(Y - X^T \beta) + \varepsilon$, il modello di regressione diviene:

$$Y = X^T \beta + B(Y - X^T \beta) + \varepsilon \quad (8)$$

In termini matriciali, è possibile riscrivere la (8) come di seguito riportato:

$$(I - B)(Y - X^T \beta) = \varepsilon \quad (9)$$

dove B è la matrice che contiene i parametri di dipendenza spaziale b_{ij} e I è la matrice identità. La matrice $I - B$ deve essere non singolare e quindi invertibile affinché il modello SAR sia definito. Il modello descritto nell'equazione (9) è chiamato *Simultaneous Autoregressive Models*, dove l'aggettivo *simultaneous* descrive le simultanee applicazioni a ogni *data*

location e distingue questa classe di modelli spaziali dai *Conditional Autoregressive Model (CAR)*.

Nel modello (9), Y è distribuita secondo una normale multivariata di media $E[Y] = X^T \beta$ (9a) e matrice di varianze e covarianze: $Var(Y) = (I - B)^{-1} \Sigma_\varepsilon (I - B^T)^{-1}$ (9b). Se la matrice Σ_ε dipende da un singolo parametro σ^2 (ciò che avviene spesso nelle applicazioni), ne consegue che $\Sigma_\varepsilon = \sigma^2 I$ e $Var[Y] = \sigma^2 (I - B)^{-1} (I - B^T)^{-1}$. Un'utile riparametrizzazione del modello SAR può essere ottenuta ponendo $B = \lambda W$, dove λ identifica il parametro di autocorrelazione spaziale e W è la matrice di contiguità spaziale, spesso assunta simmetrica, introdotta nel paragrafo 3. Con queste notazioni il modello SAR può essere così riscritto:

$$Y = X^T \beta + \lambda W(Y - X^T \beta) + \varepsilon \quad (10)$$

La varianza di Y diventa $Var[Y] = \sigma^2 (I - \lambda W)^{-1} (I - \lambda W^T)^{-1}$.

Stima dei parametri dei modelli SAR - Il modello SAR riparametrizzato può essere stimato efficientemente mediante la funzione massima verosimiglianza: gli stimatori di massima verosimiglianza di β e σ^2 sono equivalenti agli stimatori dei minimi quadrati generalizzati. Se λ non è noto lo stimatore dei minimi quadrati è inconsistente (Waller et al., 2004, p. 366; Haining, 1990, p. 130). Diversamente se $\Sigma_\varepsilon = \sigma^2 V_\varepsilon$, la matrice di varianze e covarianze del modello SAR diventa $\Sigma_{SAR} = \sigma^2 (I - B)^{-1} V_\varepsilon (I - B^T)^{-1} = \sigma^2 V_{SAR}(\theta)$ dove θ è il vettore che contiene i parametri di dipendenza spaziale b_{ij} e i parametri del vettore V_ε . Se $V_{SAR}(\theta)$ è nota i risultati delle stime di massima verosimiglianza e dell'analisi di inferenza per modelli lineari generali con errori autocorrelati si applicano direttamente per le stime e l'inferenza del modello SAR (Whittle, 1954, Haining, 1990, Cliff et al., 1973, Cliff et al., 1981). Alternativamente, se $V_{SAR}(\theta)$ è non nota per le stime dei parametri si utilizza la funzione di log-verosimiglianza (concentrata). In definitiva, si conduce un test per verificare la presenza di autocorrelazione spaziale residua presente nei residui calcolati con i minimi quadrati ordinari. Nello specifico, si considerano due modelli, uno tradizionale, cioè il modello di regressione lineare con errori indipendenti, e l'altro il modello SAR, con la matrice Σ_ε dipendente da un solo parametro, secondo il seguente schema:

$$Y = \begin{cases} X\beta + e, & e = \lambda W e + \varepsilon, & Var[Y] = \sigma^2 I & (11) \\ X\beta + (I - \lambda W)^{-1} \varepsilon, & Var[Y] = \sigma^2 (I - \lambda W)^{-1} (I - \lambda W^T)^{-1} & (12) \end{cases}$$

Se nell'equazione (11) si pone $\lambda = 0$ si ottiene il modello di regressione lineare con errori indipendenti. Il test di autocorrelazione spaziale da eseguire è il seguente:

$$\begin{cases} H_0: \lambda = 0 \\ H_1: \lambda \neq 0 \end{cases}$$

Per una breve rassegna sui metodi per costruire i test di autocorrelazione spaziale (calcolo dell'indice I di Moran in presenza di residui nei minimi quadrati ordinari, test di Wald e test sul rapporto di verosimiglianza) si rinvia a Waller et al. (2004, pp. 367-370). I vari test sono tutti asintotici, vale a dire i risultati sono validi per un numero elevato N di partizioni contenute nell'area oggetto di studio. La dimensione di N dipende da un insieme di fattori tra cui la struttura della matrice di prossimità spaziale W .

I modelli CAR - In alcune applicazioni è più intuitivo e utile specificare dei modelli di distribuzione di probabilità condizionata per ogni osservazione $Y(s_i)$, dati i valori di tutte le altre osservazioni. In altri termini, si introduce il seguente modello $f[Y(s_i)|Y_{-i}]$, in cui Y_{-i} denota il vettore di tutte le osservazioni eccetto $Y(s_i)$. Come nel modello SAR, è possibile semplificare l'analisi assumendo che $Y(s_i)$ dipenda solamente dai vicini, cioè $Y(s_i)$ dipende da $Y(s_j)$ solo se la localizzazione di s_j è nell'insieme dei vicini N_i , di s_i . Si assume, inoltre, che ogni distribuzione condizionata sia Gaussiana; il modello CAR (*Conditional Autoregressive Model*) è, quindi, identificato dai seguenti parametri:

$$E[Y(s_i)|Y_{-i}] = x(s_i)' \beta + \sum_{j=1}^N c_{ij} [Y(s_j) - x(s_j)' \beta] \quad (13)$$

$$Var[Y(s_i)|Y_{-i}] = \sigma^2, \quad i = 1 \dots N \quad (14)$$

dove i termini c_{ij} rappresentano i parametri di dipendenza spaziale e sono diversi da zero solo se $s_j \in N_i$. Per convenzione $c_{ij} = 0$, dal momento che non si regredisce alcuna osservazione su se stessa. Un modo alternativo per rappresentare il modello CAR, utilizzando i termini di errore del modello multivariato normale e indicando con $e_{j \sim i}$ l'insieme dei vicini di i , è dato dalla seguente espressione:

$$e_i | e_{j \sim i} \sim N \left(\sum_{j \sim i} \frac{c_{ij} e_j}{\sum_{j \sim i} c_{ij}}, \frac{\sigma_{e_i}^2}{\sum_{j \sim i} c_{ij}} \right) \quad (15)$$

