POLITECNICO DI TORINO

DIPARTIMENTO DI INGEGNERIA GESTIONALE E DELLA PRODUZIONE

CORSO DI LAUREA MAGISTRALE IN INGEGNERIA GESTIONALE

Raffaele Scarlata

Il principio di Relatedness e lo Spazio delle Tecnologie: studio empirico sulla complessità delle attività brevettuali in Italia

Tesi di Laurea

Relatore: Prof. Luigi Buzzacchi

Correlatore: Prof. Antonio De Marco

Anno Accademico 2019/2020

ABSTRACT

Il tema centrale del presente lavoro di tesi è l'analisi dell'allocazione spaziale delle conoscenze tecnologiche in Italia, nonché il possibile impatto di tale attuale stato di distribuzione delle conoscenze tecnologiche sul "percorso" (inteso in termini di future tecnologie realizzabili) che i differenti territori "intraprenderanno" nel prossimo futuro e l'analisi dell'impatto che le diverse conoscenze tecnologiche hanno sull'aggregazione urbana dei territori stessi.

Lo studio dei succitati fenomeni si è reso possibile grazie a opportune metriche introdotte dalla *Relatedness Theory*, una recente branca dell' Economia dello Sviluppo, la quale, sebbene inizialmente nata per essere applicata alle varie nazioni del mondo al fine di studiare le cause del loro differente "sviluppo" - qui inteso in senso lato, ovvero in termini non solo reddituali ma anche umani e sociali – presenta una formalizzazione matematica e delle ipotesi applicative, che, sotto opportune condizioni, sono applicabili più in generale a una generica ripartizione spaziale (che non sia solo quella dei confini politici e amministrativi) di un territorio (consentendoci così di effettuare confronti anche per dimensioni territoriali minori) e a un'idea di prodotto ben più generica (intendendo infatti per prodotto un generico artefatto dell'ingegno umano, catalogabile secondo un'opportuna classificazione esaustiva ed esclusiva).

In quest'elaborato si adotteranno le ipotesi e le metriche della *Relatedness Theory* impiegando la ripartizione territoriale del territorio nazionale in *Sistemi Locali del Lavoro SLL* (i quali, essendo per definizione dei poli di attività produttive e sociali, rappresentano delle significative partizioni spaziali "autonome", parte di una realtà produttiva più grande – l'Italia – così come lo sono le singole nazioni – per il pianeta – nella teoria originaria qui mutuata) al fine di studiare la diversa distribuzione delle conoscenze tecnologiche in Italia. Come *proxy* di tali conoscenze sono stati adottati i brevetti, i cui volumi, ripartiti per *SLL* e per e per "classe" secondo la classificazione brevettuale *IPC* nel quindicennio 1999 – 2013, sono stati forniti dal *Politecnico di Torino*.

Da tale database, è stato possibile costruire lo "Spazio delle Tecnologie" dei brevetti Italiani (rappresentativo di tutte le combinazioni di similarità tra conoscenze tecnologiche necessarie per realizzare due specifici brevetti) ed è stato altresì possibile valutare, per tutti i *SLL* e per

tutte le categorie di brevetti (adottando, a tal proposito, sia lo standard IPC che la classificazione *NACE*) i valori delle metriche rappresentative della "diversità di conoscenza tecnologica" racchiusa nel territorio, delle sue "Opportunità", e della "complessità tecnologica" delle varie categorie brevettuali.

Infine, combinando il già citato database dei volumi brevettuali con il database contenente la ripartizione della popolazione nazionale per *SLL*, nell'arco temporale considerato, è stato valutato l'impatto che le singole categorie tecnologiche, oltreché ovviamente il volume di brevetti nella sua interezza, hanno sulla concentrazione demografica dei *Sistemi del Lavoro* Italiani.

ABSTRACT

The central theme of this thesis is the analysis of the spatial allocation of technological knowledge in Italy, as well as the possible impact of this current state of distribution of technological knowledge on the "path" (understood in terms of feasible future technologies) that different territories will "undertake" in the near future and the analysis of the impact that the different technological knowledge have on the urban aggregation of the territories themselves.

The study of the aforementioned phenomena was made possible thanks to appropriate metrics introduced by Relatedness Theory, a recent branch of Development Economics, which, although initially created to be applied to the various nations of the world in order to study the causes of their different "Development" - here understood in a broad sense, or in terms not only of income but also of human and social terms - presents a mathematical formalization and application hypotheses, which, under appropriate conditions, are more generally applicable to a generic spatial distribution (which is not is only based on the political and administrative boundaries) of a territory (thus allowing us to make comparisons even for smaller territorial dimensions) and to a much more generic product idea (in fact, by product we mean a generic artefact of human ingenuity, catalogable according to an appropriate exhaustive and exclusive classification).

In this paper, the assumptions and metrics of the Relatedness Theory will be adopted by employing the territorial division of the national territory in Local Labor Systems SLL (which, being by definition the poles of productive and social activities, represent significant "autonomous" spatial partitions, part of a larger productive reality - Italy - as well as individual nations - for the planet - in the original theory borrowed here) in order to study the different distribution of technological knowledge in Italy. Patents have been adopted as proxy for this knowledge, the volumes of which, broken down by SLL and by and by “class” according to the *IPC* patent classification in the fifteen years 1999 - 2013, were provided by the *Politecnico di Torino*.

From this database, it was possible to build the "Space of Technologies" of Italian patents (representative of all the combinations of similarity between technological knowledge

necessary to create two specific patents) and it was also possible to evaluate, for all SLL and for all categories of patents (adopting, in this regard, both the *IPC* standard and the *NACE* classification) the values of the metrics representing the "diversity of technological knowledge" contained in the territory, its "Opportunities", and the "technological complexity" of the various categories patents.

Finally, by combining the aforementioned database of patent volumes with the database containing the distribution of the national population by *SLL*, over the time period considered, the impact that the individual technological categories, as well as obviously the volume of patents in its entirety, was assessed have on the demographic concentration of the *Italian Labor Systems*.

INDICE

CAPITOLO I

INTRODUZIONE

Ripercorrendo gli annali della storia e prestando attenzione ai volumi di invenzioni prodotte nel tempo, un occhio attento potrà sicuramente notare come gli ultimi due secoli, rispetto al resto della storia dell'umanità, siano stati caratterizzati da un esponenziale accrescimento della "*conoscenza*" umana.

Se è indubbiamente vero che il tempo ha sempre, sin dalle epoche più remote, permesso alla conoscenza acquisita di diffondersi e, di conseguenza, migliorare la qualità della vita dell'uomo, alleggerendone i gravi e migliorandone la produttività, è innegabile come, negli ultimi due secoli, l'umanità abbia raggiunto successi fino ad allora neppure lontanamente immaginabili e reputati alla stregua di sogni: si pensi, a titolo d'esempio, al sogno di *Icaro* di volare e librarsi nel cielo, sogno rimasto tale fino alla conquista dei cieli da parte dei fratelli *Wilbur e Orville Wright*, ad inizio Novecento, o al viaggio di *Astolfo*, nell'*Orlando Furioso* di *Ariosto*, che approda sulla Luna per ritrovare il senno perduto di *Orlando*, viaggio considerato alla stregua di mito fino alla conquista della Luna, sul finire degli anni '60.

Questi, come molti altri riscontrabili in letteratura, sono solo alcuni degli esempi di come, in tempi recenti, l'uomo sia riuscito a mettere a frutto le sue conoscenze superando, in certi casi, anche i confini della propria immaginazione.

È possibile infatti oggi asserire, con sufficiente convincimento e scarsa perplessità, che la *conoscenza produttiva* dell'uomo, conoscenza intesa come l'insieme delle nozioni, teorie e pratiche adottabili e sfruttabili dallo stesso per alleggerire le fatiche della vita quotidiana come per permettergli di raggiungere ed esaudire i suoi più ambiziosi e reconditi desideri, abbia subito, e viva tuttora, una rapidissima *accelerazione*.

Al di là delle più prosaiche ed estreme conquiste dell'umanità, come le due succitate, sarebbe in realtà sufficiente confrontarci coi nostri nonni su quali fossero, in altri tempi, le loro abitudini e su cosa potessero o non potessero facilmente e rapidamente fare in gioventù, per intuire come la nostra vita

sia profondamente differente dalla loro e su come strumenti per loro un tempo rivoluzionari (i macchinari e gli utensili elettrici, la telefonia, i mezzi di trasporto a motore, ecc.) siano per noi strumenti il cui uso consideriamo oramai "innato" nelle nostre vite, o strumenti addirittura reputabili obsoleti e surclassati dalla incipiente tecnologia che avanza.

Di fronte a tale evidenza non può però che sorgere spontanea una domanda: qual è la ragione alla base dell'accelerazione subita dalla tecnologia negli ultimi due secoli? È l'uomo che è diventato più intelligente rispetto al passato o la risposta a tale domanda va ricercata in altre cause? La spiegazione di tale fenomeno va investigata probabilmente nel diverso approccio che alcune società moderne hanno avuto circa la diffusione della conoscenza tra i membri della collettività.

In tempi recenti, infatti, in certe società, a differenza delle precedenti epoche in cui l'innovazione avveniva per mano della genialità di sparuti, singoli, individui, le innovazioni sono state realizzate grazie all'opera congiunta di molti membri della collettività: tale diverso paradigma, oggi diffuso pressoché in tutto il globo grazie ai moderni mezzi di comunicazione di massa, si è potuto affermare grazie alla maggior celerità con cui le innovazioni e, più in generale, gli incrementi di *conoscenza produttiva* apportati da un singolo membro della collettività vengono, rispetto al passato, recepiti e adottati dalla collettività stessa; a fronte di un'innovazione da qualcuno introdotta in un certo *ambito*, i membri della collettività sfruttando, a loro volta, tale innovazione nei rispettivi *ambiti* della conoscenza (ambiti potenzialmente anche molto distanti da quello nel quale per primo l'innovazione è stata introdotta) possono apportare nuove innovazioni, a loro volta da altri sfruttabili, che arricchiranno ulteriormente la conoscenza produttiva della società.

Tale diverso approccio, rispetto al passato, verso la diffusione e l'adozione delle innovazioni è, indubbiamente, tra le ragioni alla base dell'accelerazione della conoscenza produttiva della società contemporanea, e, in virtù di ciò, possiamo pertanto asserire, in risposta alla domanda precedente, che non è tanto l'uomo a esser diventato più intelligente ma sono le società contemporanee a esser diventate più sagge nel foraggiare la diversità della loro conoscenza produttiva e le possibilità di ricombinazione della stessa alla ricerca di *prodotti* migliori e di più agevole ed efficiente uso.

Come già in precedenza accennato, è da evidenziare come il fenomeno dell'"*accumulazione sociale di conoscenza produttiva*" non sia stato un fenomeno dalla portata universale ma sia, invece, avvenuto in misura massiccia in certe parti del globo e in misura marginale, o solo in tempi relativamente recenti, in altre, generando una difforme distribuzione del benessere che ne consegue. A riprova di ciò, è infatti innegabile che, a livello globale, esistano enormi divari (in termini di conoscenza, reddito pro-

capite, sviluppo umano, ecc.) tra le diverse nazioni sul pianeta, divari, tra tante possibili spiegazioni, anche imputabili alla diversa conoscenza produttiva accumulata dalle nazioni nel corso della loro storia.

Tali divari, nell'accezione più generale percepiti come differenze reddituali o differente qualità della vita, nel corso della presente trattazione verranno invece intesi e calcolati come differenze, in termini di *diversità* e *raffinatezza*, tra i *prodotti* dei territori oggetti di confronto, impiegando metriche e indicatori oggi adottati nelle recente branca dell'*Economia dello Sviluppo* che ricade sotto il nome di *Relatedness Theory* (*Teoria della Relazionalità*), come sarà più ampiamente descritto in seguito; ma su quali ipotesi si fonda la *Relatedness Theory*?