Formalmente le condizioni necessarie, in termini di distribuzioni condizionate, per definire una distribuzione congiunta con autocorrelazione spaziale (positiva) sono contenute nel teorema di *Hammersley-Clifford* (Besag, 1974, Cressie, 1993). Per la costruzione di modelli Gaussiani, è possibile dimostrare, sotto condizioni non troppo restrittive (Besag, 1974; Cliff et al., 1981, p. 180), che la famiglia di distribuzioni Gaussianhe con

media condizionata e varianza condizionata rappresentate dalle equazioni (13) e (14) genera una distribuzione congiunta multivariata Gaussiana con media $X\beta$ e varianza fornita dalla seguente equazione: $Var(Y) = (I - C)^{-1}\Sigma_C$, dove $\Sigma_C = diag[\sigma_1^2 \dots \dots \sigma_n^2]$. Per assicurare che la matrice di varianze e covarianze sia simmetrica si impone la seguente restrizione: $\sigma_j^2 c_{ij} = \sigma_i^2 c_{ji}$. Ponendo $\Sigma_C = \sigma^2 I$ e sviluppando un modello SAR tale che $\Sigma_\varepsilon = \sigma^2 I$, dopo aver confrontato le varianze espresse dalle equazioni (9b) e (14), è possibile esprimere ogni modello SAR, con matrice di dipendenza spaziale B , come un modello CAR, con matrice di dipendenza spaziale $C = B + B^T - BB^T$. Parimenti, ogni modello CAR può essere espresso come un modello SAR, sebbene le relazioni tra le matrici B e C *are somewhat contrived* (Haining, 1990, p. 89) e le strutture di vicinato dei due modelli possono non essere le stesse (Cressie, 1993, p. 409).

Riparametrizzando la matrice diagonale Σ_C come $\Sigma_C = \sigma^2 V_C$, la matrice di varianze e covarianza di un modello CAR può essere riscritta nel seguente modo:

$$\Sigma_{CAR} = \sigma^2 (I - C)^{-1} V_C = \sigma^2 V_{CAR}(\theta) \quad (16)$$

dove θ è un vettore che contiene i parametri di dipendenza spaziale c_{ij} e i parametri di V_C .

Stima dei parametri dei modelli CAR – Come già introdotto per i modelli SAR, per la stima dei parametri e le procedure di inferenza, anche per i modelli CAR la stima parametrica e l'applicazione dei test di autocorrelazione spaziale avviene attraverso metodologie simili, a seconda che $V_{CAR}(\theta)$ sia nota o meno (Waller et al., 2004, pp. 372-373; Cliff et al., 1981, p. 242). È possibile, come nei modelli SAR, riparametrizzare il modello CAR dipendente da un solo parametro ponendo $C = \lambda W$ e $V_C = \sigma^2 I$. I metodi per il test di autocorrelazione spaziale per il modello CAR a un solo parametro sono gli stessi di quelli discussi per il modello SAR.

Differenze tra i due modelli - La principale differenza tra i modelli CAR e SAR basati su osservazioni Gaussiane consiste nella differente definizione della matrice di varianze e covarianze. Nel caso di un solo parametro, per il modello CAR, $Var[Y] = \sigma^2 (I - \lambda W)^{-1}$ mentre per il modello SAR si ha: $Var[Y] = \sigma^2 (I - \lambda W)^{-1} (I - \lambda W^T)^{-1}$. Poiché gli errori non sono indipendenti dalle osservazioni, gli stimatori dei minimi quadrati non sono necessariamente consistenti nel modello SAR, come già evidenziato. Al contrario, gli stimatori dei minimi quadrati dei parametri del modello CAR sono consistenti. Nel modello CAR, i minimi quadrati generalizzati pesati

possono quindi essere utilizzati per stimare tutti i parametri del modello. Inoltre la struttura condizionata del modello CAR fornisce predittori più efficienti in termini di MSE.

Conclusioni – Analogamente a quanto introdotto per i modelli SAR, l'obiettivo primario dei modelli CAR è di definire un meccanismo che tenga conto della correlazione spaziale residua. La correlazione residua può comportare distorsione nella stima dei parametri e nei loro errori standard oltre che un'*inferenza non corretta* nel valutare l'impatto delle covariate sul risultato ottenuto. I modelli spaziali autoregressivi sono stati sviluppati per tener conto dei dati geografici aggregati a differenza dei modelli tipici della geostatistica e della *point pattern analysis* che studiano i dati spaziali continui, ovvero misurano le osservazioni punto per punto all'interno dell'area oggetto di analisi (Ripley, 1981; Diggle, 2003). D'altra parte pur essendo dei modelli di regressione con autocorrelazione spaziale dei dati, la previsione di nuove osservazioni non è il focus dei modelli SAR e CAR: l'interpretazione dei valori previsivi localizzati tra le unità di aree adiacenti ancora non è chiara e il dibattito resta aperto. In realtà, il problema ancor prima che di natura metodologica riguarda l'analisi e la selezione dei dati, come già dibattuto nel presente lavoro. Non necessariamente un dato di natura amministrativa ha implicazioni di carattere territoriale vale a dire spaziale nel caso in cui si costruiscono partizioni delle aree a livello mesoeconomico. In altri termini, un dato amministrativo e/o politico, esogeno, ha implicazioni trasversali sulle aree oggetto di studio e non ne evidenzia le differenze strutturali; un dato territoriale endogeno, al contrario, denota la particolare struttura sociale ed economica dell'area osservata. Ciò implica che nell'analisi dei dati spaziali bisogna osservare prima di tutto la natura dei dati stessi, poiché non tutte le informazioni possono essere trasformate in dati territoriali e, successivamente, una volta verificate le caratteristiche, costruire degli indicatori (variabili) con appropriate funzioni statistiche che riflettano la natura (concentrazione) spaziale del fenomeno oggetto di studio. Infine, si utilizzano tali predittori in un modello di regressione con autocorrelazione spaziale degli errori, che tenga conto eventualmente anche dell'eteroschedasticità nel caso in cui ci sia una *mispesificazione* dei dati.

7. Un esercizio di previsione: la contraffazione dei beni

Nella presente sezione si provvederà a simulare un esercizio di previsione: utilizzando come predittori alcune variabili introdotte per la costruzione

della matrice di vulnerabilità a livello provinciale, si verificherà se il fenomeno della contraffazione dei beni del *made in Italy* (sono esclusi i prodotti alimentari e farmaceutici) è funzione di variabili caratteristiche della criminalità organizzata ovvero di variabili legate alla vulnerabilità socio-economica del territorio. Come regressori sono state considerati i seguenti predittori (della matrice di vulnerabilità): indice della criminalità organizzata, indice dei reati di contraffazione, indice di sofferenza delle imprese, indice di credito alle famiglie. La variabile dipendente è data dal rapporto tra offerta dei beni contraffatti del *made in Italy* rilevata al 2012 (Iperico, 2013)⁷ ponderata con i consumi di tali beni al 2011 (Tagliacarne, 2012). Se il valore di tale rapporto è superiore a 1 ciò implica che l'offerta di tali beni non è destinata solo al mercato locale ma, al contrario, la provincia in questione è mercato potenziale dei beni contraffatti o luogo di transito della merce ovvero centro di specializzazione della produzione dei beni destinati sia all'Italia che all'estero. Alternativamente, se il valore di tale rapporto è inferiore o uguale a 1 ciò significa che l'offerta di beni contraffatti è destinata solo al mercato locale. Tra i predittori, oltre all'indice dei reati di contraffazione, è stato inserito l'indice dei reati tipici della criminalità organizzata (ponderato con le pene edittali) per verificare se esiste una correlazione tra le tradizionali aree a carattere mafioso e le provincie in cui avviene la distribuzione e/o produzione di beni contraffatti. Le variabili socio-economiche scelte riflettono l'attuale "stato di salute" sia delle famiglie sia delle imprese che si rivolgono al mercato dei beni contraffatti come antidoto contro la crisi: le famiglie pensano di fare buoni acquisti risparmiando e alimentando al contempo l'attività imprenditoriale della criminalità organizzata, le imprese sfruttano tali mercati per ridurre i costi di produzione competendo illegalmente con le imprese virtuose.

Risultati – L'analisi effettuata ha evidenziato che non vi è una significativa correlazione tra le provincie (meridionali) ad alta intensità criminale per reati di stampo mafioso e le provincie che rappresentano i mercati potenziali dei beni contraffatti. D'altra parte, ciò è intuibile anche dall'analisi grafica in cui si osserva che le provincie del Sud Italia rappresentano prevalentemente dei mercati locali per la distribuzione di merci contraffatte (Fig. 25). Diversamente, le altre variabili (predittori) analizzate confermano

⁷ Non sempre si è potuto disporre dei dati sul volume dei beni contraffatti specialmente per le provincie di recente costituzione: in tal caso si è proceduto a stimare i valori mancanti mediante una semplice operazione di ripartizione per quote del valore di tali beni.

le ipotesi di partenza: in un periodo di forte crisi economico-finanziaria sia le famiglie che le imprese ricorrono per ragioni diverse al mercato della contraffazione, rendendo ancora più vulnerabile l'economia del territorio. Tra le aree metropolitane italiane Roma e Napoli si confermano ai primi posti sia per numero di reati di contraffazione che per valore di beni sequestrati. L'analisi condotta ha confermato i risultati applicando sia il modello SAR che il modello CAR: le stime dei coefficienti e le statistiche test, presentano valori simili tra loro nei due modelli (Tab. 3-5).

Conclusioni

L'economia illegale altera le regole dell'economia e distorce il mercato, svilendo il lavoro, mortificando gli investimenti, distruggendo la proprietà intellettuale, ostacolando il credito, intimidendo la libertà di impresa. La presenza di criminalità modifica insomma la struttura del circuito economico e comporta un allontanamento strutturale dal modello di efficienza dell'economia di mercato, pregiudicandone la possibilità di conseguire l'*ottimo*.

L'obiettivo del presente lavoro è stato quello di esaminare i fattori di vulnerabilità del sistema socio-economico delle province italiane, coniugando la teoria assiomatica dei numeri indice con le metodologie tipiche della statistica spaziale e dell'econometria spaziale, applicate a *open data* "ufficiali", provenienti da diverse fonti e opportunamente armonizzati. La costruzione della matrice di vulnerabilità provinciale è in fase di sperimentazione non potendo ancora disporre completamente di tutta una serie di dati legati sia ai cosiddetti "reati spia", sottostimati, sia ai "reati sommersi", non catturati dalle statistiche ufficiali. La stessa matrice dovrà ulteriormente arricchirsi di dati temporali oltre che di micro-dati territoriali, sempre più puntuali. Le tecniche statistiche di analisi spaziale hanno consentito di validare e completare il quadro emerso sulla criminalità (organizzata) e sulla vulnerabilità del territorio tramite l'utilizzo dell'indice di Jevons, identificando le aree più vulnerabili e quelle in cui il fenomeno è ancora all'inizio ovvero sottostimato.

Le azioni di contrasto volte a frenare la diffusione di fenomeni criminali e l'adozione di strumenti di supporto alle imprese, specialmente nell'attuale situazione di crisi, costituiscono un valido deterrente sia per proteggere gli operatori economici dai condizionamenti della pressione mafiosa, sia per ricondurre il mercato nell'alveo delle "normali" regole di concorrenza economica. Ciò è tanto più vero in quei territori, non solamente delle

regioni meridionali ma inerenti anche diverse province dell'Italia centrale, che più di altri avrebbero bisogno di far leva sui propri fattori di competitività per superare le penalizzanti fragilità intrinseche del sistema produttivo e sociale, e che sono “ostaggio” di un'illegalità diffusa, che trova peraltro terreno fertile nella crisi economica. Il controllo delle istituzioni e del tessuto produttivo da parte della criminalità risponde non solo a motivazioni economiche ma alla volontà di rendere ben visibile alle comunità che il proprio dominio si estende anche alla *governance* locale.

Appendice

Tabella 1 - Matrice di Vulnerabilità provinciale							
Provincie	J1	J2	J3	J4	J5	VINDEX	V'INDEX
Torino	15,7	3,0	34,9	26,8	40,3	15,0	32,5
Vercelli	28,6	4,7	19,2	24,0	45,5	19,6	42,6
Novara	17,6	5,4	13,6	30,2	37,9	18,2	39,5
Cuneo	33,3	4,8	16,4	19,6	29,9	17,5	37,9
Asti	25,2	6,2	16,8	22,8	39,3	19,3	41,9
Alessandria	22,2	8,9	30,6	25,7	39,6	21,2	45,9
Aosta	45,2	3,1	19,4	25,4	23,1	16,9	36,6
Imperia	19,4	33,4	64,5	24,5	30,7	26,4	57,3
Savona	13,6	13,0	37,6	23,4	27,6	18,4	39,8
Genova	10,9	8,6	71,0	25,0	29,6	16,2	35,1
La Spezia	12,7	12,9	31,5	35,7	31,1	20,6	44,7
Varese	11,8	4,2	29,4	26,5	33,3	14,5	31,4
Como	22,4	8,9	17,5	27,4	25,7	19,4	42,0
Sondrio	49,5	10,2	23,0	24,8	26,0	23,9	51,8
Milano	11,5	4,2	38,0	30,6	29,7	14,5	31,5
Bergamo	18,5	7,5	19,5	28,9	35,5	19,4	42,1
Brescia	21,2	4,1	29,8	29,4	32,9	17,0	37,0
Pavia	22,0	4,1	22,1	26,1	46,1	18,2	39,4
Cremona	21,6	2,7	21,3	26,6	40,5	15,8	34,3
Mantova	28,7	1,7	20,2	26,4	43,2	15,2	33,1
Bolzano	37,4	2,0	14,9	23,7	17,3	13,2	28,7
Trento	31,5	10,0	11,8	24,0	22,9	20,4	44,2
Verona	18,0	3,6	26,6	25,8	23,6	14,1	30,5
Vicenza	19,6	5,7	12,4	26,5	33,4	17,8	38,6
Belluno	46,4	2,8	3,5	22,4	27,2	16,8	36,5
Treviso	19,2	2,9	11,0	26,9	31,2	14,7	31,9
Venezia	10,1	4,1	31,9	33,4	23,3	13,4	29,0
Padova	14,4	2,4	19,0	26,6	28,2	12,7	27,6
Rovigo	29,7	3,4	20,4	28,6	38,4	18,2	39,5
Udine	25,5	4,3	16,7	26,0	28,4	16,9	36,7
Gorizia	8,5	10,4	19,8	36,2	31,5	17,8	38,6