In primis, si assume che un certo generico *prodotto*, per poter essere realizzato in uno specifico *territorio*, necessiti di una certa *qualità* e *quantità* di conoscenza produttiva già accumulata, ed è intuitivo dedurre che le citate caratteristiche di qualità e quantità di conoscenza produttiva richiesta varino enormemente tra i diversi prodotti, in relazione alla *complessità* degli stessi.

In secundis, poiché tipicamente, in particolare per i prodotti più moderni, la quantità di conoscenza richiesta per realizzare un certo prodotto è maggiore della conoscenza che un singolo uomo può accumulare nell'intero arco della sua vita, affinché tale prodotto possa essere effettivamente realizzato è fondamentale che vi siano delle relazioni tra soggetti detentori di conoscenza in differenti ambiti, e che, come risultato di tali interazioni, la conoscenza totale globalmente detenuta possa essere sfruttata e riassemblata per realizzare il nuovo prodotto.

Si pensi, ad esempio, a titolo chiarificatore, alla conoscenza richiesta per produrre un computer: la vastità e la complessità della conoscenza richiesta per realizzare un computer è così grande e specifica (si pensi alla tecnologia della batteria, a quella dello schermo, a quella dei processori, ecc.) che nessuno, neanche il più abile imprenditore del settore o il *geek* più esperto, può detenerla interamente. Da qui la necessità di un "incontro" tra molteplici individui, coordinati tra loro, dalle cui proficue relazioni nasce un prodotto ben più complesso del prodotto che un singolo individuo, solo, riuscirebbe a realizzare.

Da quanto fin qui detto potrebbe però sembrare che sia sufficiente il processo di accumulazione della conoscenza produttiva necessaria per realizzare un certo prodotto affinché questo possa essere realizzato in uno specifico territorio e che quindi, potenzialmente, qualunque prodotto possa essere realizzato in qualunque territorio. Tale deduzione è però in contrasto con l'evidenza empirica che ci dimostra come, nonostante vi siano prodotti in grado di generare maggior "valore aggiunto" e quindi più "appetibili" per un territorio, la produzione di questi si concentra solo in specifiche parti del globo

(si pensi, oggi, all'*high-tech* nella *Silicon Valley*, o, in passato, alla *Rivoluzione Industriale*, nell'*Inghilterra* del *'700*, a al *Rinascimento*, in *Italia*, nel *XV* e *XVI* secolo) dando luogo a sistematici miseri fallimenti dei tentativi di "attecchimento" della stessa in altre parti del pianeta.

L'evidenza euristica qui citata è dovuta al fatto che il processo di accumulazione di conoscenza produttiva è un percorso irto di difficoltà, in quanto questa è perlopiù una *conoscenza tacita* degli individui, non disponibile sui libri o su internet, e acquisibile più tramite l'esperienza che con lo studio, e in quanto tale difficile (oltre che costoso) da trasmettere e acquisire.

Introducendo un'opportuna suddivisione in *n* finite categorie, esaustive ed esclusive, di tutti i possibili prodotti dell'ingegno umano (cosicché non vi siano prodotti che non sia possibile classificare e che non vi siano prodotti ricadenti in più di una categoria) è possibile calcolare, sfruttando la *Relatedness Theory*, uno "Spazio dei Prodotti", ovvero un piano rappresentante le n^2 probabilità che due generici prodotti, *i* e *j*, della suddetta classificazione possano essere co-prodotti, ovvero che dalla produzione di uno si possa passare alla produzione dell'altro.

Nella caratterizzazione matematica che permette di costruire lo *Spazio dei Prodotti* si tiene ovviamente conto del fatto, già di per sé intuitivo, che, affinché un'industria possa attecchire in un certo territorio in questo debba già essere presente tutta o quantomeno buona parte della conoscenza produttiva richiesta: il lettore, infatti, converrà con lo scrivente che se provassimo a far attecchire, da un giorno all'altro, la produzione di reattori nucleari in un territorio la cui produzione prevalente è quella agricola, nella quale non risiedono soggetti con competenze nel settore dell'ingegneria meccanica e della fisica nucleare, e geograficamente lontano da centri nei quali tali competenze sono già presenti, le probabilità di un attecchimento di tale produzione sarebbero prossime allo zero; analogamente se provassimo a far attecchire l'industria della produzione di aeroplani in un territorio in cui l'industria prevalente è quella afferente la produzione di autoveicoli, essendo le conoscenze produttive richieste dalle due industrie non così dissimili tra loro, le probabilità di attecchimento sarebbero, invece, elevate (e l'esempio ora menzionato non è casuale essendo storicamente avvenuto che la *FIAT*, nel 1912, sfruttando le competenze già acquisite nel settore della meccanica, fondò con successo la *FIAT Avio*, oggi *Avio S.p.a.*, per produrre anche velivoli a motore).

La strategia di accumulo di conoscenza produttiva si presenta, quindi, come un "*Chicken and Eggs Problem*": il *Chicken and Eggs Problem* è il classico problema tipicamente riscontrabile nelle "*piattaforme a più versanti*" (quale ad esempio *eBay*, nella quale si incontrano, ai due versanti opposti, acquirenti e venditori), nella quale la creazione di valore per un gruppo è dipendente dal grado di

penetrazione della piattaforma nell'altro gruppo e viceversa, con tale creazione di valore che evolve in un circolo vizioso, in grado di autosostenersi e incrementarsi, una volta raggiunto un certo numero di individui in entrambi gruppi. Il *Chicken and Eggs Problem* consiste, appunto, nella ricerca della migliore strategia atta a innescare questo "effetto rete" quando non si hanno membri in nessuno dei gruppi.

Analogamente, uno statista alla guida di un paese si trova di fronte alla seguente domanda: dato lo stato attuale delle conoscenze produttive nel paese e noti i prodotti che determinano maggior valore aggiunto, come e quali conoscenza produttive provare a importare nel paese al fine di avviare il circolo virtuoso atto a incamminare il paese verso i *prodotti* "migliori"? Nella realtà, i paesi ottimizzano implicitamente la soluzione al *Chicken and Eggs Problem* "spostando" la propria produzione verso prodotti "vicini" (ovvero con elevato valore di prossimità all'interno dello Spazio dei Prodotti), in termini di conoscenza produttiva richiesta, a quelli che già produce (in quanto questi richiederanno una modesta aggiunta di conoscenza produttiva) creando però un'ovvia "dipendenza di percorso" (poiché il percorso che tali paesi dovranno intraprendere verso la produzione dei prodotti a più elevato valore aggiunto sarà fortemente dipendente dallo stato corrente della loro conoscenza produttiva e dalle possibilità di incremento della stessa per produrre i nuovi prodotti).

Tale soluzione ottima, al suddetto *Chicken and Eggs Problem* affrontato dai paesi, è stata più prosaicamente descritta dall'economista venezuelano *Ricardo Haussman*, uno dei padri della *Relatedness Theory*, mediante la seguente metafora:

"Si immagini che i prodotti siano gli alberi di una foresta e che due alberi siano vicini o distanti in base alla similarità delle conoscenze produttive necessarie per produrre i rispettivi prodotti (ovvero in base alla corrispondente *prossimità* nello *Spazio dei Prodotti*). Le aziende operanti in un paese sono come delle scimmie che traggono il loro sostentamento dallo sfruttamento dell'albero che occupano e assumiamo la foresta (corrispondente allo Spazio dei Prodotti) come data e identica per tutti i paesi.

Per massimizzare la propria utilità, le scimmie cercheranno di saltare dalla parte più povera alla parte più ricca della foresta ma la probabilità di farlo con successo non potrà che dipendere dalla produttività prevista degli alberi e da quanto "prossime" sono le scimmie ad alberi, non già occupati, che desiderano raggiungere, dove la "prossimità" tra gli alberi corrisponde alla possibilità di reimpiego degli *asset* già esistenti nel paese nella produzione del nuovo prodotto (maggiore è, più vicini i due alberi sono)".

La metafora di *Hausmman* descrive in modo esaustivo quella "dipendenza di percorso" nella quale i paesi si trovano nel tentativo di riuscire a produrre i prodotti caratterizzati dal maggior valore aggiunto.

Nella loro opera, "*The Atlas of Economic Complexity*", il già citato *Ricardo Haussman, Cesar Hidalgo e al*. misurano, tramite metriche opportune, la quantità di conoscenza produttiva detenuta da ogni paese (stimando l'*Indice di Complessità Economica* (*Economic Complexity Index o ECI*) di 128 paesi del mondo, stilando una classifica degli stessi sotto il profilo della conoscenza produttiva da questi detenuta) nonché la complessità insita nella produzione di specifiche famiglie di prodotti stimando altresì l'*Indice di Complessità di Prodotto* (*Product Complexity Index o PCI*) dei prodotti dell'industria, opportunamente categorizzati, stilandone una classifica)[1].

Sulle orme di quanto da loro già fatto, ricercando le stesse ipotesi e impiegando le medesime metriche, è obiettivo del presente lavoro di tesi applicare la letteratura da loro introdotta ad un contesto come quello italiano, sfruttando un apposito database, messo a disposizione dal *Politecnico di Torino*, contenente le informazioni circa tutti i brevetti registrarti, in Italia, nel quindicennio tra il 1999 e il 2013, e, sulla base dei conseguenti risultati, trarre le dovute conclusioni.

[1] L'ultima versione dell'*Atlas of Economic Complexity* contiene i dati del commercio interazionale di 250 tra paesi e territori del mondo, dati classificati in 20 categorie di prodotti e 5 categorie di servizi: i dati relativi al commercio di beni sono tratti dall' *United Nations Statistical Division* (*COMTRADE*), mentre i dati relativi ai servizi sono tratti dall' *International Monetary Fund* (*IMF*). Una precisazione va fatta in merito a quest'ultimi: a causa della loro natura intangibile il commercio dei servizi non viene registrato dagli uffici doganali a differenza dei dati sul commercio dei beni, e il volume dei servizi imporrati o esportati viene unilateralmente comunicato all'*IMF* dai singoli paesi; per tale votivi l'affidabilità di quest'ultimi è ritenuta inferiore rispetto ai dati del commercio dei beni.

CAPITOLO II

ANALISI E DESCRIZIONE DELLA LETTERATURA DI RIFERIMENTO

Nel presente capitolo verranno introdotte le nozioni, i criteri e la modellazione matematica degli strumenti essenziali adottati nella *Relatedness Theory*.

II.1. Descrizione logica delle metriche introdotte dalla *Relatedness Theory*

La *Complessità Economica* di un certo *territorio* dipende dalla varietà e profondità della conoscenza produttiva detenuta dai membri della collettività che opera nello stesso; tale *Complessità Economica* si riflette ovviamente su ciò che quel territorio produce o può produrre, esprimendosi quindi nella composizione del paniere di prodotti in esso realizzati.

L'evidenza empirica infatti dimostra che in un territorio non si producono tutti i prodotti di cui quel territorio necessita, ma solamente alcuni, ovvero quelli che in quel territorio si possono effettivamente produrre: tale possibilità dipende dalla complessità economica del territorio stesso, essendo oggettivo che prodotti semplici (ovvero richiedenti minori diverse competenze), quali ad esempio il legname o il caffè, possono essere realizzati da più territori a differenza di prodotti ben più complessi, quali i prodotti elettronici o farmaceutici, i quali richiedono la congiunta presenza di specifiche e rilevanti competenze.

Noto, quindi, cosa in un certo territorio si produce, come calcoliamo quali conoscenze produttive questo possiede? Per rispondere a questa domanda possiamo partire dalla semplice osservazione che, se i residenti e le organizzazioni operanti in un *territorio* sono in grado di produrre un certo *prodotto* allora in tale territorio devono necessariamente esservi le conoscenze produttive richieste per produrlo.

Inoltre, è anche intuitivo desumere che, al crescere del numero di diverse conoscenze produttive presenti in un territorio, cresce anche il numero di prodotti distinti che in quel territorio si è in grado di produrre, aumentando il numero di possibili ricombinazioni di tecnologia realizzabili: in altre parole possiamo affermare che la conoscenza produttiva presente in un territorio, è, in prima battuta, desumibile dalla diversità dei prodotti che in esso si si producono.

Analogamente è ragionevole assumere che, tanto più un prodotto è "complesso", tanto maggiori e tanto più diverse saranno le conoscenze produttive necessarie per produrlo e, di conseguenza, tanto più rari saranno i territori in cui, realizzandosi la congiuntura della contemporanea presenza di tutte tali conoscenze produttive, tale si prodotto sarà realizzabile: anche qui, con altre parole, possiamo asserire che prodotti che richiedono grandi volumi e variegate tipologie di conoscenza sono verosimilmente realizzabili solo laddove tali conoscenze sono presenti.

Dalle precedenti constatazioni, ne segue che l'*Ubiquità* di un prodotto, ovvero il numero di territori in grado di produrlo, ci può quindi già dare delle indicazioni circa i volumi e la diversità delle conoscenze produttive necessarie per produrlo, così come la *Diversità* produttiva di un territorio, ovvero il numero di distinti prodotti che in questo vengono prodotti, ci dà delle informazioni su quanto variegate siano le diverse conoscenze produttive presenti nello stesso. Appare inoltre ragionevole assumere, in virtù delle considerazioni di cui sopra, che è più probabile che prodotti più ubiqui (in quanto prodotti da un maggior numero di territori) siano quelli che richiedano un minor numero di distinte conoscenze produttive e che, viceversa, prodotti meno ubiqui (poiché realizzati da un più ristretto numero di territori) siano probabilmente quelli che richiedono una più ampia varietà di conoscenze produttive.

L'uso delle parola "probabile" nell'asserzione procedente non è causale; occorre infatti essere accorti nell'uso della metrica *Ubiquità* come *proxy* della quantità di diverse competenze produttive presenti in un territorio: potrebbe infatti accadere che un territorio produca un prodotto poco ubiquo non perché tale territorio sia uno dei pochi a possedere le competenze per produrlo ma perché il prodotto in questione sia "raro", ovvero sia, per specifiche implicazioni geografiche, producibile solo in quel paese e/o in pochi altri.

Tipico esempio di prodotti con il quale si può incorrere in tale errore sono le risorse naturali (carbone, oro, uranio, diamanti, ecc.): è ovvio che, essendo l'estrazione di tali risorse condizionate dalla loro effettiva presenza sul territorio, un altro territorio, pur avendo le competenze necessarie per produrlo, anche potendo non potrebbe farlo non avendo a disposizione la risorsa stessa.

A tale problema si può ovviare osservando contemporaneamente anche la metrica *Diversità* del territorio stesso: se il territorio in questione non è in grado di produrre anche altri prodotti a bassa *Ubiquità* allora, probabilmente, il motivo per il quale è in grado di produrre il primo prodotto poco ubiquo risiede nella "rarità" di questo e non nel fatto che tale paese abbondi di diverse conoscenze produttive.

Analogamente, così come è possibile sfruttare la metrica *Diversità* per correggere eventuali erronee considerazioni derivanti dalla sola osservazione dell'Ubiquità dei prodotti, è anche possibile il contrario: per un territorio avere un'elevata *Diversità*, ovvero produrre molti prodotti distinti, potrebbe indurci a pensare che il territorio esaminato abbondi di molte diverse conoscenze produttive e che sia quindi economicamente complesso. Se però, osservando l'*Ubiquità* dei suoi prodotti, questi sono pressoché tutti ubiqui è allora verosimile che quel territorio non disponga in realtà di variegate conoscenze produttive non riuscendo a produrre prodotti meno ubiqui e più complessi, e che non sia quindi economicamente complesso come si poteva in prima battuta immaginare, producendo sì molti diversi prodotti, ma tutti poco complessi.

Le due succitate metriche possono quindi essere usate congiuntamente, "correggendo" con le misure di *Ubiquità* quelle di *Diversità*, e reimpiegando poi quest'ultime per "correggere" l'*Ubiquità* stessa (o viceversa); tali mutue "correzioni" possono essere reiterate e, convergendo dopo poche iterazioni, ci restituiscono una misura implicitamente rappresentativa di "quante" e "quanto rare" siano le conoscenze produttive presenti in un territorio, metrica che sarà nel seguito chiamata *Indice di Complessità Economica ECI* (*Economic Complexity Index*).

Mutuando le considerazioni sin qui fatte con i territori, possiamo, per i prodotti, analogamente calcolare l'*Indice di Complessità del Prodotto PCI* (*Product Complexity Index*), ovvero una misura di "quante" e "quanto rare" siano le competenze necessarie per produrre un certo prodotto. A titolo chiarificatore, si riporta un esempio d'uso degli *Indici di Complessità Economica*, per realizzare un confronto tra due paesi:

Il *Pakistan* e *Singapore*, nel 2017, avevano un simile *PIL* (circa 300 Miliardi di *USD* riportati a prezzi di mercato) e, in base ai dati dell'*Atlas of Economic Complexity* del 2017, esportavano un simile numero di diversi prodotti (circa 133, adottando lo *Standard International Trade Classification 4* come criterio per la classificazione dei prodotti). Considerando, però, che la popolazione del *Pakistan* è circa 38 volte quella di *Singapore*, il reddito pro capite dei cittadini singaporiani è quindi 38 volte maggiore di quello dei pakistani.

Appare quindi lecita la domanda, perché tali divergenze di reddito tra i due paesi nonostante questi esportino lo stesso numero di diversi prodotti? Studiando l'*Ubiquità* dei prodotti esportati dal *Pakistan*, si può notare come i prodotti da questo esportati siano, a loro volta, esportati, in media, da 28 paesi nel mondo (posizionando il Pakistan nel 60° percentile dei paesi in termini di ubiquità media dei prodotti esportati); replicando tale calcolo per Singapore si ottiene, invece, che i prodotti che Singapore esporta,

sono a loro volta in media esportati da 17 paesi del mondo (1 ° percentile). Si ha quindi che il Pakistan esporta prodotti in media più ubiqui rispetto a quelli esportati da Singapore.

Per verificare che tale minore *Ubiquità* dei prodotti singaporiani rispetto a quelli pakistani sia dovuta alle più variegate conoscenze produttive detenute da Singapore, e non ad eventuali risorse "rare" da esso detenute, si può studiare l'*Ubiquità* media dei prodotti esportati da entrambi i paesi: da tale studio si osserva che i prodotti che Singapore esporta sono anche esportati, in media, da paesi che presentano elevati valori di *Diversità*, mentre quelli il *Pakistan* esporta sono esportati, a loro volta, da paesi scarsamente diversificati. Poiché i prodotti esportati da Singapore sono anche esportati da altri paesi con elevata *Diversità* (che, per quanto fin qui detto, presentano quindi tante e variegate conoscenze produttive), tale evidenza, insieme alla precedente minore *Ubiquità* dei suoi prodotti, ci spinge ad affermare che Singapore è "*economicamente più complesso*" del *Pakistan*.

II.2 Formulazione matematica delle metriche della *Relatedness Theory*

Nel seguito la rappresentazione formale, in termini matematici, delle metodologie di calcolo delle metriche fin qui introdotte.

II.2.1. Chi produce cosa?

Anche in virtù degli esempi fin qui citati, in base a quale criterio possiamo asserire che in un certo territorio si "produca" effettivamente qualcosa? È sufficiente che venga realizzata anche una sola unità di quello specifico prodotto o è più opportuno introdurre criteri più stringenti?

Secondo la *Relatedness Theory*, per associare un prodotto a un territorio si ritiene opportuno rendere tale associazione dipendente dal confronto del volume di quello specifico prodotto realizzato in quel territorio con i volumi di tutti i prodotti realizzati in tutti i territori considerati (ovvero la produzione globalmente realizzata).

Per far ciò, si adotta la definizione, data dall'economista *Balassa*, di *Vantaggio Comparato Rivelato RCA* (*Revealed Comperative Advantage*): secondo tale definizione un territorio rivela, quando comparato con i restanti territori, un vantaggio nella produzione di un certo prodotto quando la quota di prodotto realizzata in quel territorio è maggiore della quota "equa" del prodotto stesso, ovvero la quota che tale prodotto rappresenta nella produzione globale.

Facciamo un esempio: nel 2008, la *Soia* ha rappresentato lo 0,35% delle esportazioni mondiali ed il 7,8% delle esportazioni brasiliane; l'indice *RCA* per il prodotto "Soia" per il paese "Brasile" è, secondo la definizione di *Balassa*, pari a "21" (7,8% / 0,35%). Essendo tale misura maggiore di "uno", possiamo affermare che il *Brasile* gode di un *Vantaggio Comparato Rivelato* nella produzione della *Soia*.

In formule, indicando con X_{CP} la produzione del prodotto *P* in un territorio *C*, possiamo formalmente esprimere il *Vantaggio Comparato Rivelato RCA* che il territorio *C* ha nella produzione del prodotto *P* come:

$$RCA_{cp} = \frac{X_{cp}}{\sum_c X_{cp}} \Big/ \frac{\sum_p X_{cp}}{\sum_{c,p} X_{cp}}$$

Calcolando tale indice *RCA* per i tutti i territori, e, per ciascuno territorio, per tutte le categorie di prodotti, è possibile costruire una matrice M_{CP} che collega ogni paese a tutti i prodotti (specificando in quale prodotto il paese ha un vantaggio comparato e in quale no) o, viceversa, ogni prodotto a ciascun paese (indicando così quale paese rivela un vantaggio comparato nella produzione di quel prodotto e quale no): in tale matrice, di dimensioni *C* x *P* e composta interamente di *0* e *1*, il generico paese *C'* rivela una vantaggio comparato nel prodotto *P'* se l'incrocio della riga C' e della colonna P' è pari a 1, non lo rivela se tale incrocio è pari a 0. Tale matrice risponde quindi alla domanda precedente, specificando "chi produce cosa".

II.2.2. Come misuriamo la *Complessità*?

La *Complessità* di un *Territorio* o di un *Prodotto*, come logicamente descritto al *Paragrafo 2.1.*, la possiamo stimare, almeno in prima battuta, a partire dalle metriche di *Ubiquità* e *Diversità*. Queste, analiticamente, le possiamo calcolare come segue: a partire da un certo *dataset* contenente i volumi delle produzioni di un insieme di territori (produzioni categorizzate secondo una classificazione *esclusiva* ed *esaustiva* dei prodotti), si può calcolare la precedentemente descritta matrice M_{CP} indicativa di "chi produce cosa"; da questa possiamo derivare le metriche *Ubiquità* e *Diversità* semplicemente sommando le righe o le colonne della suddetta matrice. Formalmente:

$$Diversità = k_{C,0} = \sum_P M_{CP}$$

$$Ubiquità = k_{P,0} = \sum_C M_{CP}$$

dove lo "zero" ivi presente (la seconda cifra indicata in entrambi i pedici) sottolinea come che tali metriche di *Diversità* e *Ubiquità* siano calcolate all'inizio del processo iterativo, ovvero semplicemente a partire dalla semplice matrice M_{CP}.

Si rende necessario sottolineare ciò poiché, come descritto in precedenza, l'uso delle metriche, prese singolarmente, può indurre a grossolane ed errate valutazioni dell'effettiva *Complessità* di territori e prodotti; per generare, allora, una più accurata misura del numero di "conoscenze" disponibili in un certo territorio, o richieste per realizzare un certo prodotto, occorre correggere le informazioni contenute in $k_{C,0}$ e $k_{P,0}$, usando l'una per correggere l'altra. Tale processo di correzione implica:

- La correzione della *Diversità* di un territorio tramite l'*Ubiquità Media* dei prodotti che esso produce, a sua volta corretta con la *Diversità Media* dei paesi che producono tali prodotti.
- La correzione dell'*Ubiquità* di un prodotto tramite la *Diversità Media* dei territori che producono tale prodotto, a sua volta corretta con l'Ubiquità Media dei prodotti realizzati in tali paesi.