Tabella 1 - Matrice di Vulnerabilità provinciale

Province	J1	J2	J3	J4	J5	VINDEX	V'INDEX
Trieste	6,3	9,2	35,5	24,8	29,0	14,3	31,0
Piacenza	26,2	2,6	16,6	26,4	34,8	15,8	34,3
Parma	19,2	2,7	26,8	25,5	28,0	13,8	30,0
Reggio Emilia	23,1	2,7	19,8	25,4	35,7	15,4	33,4
Modena	18,3	4,0	19,0	26,5	36,1	16,3	35,3
Bologna	14,5	3,5	25,3	28,9	31,6	14,7	31,9
Ferrara	26,7	3,2	16,4	26,4	36,8	16,9	36,7
Ravenna	11,7	4,0	29,9	22,5	30,9	13,4	29,1
Forlì-Cesena	20,6	5,5	23,2	28,8	32,7	18,1	39,2
Pesaro e Urbino	23,1	4,2	29,0	33,4	32,4	18,0	39,1
Ancona	15,2	9,8	34,7	32,0	38,1	20,6	44,8
Macerata	27,9	6,3	27,6	32,8	36,1	21,3	46,3
Ascoli Piceno	21,8	10,8	26,6	28,6	39,7	22,8	49,4
Massa-Carrara	20,8	10,7	7,7	24,6	44,1	22,2	48,1
Lucca	16,7	9,2	25,9	30,7	31,0	19,6	42,5
Pistoia	19,9	5,1	29,8	34,7	32,1	18,3	39,8
Firenze	10,5	9,3	31,8	29,4	23,9	16,2	35,1
Livorno	9,3	17,4	43,1	26,4	31,9	19,2	41,7
Pisa	15,6	4,8	16,3	26,8	26,5	15,2	32,9
Arezzo	27,6	7,3	29,9	27,6	30,9	20,4	44,2
Siena	32,7	10,4	22,7	28,7	25,3	22,3	48,4
Grosseto	38,9	6,1	27,9	32,3	29,6	21,8	47,4
Perugia	26,6	13,6	21,6	31,9	34,4	25,1	54,5
Terni	24,2	12,5	18,4	30,5	37,8	24,3	52,7
Viterbo	32,8	8,5	8,6	27,4	45,8	24,3	52,8
Rieti	44,3	13,7	25,4	24,7	47,4	29,0	63,0
Roma	8,4	8,2	46,7	39,4	30,9	17,0	36,9
Latina	23,4	21,4	45,4	26,8	48,8	28,4	61,7
Frosinone	24,2	9,5	30,7	27,1	44,3	22,9	49,7
Caserta	23,3	10,2	33,2	27,6	53,4	24,3	52,7
Benevento	33,5	25,0	8,3	33,2	60,5	36,0	78,1
Napoli	11,4	22,3	46,7	29,9	49,3	24,7	53,6
Avellino	28,0	32,6	27,0	27,8	52,5	34,0	73,7

Tabella 1 - Matrice di Vulnerabilità provinciale

Province	J1	J2	J3	J4	J5	VINDEX	V'INDEX
Salerno	23,1	20,8	29,8	30,7	42,5	28,1	61,1
L'Aquila	40,1	14,0	23,5	26,1	41,2	27,9	60,4
Teramo	28,0	9,8	27,9	35,4	38,7	24,7	53,7
Pescara	17,7	7,3	35,7	34,9	43,4	21,0	45,6
Chieti	22,8	8,2	17,5	21,6	45,3	20,7	44,9
Campobasso	40,6	11,2	21,4	31,5	52,9	29,5	64,0
Foggia	38,2	13,7	52,1	26,1	47,5	28,4	61,6
Bari	17,9	15,8	35,5	28,9	45,4	24,7	53,6
Taranto	26,7	17,9	29,3	23,9	59,0	28,6	62,1
Brindisi	20,6	13,0	44,6	20,7	46,1	22,5	48,8
Lecce	24,9	15,7	29,1	28,7	45,1	26,7	57,9
Potenza	46,3	16,4	26,4	26,9	59,1	33,2	72,0
Matera	52,8	16,2	27,6	31,3	46,2	33,4	72,4
Cosenza	33,4	21,9	35,7	44,3	58,3	37,1	80,4
Catanzaro	28,1	9,3	38,3	38,3	51,8	26,8	58,2
Reggio Calabria	22,9	32,0	30,6	43,9	59,7	37,2	80,7
Trapani	22,9	7,8	26,7	36,8	49,3	23,9	51,7
Palermo	20,2	9,8	23,0	44,1	49,0	25,5	55,4
Messina	21,0	11,7	29,9	30,3	41,2	23,5	51,0
Agrigento	34,7	7,4	23,3	30,0	44,7	24,3	52,6
Caltanissetta	35,1	2,8	32,9	28,8	71,4	21,2	46,1
Enna	53,1	3,3	26,0	45,6	54,3	25,7	55,8
Catania	21,0	8,1	40,6	34,2	50,4	23,3	50,5
Ragusa	32,0	5,3	32,2	35,9	54,5	23,9	51,9
Siracusa	23,8	14,2	42,6	18,4	54,3	24,1	52,3
Sassari	27,4	9,6	46,2	33,6	45,0	25,1	54,5
Nuoro	76,4	12,4	53,1	30,7	39,1	32,6	70,8
Cagliari	25,3	11,8	33,2	23,9	42,7	23,5	50,9
Pordenone	31,8	4,5	9,8	30,1	32,1	19,3	41,8
Isernia	47,3	11,1	12,9	32,3	43,4	29,3	63,6
Oristano	45,1	4,7	8,5	33,3	43,6	23,5	51,0
Biella	28,3	3,8	22,6	29,1	45,5	19,4	42,0
Lecco	24,3	5,0	17,7	26,8	33,6	18,2	39,5

Tabella 1 - Matrice di Vulnerabilità provinciale

Province	J1	J2	J3	J4	J5	VINDEX	V'INDEX
Lodi	22,3	1,2	13,0	27,8	42,6	13,3	28,8
Rimini	12,5	18,5	35,8	31,9	26,1	20,9	45,4
Prato	21,2	15,4	29,8	36,1	36,4	25,6	55,5
Crotone	46,7	26,2	26,1	49,6	74,2	46,1	100,0
Vibo Valentia	29,1	45,3	45,0	43,2	44,0	39,8	86,3
Verbano-Cusio-Ossola	36,6	4,4	19,9	11,6	25,1	14,7	31,9
Olbia-Tempio	52,3	16,5	6,5	46,0	33,5	34,0	73,7
Ogliastra	85,4	15,7	53,1	35,8	36,1	36,3	78,7
Medio Campidano	50,6	3,3	33,2	38,5	56,6	24,5	53,2
Carbonia-Iglesias	32,7	5,8	33,2	38,5	53,2	25,0	54,2
Monza e della Brianza	11,5	0,5	15,3	25,2	34,7	8,4	18,3
Fermo	21,8	0,5	21,5	25,4	37,6	10,2	22,1
Barletta-Andria-Trani	40,4	0,7	31,7	36,3	39,5	14,1	30,6

Legenda - J1: Indicatore di vulnerabilità del territorio; **J2:** Indicatore di criminalità organizzata del territorio; **J3:** Indicatore spia di infiltrazione dell'illegalità economica (o della criminalità organizzata); **J4:** Indicatore di vulnerabilità delle imprese; **J5:** Indicatore di vulnerabilità delle famiglie; **VINDEX:** Indice di sintesi di vulnerabilità provinciale. Nel **VINDEX** si riportano i valori ottenuti direttamente dal calcolo dell'indice di Jevons, che variano tra 0 (escluso) e 100 (massimo valore di vulnerabilità che una provincia può assumere). Al contrario, nel **V'INDEX** i valori vengono riscalati da 0 (provincia meno vulnerabile) a 100 (provincia più vulnerabile).