Ciò implica l'instaurarsi di un ciclo ricorsivo, che converge dopo poche iterazioni, nella quale ad ogni *N -esima* iterazione può essere espresso, formalmente, dalla seguente espressione ricorsiva:

$$k_{c,N} = \frac{1}{k_{c,0}} \sum_p M_{cp} \cdot k_{p,\ N-1}$$

$$k_{p,N} = \frac{1}{k_{p,0}} \sum_c M_{cp} \cdot k_{c,\ N-1}$$

Inserendo nel seguito, per semplicità di lettura, solo le espressioni di $k_{C,N}$ (sottolineando che per $k_{P,N}$ i procedimenti sono mutuabili scambiando semplicemente l'indice *C* con l'indice *P*), indichiamo con *N* la *N – esima* iterazione dell'iter ricorsivo. Se sostituiamo la (4) alla (3) otteniamo:

$$k_{c,N} = \frac{1}{k_{c,0}} \sum_p M_{cp} \frac{1}{k_{p,0}} \sum_{c\prime} M_{c\prime p} \cdot k_{c',N-2}$$

la quale la possiamo riordinare come:

$$k_{c,N} = \sum_{c'} k_{c',N-2} \frac{1}{k_{p,0}} \sum_p \frac{M_{cp}\, M_{c'p}}{k_{c,0}\, k_{p,0}}$$

che possiamo riformulare come:

$$k_{c,N} = \sum_{c\prime} \widetilde{M_{cc\prime}} k_{c\prime,N-2}$$

dove:

$$\widetilde{M_{cc\prime}} = \frac{1}{k_{p,0}} \sum_{p} \frac{M_{cp}\, M_{c\prime p}}{k_{c,0}\, k_{p,0}}$$

Si osserva che la penultima precedente equazione converge quando il vettore $k_{C,N}$ rispetta la condizione $k_{C,N} = k_{C,N-2} = 1$ e, sotto tali condizioni, tale vettore $k_{C,N}$ è l'autovettore di $M_{CC'}$ associato al suo più elevato autovalore: in realtà, l'autovettore associato al più elevato autovalore è un vettore di 1 e un tale vettore non contiene valore informativo.

Per ovviare a ciò, cerchiamo quindi l'autovettore associato al secondo più elevato autovalore, che è l'autovettore che cattura la più grande varianza nel sistema e che rappresenta la nostra misura di *Complessità Economica*. Da qui, definiamo l'*Indice di Complessità Economica* (*ECI*) come:

$$ECI = \frac{\vec{K} - < \hat{K} >}{stdev(\vec{K})}$$

dove $< K >$ e $stdev(K)$ rappresentano, rispettivamente, la media e la deviazione standard dei valori dell'autovettore:

$$\vec{K} = Autovettore\ di\ \widetilde{M_{cc\prime}}\ associato\ al\ secondo\ pi\grave{u}\ elevato\ Autovalore$$

Replicando quanto fin qui fatto per i prodotti, anziché per i territori, definiamo l'*Indice di Complessità di Prodotto PCI*:

$$PCI = \frac{\vec{Q} - < \hat{Q} >}{stdev(\vec{Q})}$$

Dove $< Q >$ $stdev(Q)$ rappresentano, rispettivamente, la media e la deviazione standard dei valori dell'autovettore:

$$\vec{Q} = Autovettore\ di\ \widetilde{M_{pp\prime}}\ associato\ al\ secondo\ pi\grave{u}\ elevato\ Autovalore$$

II.2.3. Come si evolve la Complessità Economica ?

La *Relatedness Theory* pone grande enfasi sul ruolo che l'*Indice di Complessità Economica ECI* assume nello spiegare le differenze reddituali pro-capite o le future prospettive di crescita di un territorio: le analisi condotte dai suoi fautori sembrano infatti dimostrare che la *Complessità Economica* di un territorio influisce sul suo sviluppo economico, anche più di altri indici spesso usati da enti governativi, ed è proprio questa capacità nel ben spiegare le differenze tra i territori che ne rende di così pregevole l'uso.

Sorge però spontanea la domanda: ma la complessità economica di un territorio è immutabile? Ebbene, ripercorrendo mentalmente gli annali della storia focalizzandoci su un certo territorio, qui a titolo d'esempio il *Regno Unito*, e studiandone le caratteristiche produttive, possiamo osservare come questo si sia evoluto da paese prevalentemente agricolo e pastorale, come a fine '400, a paese esportatore di prodotti meccanici e tessili tra la fine del '700 e l'inizio dell''800, al paese oggi focalizzato sui servizi e i prodotti finanziari.

Se avessimo a disposizione i dati precisi circa i volumi e le tipologie di prodotti esportati dal *Regno Unito*, o da qualunque altro paese o territorio, nel corso degli ultimi 2000 o 3000 anni, potremmo calcolarne l'*Indice di Complessità Economica* nel corso degli anni, studiandone l'evoluzione nel tempo.

Sicché, sorgono spontanee le domande: ma come si evolve la *Complessità*? In che modo le società aumentano la quantità di conoscenza produttiva in esse contenuta? Cosa limita la velocità di questo processo? E perché accade in alcuni luoghi ma non in altri?

Come detto, la *Relatedness Theory* ci dice che l'economia di un certo territorio riflette la quantità di conoscenza produttiva che questo contiene. La conoscenza, costosa da acquisire e difficile da trasferire, è, almeno ideologicamente e per agevolezza d'uso, possibile modularla "in blocchi" che chiameremo nel seguito "*capacità*": le c*apacità*, come già accennato, sono difficili da accumulare poiché l'accumulo determina un *Chicken and Eggs Problem*: da una parte, infatti, i territori non possono creare prodotti che richiedono *capacità* non già hanno, d'altra parte, però, ci sono scarsi incentivi per accumulare *capacità* nei luoghi in cui le industrie che li richiedono non esistono.

Ciò è particolarmente vero quando le c*apacità* mancanti richieste da una nuova potenziale industria sono molte: in tal caso, infatti, fornire una sola delle tante c*apacità* mancanti sarà molto probabilmente non sufficiente per lanciare la nuova industria, data il permanere dell'assenza di tutte le altre *capacità* complementari richieste. Viceversa, laddove fossero poche le *capacità* mancanti richieste

per lo sviluppo di una nuova industria, queste potrebbero essere più facilmente sviluppate e introdotte poiché sarebbero minori sono i costi e gli sforzi di coordinamento necessari per l'accumulo di diverse nuove *capacità* contemporaneamente.

Per i motivi fin qui detti, i territori tendono con più probabilità a trasferirsi verso prodotti "simili" a quelli che già producono, ovvero prodotti che sfruttano le *capacità* già incorporate nel territorio stesso.

Ma come facciamo a misurare analiticamente la "somiglianza" tra i requisiti di *capacità* di prodotti diversi? Il più rigoroso degli approcci richiederebbe l'accumulo di un volume indescrivibile di informazioni, quale lo stato della tecnica, i requisiti funzionali e istituzionali di ogni prodotto, e molto, troppo, altro ancora. Tale approccio, ovviamente non praticabile, si può però superare con un semplice accorgimento derivante dall'osservazione: due prodotti che richiedono simili *capacità* saranno più probabilmente co-realizzati dal medesimo produttore, e tanto maggiore sarà la somiglianza tra le *capacità* richiesta tanto maggiore sarà la suddetta probabilità. Per fare un esempio, è più probabile che in un certo territorio si co-producano i prodotti "Camicie" e "Pantaloni", o "Camicie" e "Foulard", che "Camicie" e "Reattori Nucleari".

Dall'appena citata osservazione, possiamo calcolare la "prossimità" tra tutte le coppie di prodotti, data una certa categorizzazione degli stessi; A partire dalla matrice M_{CP}, come già descritto rappresentativa di "chi produce cosa", la *prossimità* tra due prodotti, ovvero la probabilità di produrne uno condizionata alla già esistente produzione dell'altro, la possiamo formalmente scrivere come:

$$Prossimità_{PP'} = \Phi_{PP'} = \sum_C M_{CP} M_{CP'} \Big/ \max(k_{P,0}, k_{P',0})$$

La rappresentazione di tutte le *prossimità* esistenti, che dal punto di vista algebrico è una matrice di dimensioni *n* x *n* (con *n* numero di categorie di prodotti), si potrebbe schematizzare, rappresentando gli *n* prodotti come *n* punti su un piano, alla stregua di una "rete" la cui trama si compone di tutti i collegamenti esistenti tra le coppie di prodotti, collegamenti che, se adottassimo ad esempio come criterio di rappresentazione della *prossimità* lo spessore della linea, sarebbero tanto più spessi quanto maggiore è la probabilità che quei due prodotti vengano co-prodotti. A prescindere dalle sue modalità di rappresentazione definiamo tale matrice/rete come "Spazio dei prodotti" e la utilizziamo per studiare la struttura produttiva dei paesi.

Lo *Spazio dei Prodotti* influisce infatti sulla possibilità dei paesi di passare dai prodotti correnti ai nuovi: due prodotti strettamente collegati condividono infatti la maggior parte della capacità richieste e un territorio che già ha ciò che serve per realizzare uno dei due prodotti troverà relativamente facile passare alla produzione dell'altro. Ne deduciamo quindi che tanto più connesso è lo *Spazio dei Prodotti* (ovvero tanto più "spesse" sono le corde della sua maglia) tanto più facile sarà per un territorio accrescere la sua complessità economica; viceversa, quanto più scarsamente connesso è lo *Spazio dei Prodotti* tanto più difficile è per un paese accrescere la sua complessità.

Ma, come la distanza tra alberi nella già citata metafora di Hausmann, distanza che poteva essere più o meno superabile dalla scimmie arrampicate su un certo albero, se lo *Spazio dei Prodotti* è analogamente molto eterogeneo potrebbero esservi gruppi di prodotti altamente correlati, e tali per cui l'aggiunta di nuove *capacità* e l'espansione verso tali prodotti potrebbe essere relativamente facile, e altri prodotti, più distanti e poco correlati, tali da rendere il processo di accumulazione e diversificazione delle *capacità* più difficoltoso e complicato.

A titolo chiarificatore, in analogia alla metafora di *Haussmann*, qual è la forma dello Spazio dei Prodotto in cui viviamo? È un mondo in cui la foresta è fitta o rada?

Tale forma è riportata nella *Figura 2.1.* alla pagina successiva; tale figura è una rappresentazione dello Spazio dei Prodotti costruita utilizzando i dati del commercio internazionale per gli anni 2006-2008: in esso i nodi rappresentano i prodotti, la loro dimensione è proporzionale al commercio mondiale totale di quel prodotto, mentre i collegamenti tra nodi ivi rappresentati si instaurano tra i prodotti eventi un'alta probabilità di essere co-esportati [2].

Tale rappresentazione dello *Spazio dei Prodotti* rivela che questo tende ad essere molto eterogeneo: alcune sezioni di esso sono dense di prodotti, tra loro fittamente collegati, mentre altri prodotti tendono ad essere più periferici, sparsi e poco collegati;

[2] La rappresentazione dello *Spazio dei Prodotti* dell'Immagine precedente è stata realizzata per mezzo di un algoritmo che cercasse anche di evitare di rappresentare tutte quelle prossimità tra prodotti ritenute "insignificanti". Il set di collegamenti rappresentato prende il nome di *Maximum Spanning Tree MST* della *Matrice di Prossimità*, ed è stato ottenuto cecando di minimizzare il numero di collegamenti tra nodi a massimizzando la somma delle *prossimità*; tale algoritmo prende il nome di *Algoritmo di Kruskal*.
Dopo la selezione dei collegamenti significativi, la Matrice di Prossimità è stata rappresentata tramite un algoritmo "*Force-directed layout*", nella quale i nodi si repellono a vicenda, come cariche elettriche concordi, al fine di creare una rappresentazione che rappresentasse quali gruppi di nodi sono più densi e quali meno.

In tale immagine risulta infatti evidente che alcuni prodotti tendono a raggrupparsi in “comunità”, rendendo agevole il passaggio da un prodotto all’altro, mentre altri si trovano isolati, rendendo così più difficoltoso il “*salto verso un albero più ricco*”.