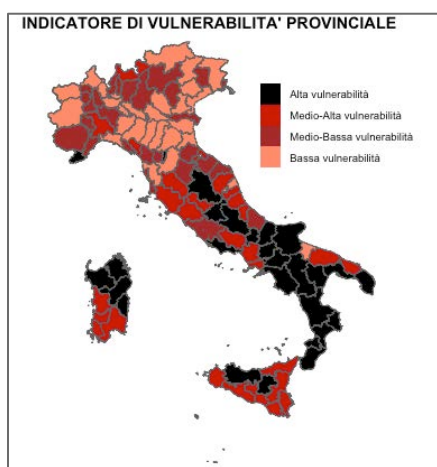


Figura 1: Elaborazione su open data

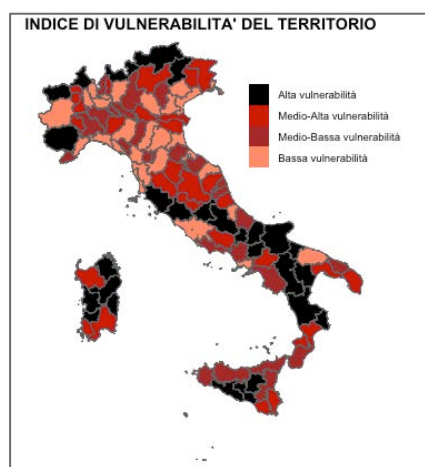


Figura 2: Elaborazione su open data

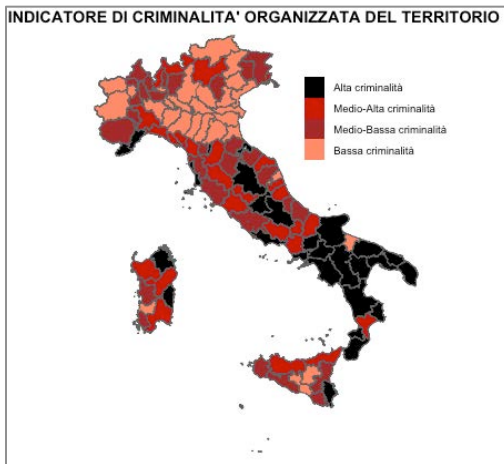


Figura 3: Elaborazione su *open data*

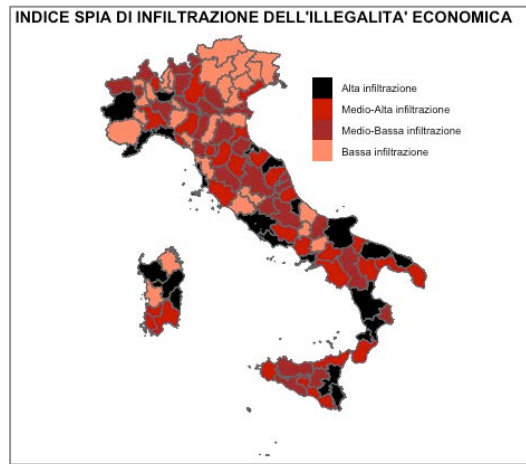


Figura 4: Elaborazione su *open data*

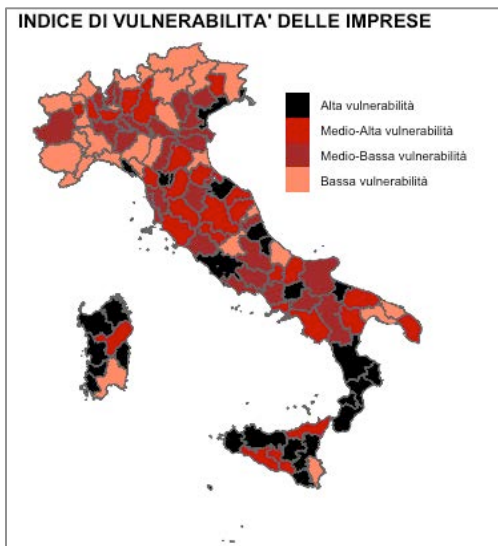


Figura 5: Elaborazione su *open data*

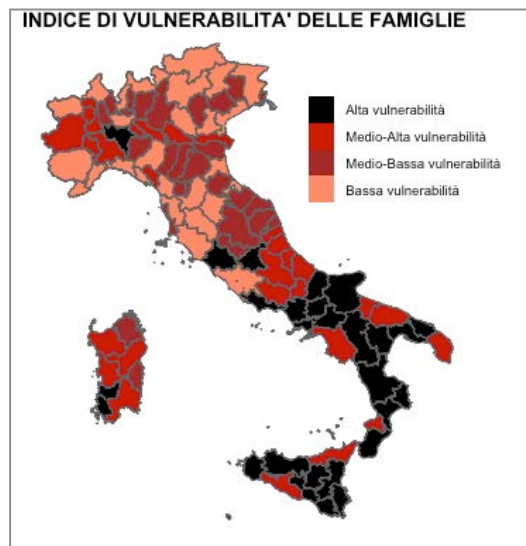


Figura 6: Elaborazione su *open data*

Tabella 2 – Statistica I di Moran

Indicatori	I di Moran	p-value (random.)	p-value (norm.)
J1	0.3617	0.0000	0.0000
J2	0.5233	0.0000	0.0000
J3	0.1150	0.0380	0.0390
J4	0.3761	0.0000	0.0000
VINDEX	0.6627	0.0000	0.0000

Distanza - Critical cut-off: 75 km

Legenda - J1: Indice di vulnerabilità del territorio; **J2:** Indice di criminalità organizzata del territorio; **J3:** Indice spia di infiltrazione dell'illegalità economica (o della criminalità organizzata); **J4:** Indice di vulnerabilità delle imprese; **J5:** Indice di vulnerabilità delle famiglie; **VINDEX:** Indice di sintesi di vulnerabilità provinciale