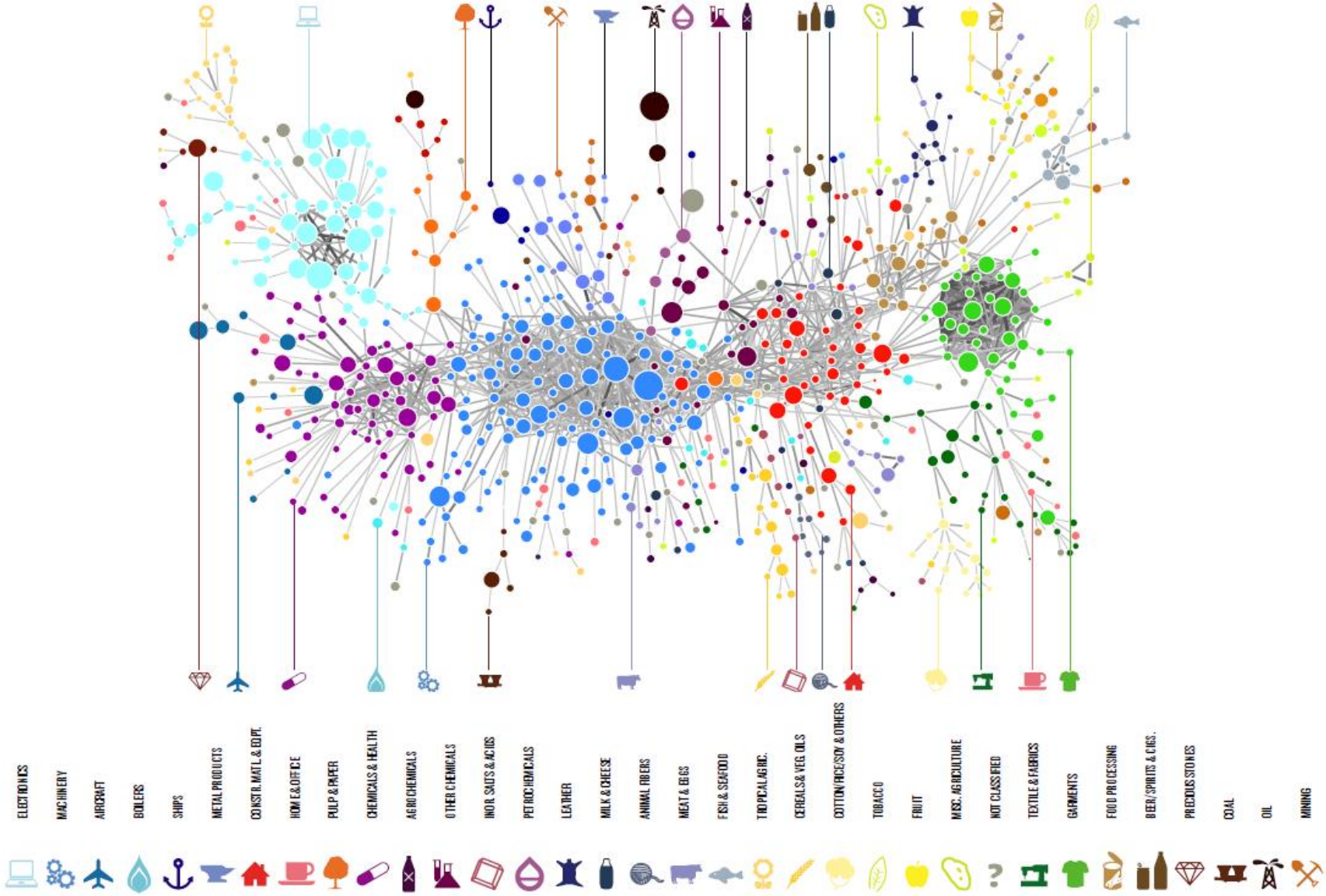

Immagine 2.1.: Spazio dei Prodotti costruito sulla base dei dati del commercio mondiale dal 20006 al 2008 (Il colore è specifico della tipologia del prodotto, la dimensione è proporzionale al volume delle esportazioni, lo spessore delle linee è proporzionale alla prossimità tra prodotti)

II.2.4. Verso quali prodotti conviene "spostarsi" ?

Assodato quindi che i prodotti che un territorio produce (e la posizione quindi assunta dalle sue aziende nello *Spazio dei Prodotti*) catturano informazioni cica lo stato delle conoscenze produttive che quel territorio possiede e la sua possibilità di ampliare tali conoscenze passando ad altri prodotti "vicini", come valutiamo, però, quanto tale territorio è "distante" dagli altri prodotti che potrebbe produrre?

Effettivamente la metrica *prossimità* precedentemente introdotta misura, più che la "vicinanza" di un territorio alla produzione di un certo prodotto, la "somiglianza" tra le capacità produttive richieste per realizzare due distinti prodotti; un territorio, però, non produce un solo prodotto ma un paniere di prodotti, e in virtù di ciò necessitiamo di una metrica che ci permetta di quantificare quanto "distanti" siano, dalla produzione corrente di un territorio, tutti i prodotti che invece correntemente non fa.

Chiamiamo tale misura "distanza" d_{cp} e la definiamo, per ogni territorio *c*, come la somma delle sue *prossimità* $\phi_{pp'}$ che collegano tutti i prodotti *p'* già prodotti dal territorio stesso al prodotto *p*, non già prodotto, verso il quale si desidera calcolare la "distanza". Tale misura è poi normalizzata divenendo per la somma delle prossimità $\phi_{pp'}$ verso il prodotto *p*.

La distanza così introdotta si può definire come una distanza pesata (e proporzionata a una scala di valori compresa tra 0 e 1) della distanza esistente tra un prodotto *p* e tutti i rimanenti prodotti *p'* che il territorio *c* già produce (dove i pesi sono dati dai valori di *prossimità*).

Se il territorio già produce la maggior parte delle merci collegate al prodotto, la distanza sarà breve (cioè prossima a 0); se altrimenti il territorio produce solo una piccola parte dei prodotti correlati a quello rispetto al quale cerchiamo di calcolare la distanza, la distanza con tale prodotto sarà grande (cioè prossima a 1). Formalmente:

$$d_{cp} = \sum_{p'} (1 - M_{cp'}) \, \Phi_{pp'} \Big/ \sum_{p'} \Phi_{pp'}$$

La metrica *distanza* d_{CP}, così introdotta, ci dà un'idea di quanto sia distante un certo prodotto *p* dato il corrente mix produttivo del territorio *c*. Tuttavia, è ancor più utile avere una misura olistica delle "opportunità" date dalla posizione di un paese nello *Spazio dei Prodotti*: avendo osservato, infatti, che i territori che producono i prodotti più complessi tendono a crescere più velocemente, ha senso tener conto non solo della distanza tra i prodotti, ma anche della loro *complessità*.

A parità di distanza d_{CP}, è infatti differente, sotto il profilo delle *opportunità*, se un paese si trovi vicino a un prodotto poco collegato e relativamente semplice, o se si trovi vicino a un prodotto complesso e altamente connesso. Questo significa che i paesi differiscono non solo per ciò che fanno, ma anche per le loro *opportunità*. Possiamo pensare a tale metrica come, per un territorio *c*, il valore della possibilità di passare alla realizzazione di altri prodotti.

Definiamo quindi la metrica "Opportunity Value", rappresentativa delle prospettive non sfruttate, che formalizziamo analiticamente come segue:

$$OV_C = Opportunity\ Value_C = \sum_{p'} (1 - d_{cp'}) \cdot (1 - M_{cp'}) \cdot PCI_{p'}$$

Dove $PCI_{p'}$ è l'"*Indice di Complessità del Prodotto*" del prodotto generico *p'* e il termine $(1 - M_{CP'})$ fa in modo che si contino solo i prodotti che il paese non già produce.

Un *Opportunity Value* più elevato implica l'essere più "vicino" a nuovi prodotti e/o a prodotti più complessi, e si propone come criterio di indirizzo nel definire il percorso dei paesi nella loro strada verso il futuro.

II.3. La relazione esistente tra la Complessità delle attività economiche e la concentrazione demografica nelle aree urbane.

Un'ulteriore evidenza che è possibile trovare in letteratura parlando di *Relatedness Theory* e *Complessità Economica*, è la relazione "superlineare" esistente tra la *Complessità* delle attività economiche e la concentrazione demografica nelle diverse aree urbane di un paese.

Ne "*Complex economic activities concentrate in large cities*", pubblicazione di *Hidalgo & C*. del gennaio 2018 nel *SSRN Electronic Journal*, si presenta una dimostrazione empirica che, quanto più un'attività economica è complessa, tanto più ampia è la sua tendenza a concentrarsi nelle grandi città.

Impiegando i dati storici sui brevetti, nella succitata pubblicazione si dimostra come la concentrazione, nei grandi centri urbani, di invenzioni inerenti i prodotti più complessi, come ad esempio nel settore delle "biotecnologie" o dei "semiconduttori", sia in continuo aumento almeno dal 1850, mentre quella inerente le tecnologie meno complesse è costantemente diminuita dagli anni '70.

Ma perché le attività economiche complesse dovrebbero concentrarsi maggiormente nelle grandi città? Da quanto fin qui discusso si può effettivamente desumere che all'aumentare della complessità di un processo economico aumenta anche il numero di diversi settori della conoscenza necessari per generarlo. Tale eterogeneità di conoscenze richieste è però, innegabilmente, più facilmente ed efficientemente ottenibile nelle grandi città, nella quale è più facile incontrare persone con esperienze e conoscenze diverse, e nella quale, di conseguenza, è più facile che si "incontrino" tutte le conoscenze richieste per "passare" a un nuovo prodotto.

Riformulando in termini più formali ed economici, possiamo asserire che, in una grande città (in termini di popolazione), data la più variegata e balcanizzata divisione della conoscenza e le molteplici opportunità di interazione offerte, si comprimono, rispetto a una città più piccola, i costi di coordinamento necessari affinché le conoscenze si incontrino.

Ma come facciamo a convalidare e a misurare tale tendenza delle tecnologie più complesse a concentrarsi maggiormente nelle più grandi città? Nella pubblicazione, utilizzando i dati storici sui brevetti negli Stati Uniti d'America a partire dal 1850, considerando tali brevetti come una *proxy* della conoscenza, si dimostra che, laddove maggiormente prosperano forme più complesse di conoscenza, lì la popolazione è aumentata in maniera più che proporzionale, almeno nell'ultimo secolo e mezzo.

Tali risultati spiegano perché alcuni processi economici si concentrano in modo sproporzionato in certe città e contribuiscono alla nostra comprensione generale dell'organizzazione spaziale dell'economia.

In termini analitici, per dimostrarlo, si è assunta l'esistenza di una generica "legge di scala" del tipo " $y \sim x^{\beta}$ ", con $\beta > 0$, nella quale x rappresenta la *popolazione* di un centro urbano e y rappresenta il *numero di brevetti* prodotti nel centro urbano stesso. Lo studio di tale funzione, a meno di un fattore costante che moltiplica *x*, essendo $\beta > 0$ e non potendo essere *x* negativo, implica:

- *y* crescente a tassi decrescenti se $\beta < 1$;
- *y* linearmente crescente se $\beta = 1$;
- *y* crescente a tassi crescenti se $\beta > 1$;

Adottando quindi tale relazione come "modello econometrico" e stimando β tramite regressione, in base alla stima di β si può dedurre l'effetto di una specifica tecnologia sull'aggregazione urbana.

Ad esempio, *Hidalgo & C.* stimano, per gli Stati Uniti e prescindendo dalla tecnologia, un generico esponente "$\beta_{Brevetti} = 1{,}26$" (ovvero un discreto effetto di super linearità che prescinde dalla tecnologia) che assume però, nel caso del settore "Computer, Hardware e Software", un valore ben più preponderante arrivando a "$\beta_{Computer} = 1{,}57$" e assumendo, invece, i connotati ben più modesti nel settore "Tubi & Giunti" con "$\beta_{Tubi\ \&\ Giunti} = 1{,}10$".

Il risultato principale dalla pubblicazione citata è quindi l'aver dimostrato la relazione positiva esistente tra la *Complessità* di una classe tecnologica, nell'accezione fin qui descritta, e la concentrazione urbana delle pubblicazioni brevettuali della stessa.