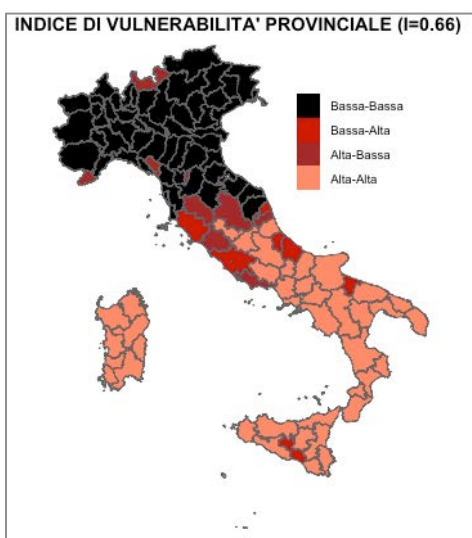


Figura 7: Elaborazione su open data

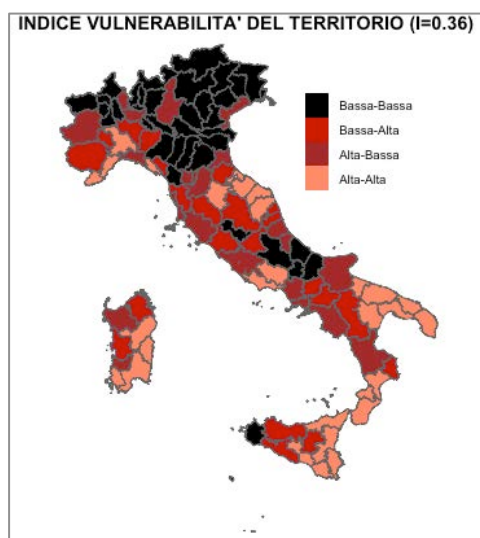


Figura 8: Elaborazione su open data

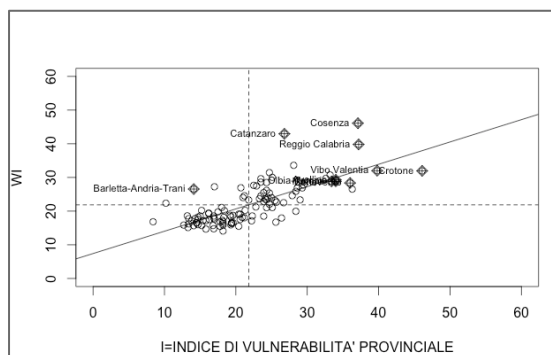


Grafico 1: Moran scatterplot – Elaborazione su fonti varie

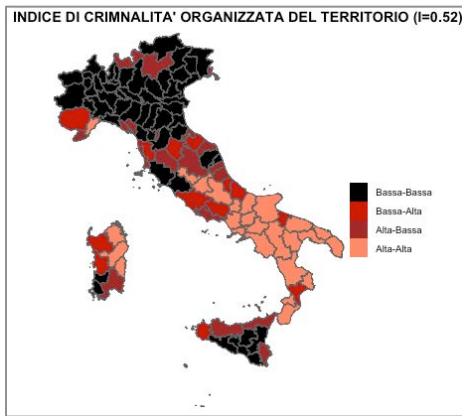


Figura 9: Elaborazione su *open data*

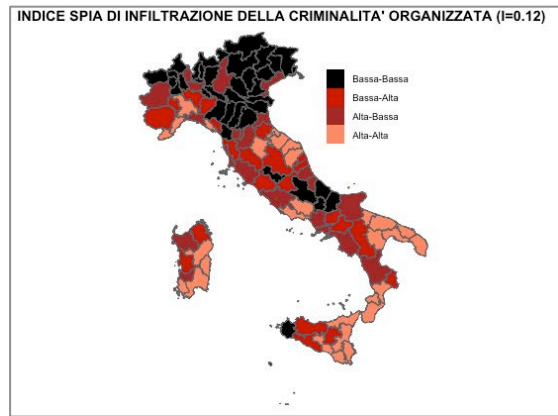


Figura 10: Elaborazione su *open data*

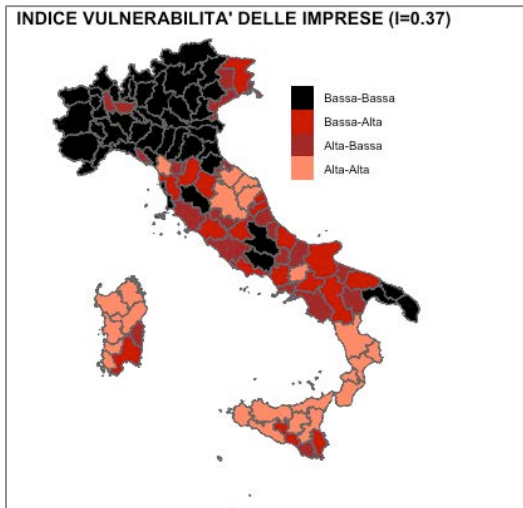


Figura 11: Elaborazione su *open data*

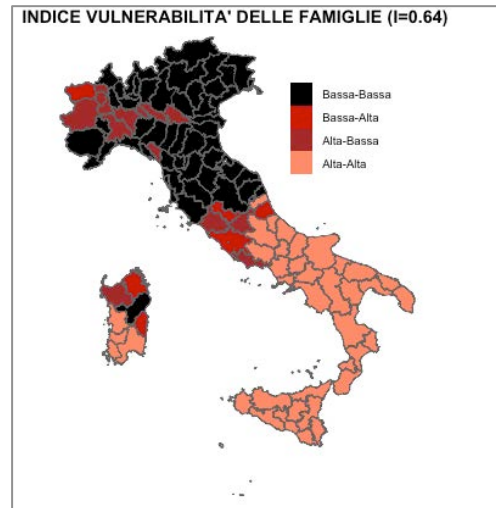


Figura 12: Elaborazione su *open data*

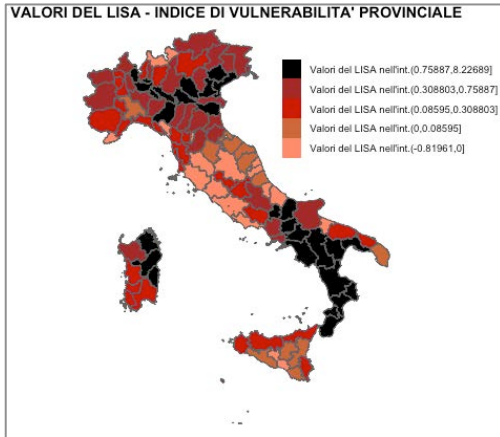


Figura 13: Elaborazione su open data

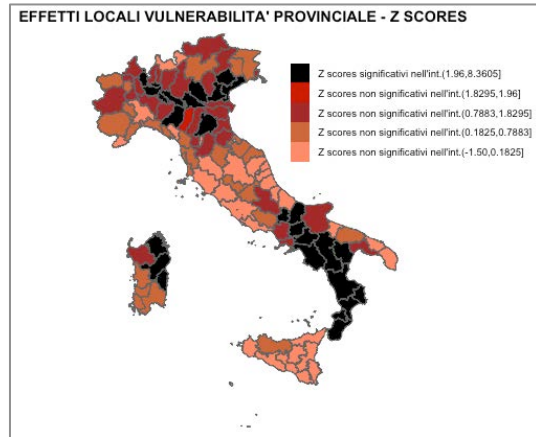


Figura 14: Elaborazione su open data

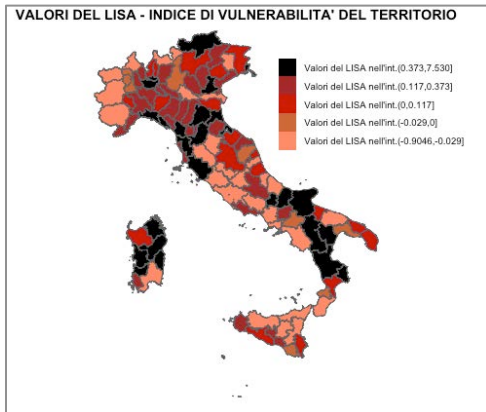


Figura 15: Elaborazione su open data

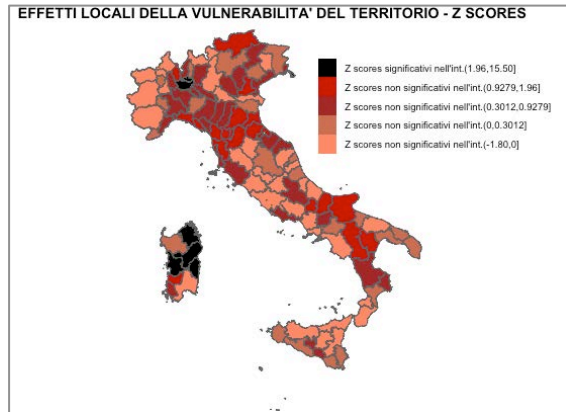


Figura 16: Elaborazione su open data

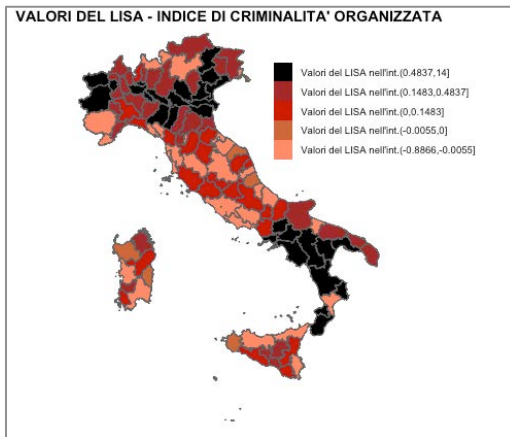


Figura 17: Elaborazione su fonti varie

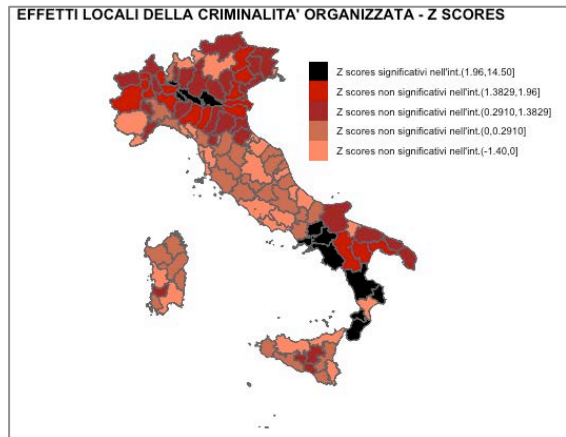


Figura 18: Elaborazione su fonti varie

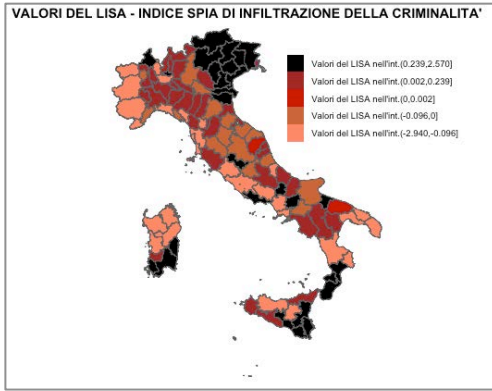


Figura 19: Elaborazione su open data

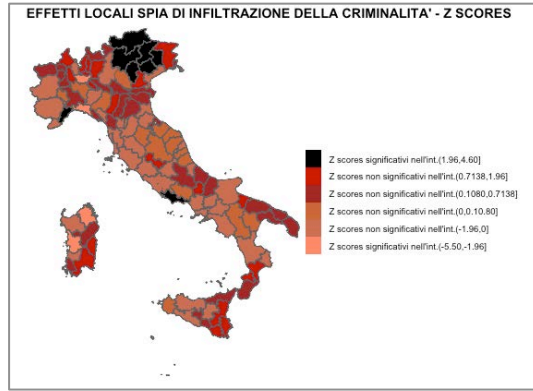


Figura 20: Elaborazione su open data

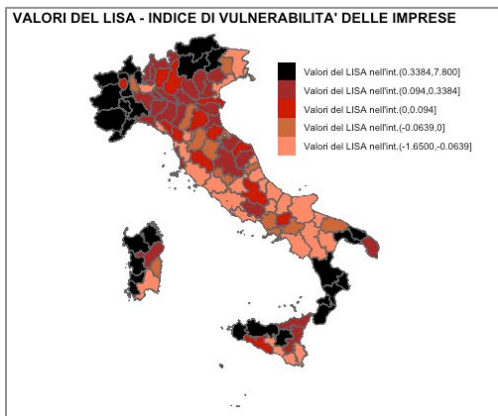


Figura 21: Elaborazione su open data

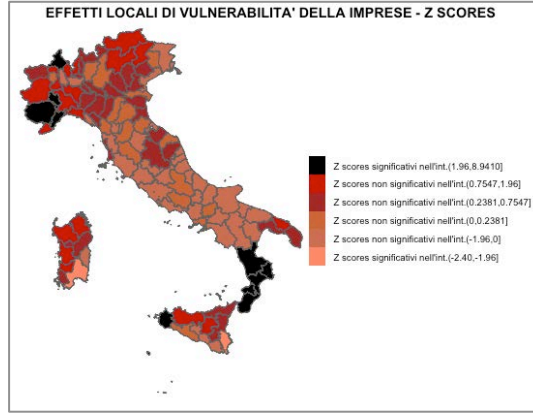


Figura 22: Elaborazione su open data

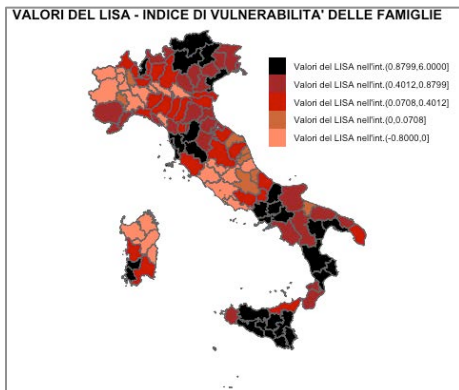


Figura 23: Elaborazione su open data

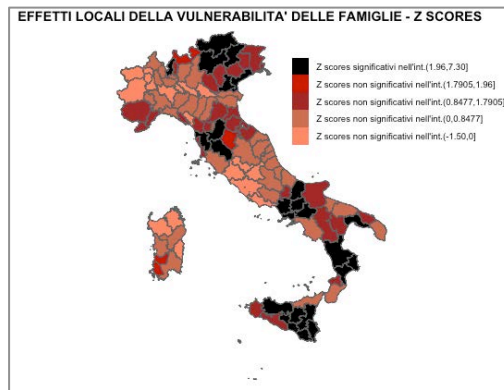


Figura 24: Elaborazione su open data

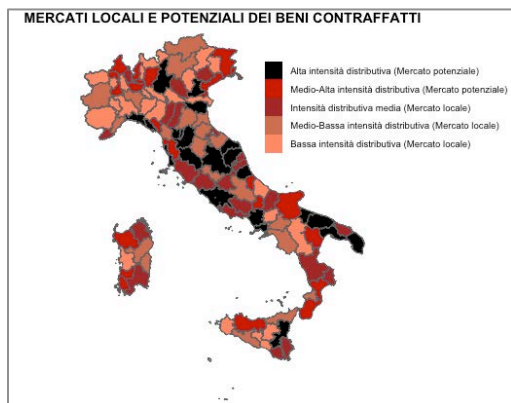


Figura 25: Elaborazione su *open data*

Tabella 3 – Modello SAR – Stima dei coefficienti

SAR Model	Stima	E. S.	Z valued	p-value
Intercetta	-19.025607	6.765646	-2.8121	0.004922
Crimine org	-0.110313	0.093660	-1.1778	0.238874
Reati contraf	0.255030	0.059791	4.2653	1.996e-05
Soff. imprese	0.282239	0.085999	3.2819	0.001031
Credito fam.	0.143818	0.082606	1.7410	0.041681

Tabella 4 – Modello CAR – Stima dei coefficienti

CAR Model	Stima	E. S.	Z valued	p-value
Intercetta	-19.025424	6.766659	-2.8116	0.004929
Crimine org	-0.110430	0.093674	-1.1789	0.238450
Reati contraf	0.255030	0.059794	4.2651	1.998e-05
Soff. imprese	0.282248	0.086014	3.2814	0.001033
Credito fam.	0.143853	0.082624	1.7411	0.041672