Si constata quindi che, nonostante il progresso nell'ambito dei trasporti e delle comunicazioni che avrebbero potuto dettare un cambio di passo nell'evidenza fin qui descritta, le disuguaglianze spaziali in termini di localizzazione produttiva dei diversi prodotti non sono tramontate, e ciò in virtù, probabilmente, del fatto che la conoscenza necessaria per svolgere le attività maggiormente complesse, è per lo più conoscenza tacita che viaggia meglio nelle reti sociali tra individui che attraverso i canali di comunicazione digitale.

CAPITOLO III

ANALISI DEI DATI E METODOLOGIA DI STUDIO

Nella letteratura fin qui descritta, seguendo l'approccio metodologico della *Relatedness Theory* e al fine di calcolare le metriche finora descritte, si pone sempre in relazione un *territorio* (che può essere un paese, un centro urbano o, più in generale, una porzione di uno spazio più grande frammentato secondo un comune e ben definito criterio), e la *produzione* dello stesso (produzione, anche qui, definita in senso lato: può essere la produzione di beni secondo specifiche categorie merceologiche, la produzione di brevetti secondo specifiche categorie tecnologiche, ecc.).

Scopo del presente elaborato è quello replicare i modelli e le metriche propri della *Relatedness Theory* fin qui descritti, applicandoli ad uno specifico *set* di dati, al fine studiare il comportamento dello stesso alla luce della letteratura scientifica in merito.

Il *dataset* impiegato, fornito dal *Politecnico di Torino*, contiene i dati brevettuali (suddivisi in base alla classe tecnologica secondo lo standard *IPC*), per ogni *Sistema Locale del Lavoro* (nel seguito anche *SLL*) italiano, dal 1999 al 2013: nell'analisi che segue considereremo pertanto come *territori* i *Sistemi Locali del Lavoro* e come *produzione* i volumi brevettuali in ciascuno di essi, raggruppati secondo il succitato standard.

Nel seguito del presente capitolo si esplorerà più nel dettaglio cose è lo *Standard IPC*, cosa sono i *Sistemi Locali del Lavoro* e quali sono le caratteristiche del *database* impiegato.

III.1. I Sistemi Locali del Lavoro SLL

Secondo quanto dettato dall'*Istat* e contenuto nella *Nota Metodologica*, consultabile sul suo sito *Istat.it*, i *Sistemi Locali del Lavoro* "rappresentano dei luoghi (precisamente identificati e simultaneamente delimitati su tutto il territorio nazionale) dove la popolazione risiede e lavora e dove, quindi, indirettamente tende a esercitare la maggior parte delle proprie relazioni sociali ed economiche".

La costruzione della griglia nazionale atta a definire univocamente su quale *SLL* ricade un certo territorio è figlia di un algoritmo, di nome *EURO* (adottato anche dall'*Eurostat* per la predisposizione di SLL armonizzati sul piano europeo), il quale costruisce la suddetta griglia aggregando più comuni perseguendo l'obiettivo della massimizzazione del livello d'interazione tra i comuni appartenenti allo stesso *SLL*; tale "interazione" è data dai flussi $f_{h,k}$ di pendolarismo giornaliero tra il *luogo di residenza* (località *h*) e il *luogo di lavoro* (località *k*).

L'obiettivo di massimizzazione delle interazioni, in virtù del quale si effettua la ripartizione, deve essere, inoltre, soggetto al rispetto dei vincoli imposti sulla dimensione minima delle aree, espressa tramite il *numero di occupati residenti* R_i

$$R_i = \sum_k f_{i,k} = f_{i\cdot}$$

e sul livello minimo accettabile di auto-contenimento dei flussi di pendolarismo, distinto tra auto-contenimento dal lato dell'offerta di posti di lavoro *SCO*

$$SCO = \frac{f_{ii}}{f_{i\cdot}}$$

e auto-contenimento dal lato della domanda di posti di lavoro *SCD*

$$SCO = \frac{f_{ii}}{f_{\cdot i}}$$

dove

$$f_{\cdot i} = \sum_h f_{h,i} = W_i$$

sono i *posti di lavoro nella località i*, e $f_{ii} = RW_i$ sono gli *occupati che risiedono e lavorano nella località i* (ovvero gli spostamenti interni).

L'introduzione dei *Sistemi Locali del Lavoro*, i cui confini esulano dai perimetri delle tradizionali aree amministrative in cui è classicamente suddiviso il territorio nazionale, nasce dal bisogno di individuare, analizzare e studiare le caratteristiche scoiali ed economiche di aree nella quale si hanno processi di auto-organizzazione della popolazione attiva, aree perimetrate grazie agli spostamenti medi giornalieri compiuti dalla stessa per coniugare la vita professionale con quella sociale.

La definizione dei principi cardine per la definizione dei "sistemi locali" nasce dall'attività di collaborazione tra gli Istituti di Statistica nazionali europei, coordinati dall'*Eurostat*, i quali propugnavano la definizione di *Labour Market Areas* armonizzate a livello europeo, mediante l'adozione di principi e metodi comuni. Da tale collaborazione sono stati nati i principi comuni da seguire per la costruzione dei *SLL*:

- *Scopo*: ciascuna zona rappresenta un mercato del lavoro;
- *Rilevanza*: le zone permettono di diffondere informazione statistica affidabile e confrontabile;
- *Completezza*: le zone sono una partizione dell'intero territorio dello stato.
- *Unitarietà*: ciascun comune può appartenere a una sola zona;
- *Contiguità*: ciascuna zona è costituita da un insieme di comuni contigui.
- *Coerenza*: ciascuna zona è costituita da un insieme di comuni non frazionati.
- *Conformità*: le zone possono non rispettare i confini amministrativi.
- *Omogeneità*: le zone non sono troppo estese territorialmente o troppo numerose in termini di occupati.

Da tali principi deriva l'importanza dei *SLL*, ovvero la possibilità di creare una geografia confrontabile e coerente dell'intero territorio italiano (ed europeo) che possa essere d'ausilio nell'analisi d'importanti fenomeni socio-economici quali quelli del mercato del lavoro.

L'algoritmo di costruzione adottato (creato da *Coombes* e *Bond* nel 2007) è di tipo deterministico iterativo *single step* e rappresenta un'evoluzione della più classica metodologia di costruzione delle "*Travel-To-Work Areas*" (già definita da *Coombes et al.* nel 1986, e adottata, sotto varie forme, in numerosi paesi europei tra cui l'Italia).

L'idea alla base dell'algoritmo, come già accennato denominato *EURO*, è quella di superare il concetto di soglia unica sia sul *numero di occupati residenti* sia sulle *funzioni di auto-contenimento* (ovvero, la

ricerca di una partizione tale da superare, per ogni SLL, un valore prefissato di entrambe le variabili) che caratterizzava la precedente metodologia.

Per far ciò si introduce invece un *trade-off* tra *occupati residenti* e *auto-contenimento*: a fronte di valori elevati (superiori a una soglia prefissata, definita *target*) di entrambe le *funzioni di auto-contenimento* (come in precedenza definite) si accettano *Sistemi Locali del Lavoro* di dimensioni "ridotte" (ovvero ai quali si richiede la sola condizione di avere un numero di occupati residenti superiore a una soglia minima prefissata), mentre per *Sistemi Locali del Lavoro* di dimensioni "maggiori" (ovvero con un numero di occupati residenti superiore a un valore definito *target* prefissato) la soglia di accettazione su entrambe le funzioni di auto-contenimento diminuisce (e sono accettati, in analogia al precedente, solo valori superiori a una soglia minima di auto-contenimento).

Tale *trade-off* tra ampiezza dimensionale, espressa dagli occupati residenti, e auto-contenimento minimo del *SLL* è governato dalla definizione di quattro parametri: un valore *minimo* e un valore *target* sia per il numero di occupati residenti, rispettivamente *minSZ* e *tarSZ*, sia per le funzioni di auto-contenimento, *minSC* e *tarSC*, dove *SC* (*Self-Containment*) è da intendersi come

$$SC = \min(\, SCO, SCD \,)$$

ovvero il minimo tra le due funzioni di auto-contenimento prima introdotte.

Il parametro *minSZ* definisce la "soglia minima di occupati residenti" che un *SLL* deve raggiungere mentre il parametro *tarSC*, rappresenta il "livello di auto-contenimento minimo" che *sistemi locali* di piccole dimensioni devono raggiungere per essere considerati accettabili; quest'ultimo valore sarà abbastanza elevato in modo da costituire sistemi che effettivamente abbiano stringenti caratteristiche di auto-contenimento.

In merito invece agli altri due parametri, l'interpretazione di questi è legata alla diminuzione del livello di auto-contenimento, fino *minSC*, che si è disposti ad accettare per *sistemi locali* che presentano dimensioni superiori a *tarSZ*. mentre la dimensione target, *tarSZ*, va intesa, più che come un obiettivo stringente del modello, come "il livello dimensionale minimo per il quale si è disposti ad accettare una riduzione del livello di auto-contenimento".

Definite tali variabili, l'algoritmo è strutturato in modo tale da creare *SLL* che rispettino i vincoli di dimensione e auto-contenimento prefissati; per raggiungere tale obiettivo si partiziona il piano cartesiano "numero di occupati" su un asse, "minimo delle funzioni di auto-contenimento" come precedentemente definito sull'altro, in due regioni: la prima di accettazione caratterizzata da *Sistemi*

Locali che soddisfano il *trade-off* sopra descritto e la seconda di rifiuto dove i vincoli assegnati non sono rispettati.

La "soglia di accettazione" è definita tramite un ramo d'iperbole che rappresenta la "condizione di validità":

$$\frac{minSC}{tarSC} \leq \left[1 - \left(1 - \frac{minSC}{tarSC}\right) \cdot max\left(\frac{tarSZ - SZ}{tarSZ - minSZ}; 0\right)\right] \cdot \left[\frac{min(SC;\ tarSC)}{tarSC}\right]$$

Il punto di partenza dell'algoritmo è una griglia che definisce dei "proto-SLL", ovvero zone costituite da una, o più località, che devono essere esaminate dall'algoritmo: un *proto-SLL* è accettato come *SLL* se i relativi valori di dimensione *SZ* (numero di occupati residenti) e auto-contenimento *SC* soddisfano la precedente condizione di validità; se ciò non accade, si cerca il *proto-SLL* o *SLL* con il quale esistano maggiori legami per effettuare l'unione delle due aree.

I *passi* dell'algoritmo iterativo possono essere riassunti schematicamente come segue:

1. Inizialmente ogni comune è considerato un *proto-SLL* e per ciascuno si calcola la funzione di validità, ovvero la quantità a destra della diseguaglianza di cui sopra;
2. Fin quando esistono *proto-SLL* non "validi", ovvero che non soddisfano la diseguaglianza sopra riportata:
    a. Determinare il *proto-SLL* che minimizza la funzione di validità e disaggregarlo nei singoli comuni costituenti;
    b. Fino a quando sono presenti comuni singoli non assegnati:
        i. Per ogni comune non assegnato identificare il *proto-SLL* (o *SLL*) dominante ossia quello che massimizza la *funzione di coesione* (*cohesion*):

$$L_{hk} = \left[\frac{(f_{hk})^2}{(f_{h\cdot}\, f_{\cdot k})}\right] + \left[\frac{(f_{kh})^2}{(f_{k\cdot}\, f_{\cdot h})}\right]$$

        ii. Assegnare la località al (*proto*)*SLL* dominante;
3. Ricalcolare la funzione di validità.

L'algoritmo quindi ripropone la medesima iterazione di separazione/aggregazione fino alla convergenza a una soluzione dove tutte le aree della partizione soddisfano i vincoli.

Il grande vantaggio del nuovo algoritmo per la determinazione dei *SLL*, risiede nel fatto che questo è più funzionale allo scopo stesso per il quale tali *SLL* sono stati introdotti, superando i limiti del precedente criterio di aggregazione di comuni, costruito sulla base del solo auto-contenimento della

domanda, con l'introduzione, invece, di un criterio più funzionale che coinvolge, oltre all'auto-contenimento minimo (sia domanda che offerta) anche la dimensione del *SLL* espressa tramite la precedente disequazione.