Tabella 5 – Modelli SAR e CAR – Statistiche

Statistiche	SAR	CAR
Lambda	-0.0044	-0.0086
LR test value	0.0164	0.0159
p-value	0.8979	0.8994
Log likelihood	-453.9222	-453.9224
ML (residual variance)	$\sigma = 14.993$	$\sigma = 14.992$

Bibliografia

- Atzeni S., Dettori B., Usai S. (2004) *L'econometria per le indagini territoriali*, Crenos
- Banca d'Italia (2012) *Statistiche - Base Informativa Pubblica On Line*.
- Besag J. (1974) Spatial interaction and the statistical analysis of lattice systems, *Journal of the Royal Statistical Society, Series B* 36, 192–225.
- Bivand R., Pebesma E.J., Gomez-Rubio V. (2013) *Applied spatial data analysis with R*, Springer, II edition, Berlino
- Cliff A. D., Ord J. K. (1969) The problem of Spatial autocorrelation, in *London Papers in Regional Science* 1, *Studies in Regional Science*, 25–55, Pion, London
- Cliff A. D., J. K. Ord (1973). *Spatial Autocorrelation*, Pion, London
- Cliff A. D., J. K. Ord (1981). *Spatial Processes: Models and Applications*, Pion, London.
- Commissione Antimafia (1993) *Relazione su: insediamenti e infiltrazioni di soggetti ed organizzazioni di tipo mafioso in aree non tradizionali*. Roma, 17 dicembre
- Cressie N. A. C. (1993) *Statistics for Spatial Data, rev. ed.* New York: John Wiley & Sons.
- Diewert W. E. (1976) Exact and superlative index numbers, *Journal of Econometrics*, Vol 4., pp. 115-145
- Diewert W. E. (1995) *Axiomatic and Economic Approaches to Elementary Price Indexes*. Cambridge: National Bureau of Economic Research. *NBER Working Papers* n. 5104
- Diggle, P. J. (2003). *Statistical Analysis of Spatial Point Patterns*, II edition Academic Press, London
- Draghi M. (2011) *Le Mafie a Milano e nel Nord: aspetti sociali ed economici*, Università degli Studi di Milano, 11 marzo
- Eichhorn W., Voeller J. (1976) Theory of price index: Fisher's test approach and generalizations, *Lectures notes in economics and mathematical systems*, Springer-Verlag, Berlino
- Espa, G., Filippini D., Giuliani D., Piacentino D. (2013) Decomposing regional business change at plant level in Italy: A novel spatial shift-share approach, *Papers in Regional Science*
- Infocamere (2012) *Banca dati Telemaco - Statistiche*
- Haining R. (1990) *Spatial Data Analysis in the Social and Environmental Sciences*, Cambridge University Press.