Tale vincolo, infatti, se da un lato risulta più restrittivo imponendo ai *SLL* di piccole dimensioni un'auto-contenimento elevato sia per la domanda sia per l'offerta, dall'altro, definendo un *trade-off* con il numero di occupati residenti del *SLL*, permette una maggiore flessibilità al *SLL* di dimensioni medio-grandi.

Le principali conseguenze di tali nuovo criteri di definizione sono la scomparsa di *SLL* di piccole dimensioni caratterizzati da un auto-contenimento dell'offerta debole, e la possibilità di modulare le aree intorno alle grandi città favorendo la formazione di *SLL* medio-grandi in modo da garantire una maggiore omogeneità della partizione.

Ad oggi (e dal 2011), la combinazione dei 4 precedenti parametri scelta per rappresentare i *SLL* è la seguente: *minSC* = 0,60; *tarSC* = 0,75; *minSZ* = 1.000 occupati residenti; *tarSZ* = 10.000 occupati residenti.

Dall'analisi cartografica si conferma che la soluzione prescelta permette di identificare *SLL*, anche di piccole dimensioni, che sono considerati da esperti del territorio come realtà e centri di riferimento consolidati, e ben si adegua all'evidenza empirica di un aumento della consistenza dei flussi di pendolarismo e del numero di connessioni tra comuni, soprattutto intorno ai grandi centri urbani del nord (evidenza che comporta l'espansione, rispetto alle configurazione del 2001, di alcuni *SLL*.

III.2. Lo Standard IPC

I dati brevettuali oggetto di analisi adottano, come criterio di classificazione, lo standard *IPC* (*International Patent Classification*, o, in italiano, *Classificazione Internazionale dei Brevetti*); questo è un sistema gerarchico di classificazione dei brevetti, ad oggi adottato in più 100 paesi al mondo, nato col proposito di uniformare i criteri di classificazione del contenuto dei brevetti.

Tale standard nasce quasi cinquanta anni fa, nell'ambito dell'*Accordo di Strasburgo* risalente al 1971, (trattato amministrato dall'*Organizzazione Mondiale della Proprietà Intellettuale OMPI*), nel quale si decise che i criteri con il quale riferirsi, in merito tecnologia specifica cui un brevetto appartiene, avrebbero dovuto adottare simboli indipendenti dalla lingua parlata dall'inventore o adottata dall'ente riconoscitore, e avrebbero dovuto garantire un facile inserimento in un opportuno "archivio" in base alla categoria tecnologica di riferimento.

La versione dello standard oggi in uso prevede che ogni simbolo di classificazione abbia una forma del tipo "A01B 1/00" (che, nel caso di specie, rappresenta gli *utensili manuali*); in esso si riscontrano le seguenti caratteristiche comuni:

- La prima lettera (nell'esempio precedente la *A*) rappresenta la "sezione", e deve necessariamente essere rappresentata da una delle prime otto lettere dell'alfabeto latino, da "A" ad "H" (le 8 macro-categorie tecnologiche cui le sezioni si riferiscono sono:
  - *A* "*Necessità Umane*";
  - *B* "*Esecuzione di operazioni, trasporto*";
  - *C* "*Chimica, Metallurgia*";
  - *D* "*Tessile, Carta*";
  - *E* "*Costruzioni Fisse*";
  - *F* "*Ingegneria Meccanica, Illuminazione, Riscaldamento, Armi*";
  - *G* "*Fisica*";
  - *H* "*Elettricità*";
- Tale lettera, è poi combinata con un numero arabo di due cifre, rappresentante la "classe" (la classe *A01* dell'esempio di prima rappresenta *Agricoltura; silvicoltura; zootecnia; cattura; pesca*).
- L'ultima lettera, anch'essa dell'alfabeto latino, costituisce la "sottoclasse" (la *sottoclasse A01B* rappresenta *Il suolo che lavora in agricoltura o silvicoltura; parti, dettagli o accessori di macchine o attrezzi agricoli, in generale*).

- La *sottoclasse* è poi seguita da un numero, di "gruppo", da una a tre cifre, un tratto obliquo e un ulteriore numero di almeno due cifre che rappresentano un "gruppo principale" o un "sottogruppo".

Un esaminatore di brevetti assegna simboli di classificazione alla domanda di brevetto o ad altro documento in conformità con le regole di classificazione e generalmente al livello più dettagliato applicabile al suo contenuto.

Ad oggi, nello standard *IPC* la tecnologia presenta circa 70.000 suddivisioni; la classificazione viene aggiornata regolarmente da un comitato di esperti, composto dai rappresentanti degli Stati contraenti l'*Accordo di Strasburgo* e osservatori di altre organizzazioni, come, ad esempio, l'*Ufficio europeo dei brevetti*.

III.3. La Nomenclatura delle Attività Economiche NACE

Da una prima disamina dei dati brevettuali presenti nel database, si evince che le *classi* (codice di riferimento *IPC* a 3 cifre) brevettuali per le quali si ha avuto, nel corso dei 15 anni di osservazioni, produzioni di brevetti non nulle sono esattamente 120, e rappresentative di tutte le 8 *sezioni* brevettuali dello standard *IPC*.

Nel corso della seguente trattazione il calcolo delle metriche sopra descritte sarà effettuato con due criteri di classificazione: il primo sarà il già citato standard *IPC* in virtù del quale saranno adottate 120 categorie di prodotti; il secondo criterio di raggruppamento, invece, sarà adottato alla luce degli esempi di calcolo di *Hidalgo & C*[3]*.* nella letteratura di riferimento, ed è rappresentata dalla suddivisone dei brevetti secondo la classificazione *NACE*, in virtù del quale il numero di categorie brevettuali saranno 26, un numero in grado di garantire un grado di risoluzione opportuno.

La scelta di replicare il calcolo con entrambi i criteri di classificazione permetterà di osservare gli effetti, sulle metriche, di una differente balcanizzazione del *dataset*. Si precisa che la scelta di rifiutare la banale adozione della divisione in 8 sezioni brevettuali è dettata dal fatto che, a parere dello scrivente, una così bassa granularità delle osservazioni non permetterebbe la piena osservazione dei fenomeni d'interesse e svilirebbe il valore stesso delle metriche calcolate.

La classificazione *NACE* (acronimo di *Nomenclature statistique des Activités économiques dans la Communauté Européenne*, ovvero *Nomenclatura statistica delle Attività economiche nella Comunità Europea*), è un sistema di classificazione, oggi ampiamente diffuso, sfruttato per sistematizzare ed uniformare le definizioni delle attività economiche e industriali negli Stati facenti parte dell'*Unione Europea*.

Esso fu introdotto per la prima volta dall'*Eurostat* nel 1970, e da allora ha subito diverse revisioni, raffinazioni e rimaneggiamenti. Per favorirne la definitiva e rapida diffusione fu cruciale, a tal proposito, la revisione del 2006 (la "rev.2", implementata nel 2007), dopo il quale la classificazione *NACE*, in realtà una derivazione dell'*ISIC* (*International Standard Industrial Classification*), ottenne la sua definitiva consacrazione tra i paesi membri dell'*UE* divenendo lo standard di riferimento la raccolta e la presentazione dei dati ad interesse statistico ed economico.

[3] Il quale adotta la disaggregazione dei brevetti in 30 categorie brevettuali, per come definito dal *National Bureau of Economic Reserch*.

La classificazione *NACE*, consolidata a livello comunitario, per essere meglio recepita dai vari paesi europei, è stata mutuata a livello nazionale, talvolta integrandola laddove ritenuto opportuno: in Italia il suddetto recepimento si è avuto con l'introduzione del codice *ATECO*, ovvero la classificazione italiana delle attività economiche produttive.

Tutto ciò che è oggetto di analisi statistica, infatti, per poter essere confrontabile, fruibile e largamente impiegabile richiede un'esauriente classificazione sistematica; tale classificazione deve rendere, infatti, possibile la divisone dell'universo in "gruppi", internamente largamente omogenei, garantendo:

- Copertura esaustiva dell'universo oggetto dell'indagine (ovvero nulla deve rimanere non classificato);
- Categorie mutuamente esclusive: ogni elemento deve essere classificato in una e una sola categoria;
- Principi metodologici che consentono un'allocazione coerente degli elementi nelle varie categorie della classificazione.

La *NACE*, e di conseguenza anche l'*ATECO*, per garantire il rispetto dei succitati principi è costituita da una struttura gerarchica che, da quanto specificato sui siti dell'*Eurostat* e dell'*Istat*, si presenta come nel seguito:

- le *sezioni*, il primo livello, le cui voci sono contraddistinte da un codice alfabetico;
- le *divisioni*, il secondo livello, le cui voci sono contraddistinte da un codice numerico a 2 cifre;
- i *gruppi*, il terzo livello, le cui voci sono contraddistinte da un codice numerico a 3 cifre;
- infine le *classi*, il quarto e ultimo livello, rappresentato con un codice numerico a 4 cifre;

A queste l'ATECO aggiunge a sua volta, per esigenze nazionali:

- le *categorie*, un quinto livello, contraddistinto da un codice numerico a 5 cifre;
- le *sottocategorie*, un sesto livello, contraddistinto da un codice numerico a 6 cifre.

Nel riferirsi ad un'attività, tramite codice *NACE*, con un livello di dettaglio maggiore del primo livello (ovvero almeno specificando la *divisone*) il codice che si riferisce al livello *sezioni* non viene tipicamente integrato: nell'identificativo sarà pertanto specificata la *divisione*, il *gruppo* e la *classe* relativa ad una specifica attività. A titolo d'esempio, l'attività "Fabbricazione di colle" è identificata dal codice *20.52*, dove *20* è il codice per la *divisione*, *20.5* il codice per il *gruppo* è *20.52* il codice per la

classe; nel codice identificativo, la *sezione* cui appartiene tale *classe* (nel caso di specie la "*C*") non appare.

Come ulteriore esempio chiarificatore, l'attività di "Fabbricazione di birra" è codificata, secondo *NACE*, come la classe *11.05* perché la divisione "Produzione di bevande" (codice *11*) non è suddivisa in più *gruppi*, mentre il *gruppo* "Produzione di bevande" (codice *11.0*) è suddiviso in *classi*, e quella afferente la "Produzione di birra" è la *11.05*.

Nella tabella seguente l'elenco delle "divisioni" cui appartengono i dati brevettuali presenti nel database oggetto di analisi, con a fianco la relativa descrizione:

Codice NACE	*Descrizione Attività*
10	Industrie Alimentari
11	Produzione di bevande
12	Industria del tabacco
13	Industrie tessili
14	Confezione di articoli di abbigliamento
15	Confezione di articoli in pelle e simili
16	Industria del legno e dei prodotti in sughero
17	Fabbricazione di carta e prodotti di carta
18	Stampa e riproduzione su supporti registrati
19	fabbricazione di coke e prodotti derivati dalla raffinazione del petrolio
20	Fabbricazione di prodotti chimici
21	Fabbricazione di prodotti farmaceutici di base e di preparati farmaceutici
22	Fabbricazione di articoli in gomma e materie plastiche
23	Fabbricazione di altri prodotti della lavorazione di minerali non metalliferi
24	Attività metallurgiche
25	Fabbricazione di prodotti in metallo, esclusi macchinari e attrezzature
26	Fabbricazione di computer e prodotti di elettronica e ottica
27	Fabbricazione di apparecchiature elettriche
28	Fabbricazione di macchinari e apparecchiature n.c.a.
29	Fabbricazione di autoveicoli , rimorchi e semirimorchi
30	Fabbricazione di altri mezzi di trasporto
31	Fabbricazione di mobili
32	Altre Industrie manifatturiere
42	Ingegneria civile
43	Lavori di costruzione specializzati
62	Programmazione, consulenza informatica e attività connesse

Tabella 3.1: Elenco dei codici *NACE* e relativa descrizione cui afferiscono i brevetti nel *dataset* oggetto di analisi

III.4. Caratteristiche dei dataset oggetto di analisi

Il primo passo per la costruzione delle metriche descritte dalla *Relatedness Theory*, in virtù di quanto in precedenza detto, è la trattazione del *dataset* fornito dal *Politecnico di Torino* per renderlo idoneo allo scopo: per far ciò occorre associare a ciascun brevetto, per mezzo di un'opportuna tabella di concordanza tra i codici *IPC* a 4 *digit* e i codici *NACE* a 2 *digit*, l'associato codice *NACE*. La succitata tabella di concordanza impiegata è quella scaricabile dal sito dell'*Eurostat* sotto il titolo: *Patent Statistics: Concordance IPC V8 – Nace rev.2*, realizzata ad Ottobre 2015[4].