- Iperico (2013) *Banca dati sulla contraffazione*, MISE – Ministero dello

sviluppo economico

ISTAT (2012) *Banca dati degli Indicatori territoriali per le politiche di sviluppo*.

Legambiente (2013) *Ecomafia 2013*, Edizioni Ambiente.

Martini M. (1992) *I numeri indice in un approccio assiomatico*, Giuffrè Editore, Milano.

Martini M. (2001) *I numeri indice nel tempo e nello spazio*, Edizioni CUSL, Milano.

Mazziotta C., De Castris M., Mazziotta M. (2005) L'autocorrelazione spaziale presente in stime territoriali alternative della dotazione infrastrutturale, XXVI Conferenza italiana di scienze regionali, Bari

Mazziotta M., Pareto A. (2012) Indici sintetici per confronti spazio-temporali: un'applicazione alla dotazione infrastrutturale XXXIII Conferenza italiana di scienze regionali, Roma.

OECD (2008) *Handbook on Constructing Composite Indicators. Methodology and user guide*, OECD Publications, Paris, 2008.

Rao D. S. Prasada (2009) Generalized Eltetö-Köves-Szulc (EKS) and Country-Product-Dummy (CPD): Methods for International Comparisons, in *Purchasing Power Parities of Currencies: Recent Advances in Methods and Applications*, Ed. D. S. Prasada Rao, 86-120, Cheltenham, U.K., Edward Elgar.

Rao D.S. Prasada (2013) Computation on basic heading PPPs comparisons within and between regions, in *Measuring the real size of the world economy*, World Bank.

Ripley B. D. (1981) *Spatial Statistics*, New York: John Wiley & Sons.

Schabenberger O., Gotway C. A. (2005) *Statistical methods for spatial data analysis*, Chapman & Hall, London.

Tagliacarne (2012) *Starnet, XI Giornata dell'economia, Tavole di dati*

Tarantola A.M. (2012) *Dimensione delle attività criminali, costi per l'economia, effetti della crisi economica*. Testimonianza presso Senato della Repubblica, Camera dei Deputati, 6 giugno

Unioncamere-Ministero del Lavoro (2012) *Sistema Informativo Excelsior*

Unioncamere (2012) *XI Giornata dell'economia, Tavole di dati*

Unioncamere-Symbola (2012) *L'Italia che verrà. Industriale culturale, made in Italy e territori*

Waller L.A., Gotway C.A. (2004) *Applied spatial statistics for public health data*, John Wiley & Sons.

Whittle P. (1954) On stationary processes in the plane, *Biometrika* 41, 434–449.

Wong D.W.S., Lee J. (2005) *Statistical analysis of geographic information*,
John Wiley & Sons