L'applicazione della citata tabella di concordanza al *dataset* fornito fa emergere però alcune criticità nello stesso: secondo quanto descritto nella summenzionata pubblicazione dell'*Eurostat*, esistono alcune categorie merceologiche ricadenti sotto specifici codici *IPC* che, in virtù della loro intrinseca estrema eterogeneità non vengono "smistate" in alcun codice *NACE* ma rientrano invece nella categoria *co-IPC*. I codici *IPC* che ricadono in tali caratteristiche sono:

- *F16S* (elementi costruttivi in genere e, a loro volta, strutture costruite con tali elementi);
- *B29K* (a cui afferiscono le categorie *B29B*, *B29C* e *B29D* relative ai materiali per stampaggio, rinforzi, riempitivi e preformati);
- *B29L* (in cui ricadono alcuni particolari articoli della sottoclasse *B29C*);
- *C12R* (a cui fanno riferimento gli elementi delle sottoclassi, non già rimosse dallo standard, che vanno da *C12C* a *C12Q* e la *C12S*, tutte legate ai microorganismi);
- *-99Z* (tutti gli oggetti non diversamente previsti).

Da una prima analisi del database si riscontrano, dal 1999 al 2013, su 108.508 brevetti registrati in Italia, 479 brevetti ricadenti nelle categorie *IPC* prima citate nelle quantità riportate nella tabella seguente:

Codice IPC	*N. Brevetti*
B29K	6
B29L	2
C12P	72
C12Q	184
C12R	7
F16S	208
Totale	*479*

Tabella 3.2: Numero di brevetti afferenti ai rispettivi codici *IPC* non "smistati" in alcun codice *NACE*

[4] https://ec.europa.eu/eurostat/ramon/documents/IPC_NACE2_Version2_0_20150630.pdf

A questi si devono aggiungere, nel medesimo arco temporale, 13 brevetti che risultano registrati con codici desueti, per lo più aboliti con la revisione del 1979 o con quella del 1984 (nella tabella che segue le quantità dei brevetti registrati con tali codici e la relativa data di estinzione dello stesso).

Codice IPC	*Data Soppressione*	*N. Brevetti*
B01K	31/12/1979	1
B29F	31/12/1984	3
B29G	31/12/1984	1
B29J	31/12/1984	1
B65J	31/12/1979	2
C07M	31/12/2005	1
C12B	31/12/1979	2
C12K	31/12/1979	1
E01G	31/12/1979	1
Totale		*13*

Tabella 3.3: Numero di brevetti registrati nei rispettivi codici *IPC* e data di soppressione di quest'ultimi

Per gli scopi delle analisi oggetto del presente elaborato scegliamo di scartare, dai 108.508 brevetti registrati i 492 manifestanti le suddette frizioni con la classificazione *NACE*. Tale operazione di rimozione si può ritenere accettabile e non lesiva della validità dei risultati complessivamente ottenuti dall'elaborazione del database in quanto:

- i 13 brevetti (0,01% del totale) erroneamente indicati con codici *IPC* desueti, rappresentano, già da sé, dei dati errati inficianti il database e, in quanto tale, una loro rimozione può solo giovare, sotto il profilo della validità delle elaborazioni realizzate;
- i 479 (0,44% del totale) dati brevettuali ricadenti nella categoria *co-IPC*, data la loro eterogeneità tecnologica non sono classificabili secondo *NACE*, ma il loro numero, se rapportato al totale dei brevetti registrati e contenuti nel *database*, si può comunque ritenere esiguo;

III.4.1. Analisi e descrizione del dataset dei brevetti

Rimossi i 492 dati brevettuali citati il *database* si presenta contenente 108.016 brevetti afferenti a tutte e 8 le sezioni dello standard *IPC* (e più nello specifico ben 120 classi) e a 26 diversi codici *NACE*.

Realizzando una suddivisione merceologica dei brevetti, poiché le analisi che seguiranno saranno effettuate adottando come criterio di classificazione sia lo standard *IPC* che la classificazione *NACE*, nel seguito sarà realizzata una breve descrizione delle caratteristiche merceologiche del *dataset*, adottando, uno per volta, entrambi i criteri.

Codice NACE	*Descrizione*	*Quantità*	*Percentuale*
10	Industrie Alimentari	1.628	1,51%
11	Produzione di bevande	204	0,19%
12	Industria del tabacco	210	0,19%
13	Industrie tessili	622	0,58%
14	Confezione di articoli di abbigliamento	560	0,52%
15	Confezione di articoli in pelle e simili	1.181	1,09%
16	Industria del legno e dei prodotti in sughero	253	0,23%
17	Fabbricazione di carta e prodotti di carta	131	0,12%
18	Stampa e riproduzione su supporti registrati	303	0,28%
19	Fabbricazione di coke e prodotti derivati dalla raffinazione del petrolio	242	0,22%
20	Fabbricazione di prodotti chimici	8.046	7,45%
21	Fabbricazione di prodotti farmaceutici di base e di preparati farmaceutici	1.210	1,12%
22	Fabbricazione di articoli in gomma e materie plastiche	2.721	2,52%
23	Fabbricazione di altri prodotti della lavorazione di minerali non metalliferi	2.423	2,24%
24	Attività metallurgiche	735	0,68%
25	Fabbricazione di prodotti in metallo, esclusi macchinari e attrezzature	5.573	5,16%
26	Fabbricazione di computer e prodotti di elettronica e ottica	10.882	10,07%
27	Fabbricazione di apparecchiature elettriche	9.933	9,20%
28	Fabbricazione di macchinari e apparecchiature n.c.a.	32.901	30,46%
29	Fabbricazione di autoveicoli , rimorchi e semirimorchi	4.859	4,50%
30	Fabbricazione di altri mezzi di trasporto	2.842	2,63%
31	Fabbricazione di mobili	2.443	2,26%
32	Altre Industrie manifatturiere	13.379	12,39%
42	Ingegneria civile	413	0,38%
43	Lavori di costruzione specializzati	3.982	3,69%
62	Programmazione, consulenza informatica e attività connesse	340	0,31%
Totale		*108.016*	*100,00%*

Tabella 3.4: Numero di brevetti, e relativa percentuale sul totale, registrati nei rispettivi codici *NACE*

Come verificabile dal lettore, ad un solo codice *NACE*, il codice *28* "Fabbricazione di macchinari e apparecchiature n.c.a.", sono afferenti ben 32.901 brevetti (il 30,46% del totale). I rimanenti brevetti sono distribuiti assai poco omogeneamente tra i restanti codici, con una visibile preponderanza di altri 4 codici che, congiuntamente, contano per il 39,12% del totale; questi codici sono: *26* "Fabbricazione di

computer e prodotti di elettronica e ottica", *27* "Fabbricazione di apparecchiature elettriche", *20* "Fabbricazione di prodotti chimici", *32* "Altre Industrie manifatturiere".

La distribuzione dei brevetti tra le differenti categorie merceologiche secondo la classificazione *NACE* è quindi fortemente eterogenea: a fronte, infatti di un numero medio di brevetti per singolo codice pari 4.154, solo i primi 5 codici per numero di brevetti sono rappresentativi, congiuntamente, di ben 75.141 brevetti (ovvero il 69,56% del totale) mentre gli ultimi 5 codici contengono, insieme, l'esiguo numero di 1.040 brevetti (ovvero lo 0,96% del totale).

Tra i 5 codici *NACE* che rappresentano i numeri più esigui nel *dataset* troviamo: *17* "Fabbricazione di carta e prodotti di carta" (131 brevetti, ovvero lo 0,12% del totale), *11*"Produzione di bevande" (204 brevetti, ovvero lo 0,19% del totale), *12* "Industria del tabacco" (210 brevetti, ovvero lo 0,19% del totale), *19* "Fabbricazione di coke e prodotti derivati dalla raffinazione del petrolio" (242 brevetti, ovvero lo 0,22% del totale), *16* "Industria del legno e dei prodotti in sughero" (253 brevetti, lo 0,23% del totale).

Studiando le caratteristiche del *dataset*, adottando come criterio di classificazione i settori brevettuali dello standard *IPC* (la scelta di categorie a così altro livello è dettata dai fini descrittivi propri di questo paragrafo), si può constatare la non indifferente diversità, in termini di volumi, della produzione brevettuale tra i differenti settori; tali volumi, ripartiti come descritto, si possono più nel dettaglio osservare nella tabella che segue:

Codice	*Descrizione*	*Quantità*	*Percentuale*
A	Human Necessities	26.146	24,21%
B	Performing Operations, Transporting	31.250	28,93%
C	Chemistry, Metallurgy	7.055	6,53%
D	Textiles, Paper	3.180	2,94%
E	Fixed Construction	9.830	9,10%
F	Mechanical Engineering, Lighting, Heating, Weapons, Blasting	13.524	12,52%
G	Phisics	9.910	9,17%
H	Electricity	7.121	6,59%
Totale		*108.016*	*100,00%*

Tabella 3.5: Numero di brevetti registrati nei rispettivi codici *IPC*

Due degli otto settori, più in particolare il settore *A* "*Human Necessities*" e il settore *B* "*Performing Operations, Transporting*" rappresentano, congiuntamente, più del 50% della produzione totale di brevetti (con precisione, il 53,14%) contando, rispettivamente, per il 24,21% e 28,93%. Gli altri

settori presentano invece delle produzioni ben più modeste con primato di esiguità riconosciuto al settore *D* ("*Textiles and Paper*") il quale rappresenta solo il 2,94% della produzione brevettuale totale.

Inoltre, come è possibile notare confrontando le due tabelle, nonostante la minore balcanizzazione data dallo standard *IPC* a 1 *digit*, non si osserva un minor accentramento dei dati brevettuali all'interno delle singole categorie merceologiche: nonostante infatti un numero medio di brevetti, per singolo codice *IPC*, pari a 13.502, il codice *NACE* più numeroso presenta un numero di brevetti maggiore del codice *IPC* più numeroso nonostante il minor numero di categorie.

Per avere maggior chiarezza su come i brevetti si sono distribuiti tra i vari SLL italiani nei 15 anni di osservazione del fenomeno presentiamo una breve disamina dei volumi e dei trend brevettuali per i 10 SLL più prolifici, come desumibile dal *dataset* e come rappresentato nella tabella seguente.

Posizione	*Codice SLL*	*SLL*	*N. Brevetti*	*Percentuale*
1°	313	Milano	13.342	12,35%
2°	820	Bologna	5.752	5,33%
3°	106	Torino	5.240	4,85%
4°	1209	Roma	5.065	4,69%
5°	540	Padova	2.580	2,39%
6°	915	Firenze	2.175	2,01%
7°	315	Bergamo	1.936	1,79%
8°	321	Brescia	1.685	1,56%
9°	815	Modena	1.589	1,47%
10°	710	Genova	1.432	1,33%
Totale			*40.796*	*37,77%*

Tabella 3.6: Numero di brevetti registrati (e relativa percentuale) nei 10 *SLL* più prolifici

Utilizzando come criterio di ordinamento la quantità di brevetti registrati si può osservare un evidente fenomeno di accentramento dei volumi brevettuali intorno ai grandi centri urbani; tale evidenza si esaspera ulteriormente nel caso del *Sistema Locale del Lavoro* di *Milano*, il primo della classifica in termini di volumi brevettuali, nel quale, dal '99 al 2013, sono stati registrati ben 13.342 brevetti, ovvero ben il 12,65% del totale delle registrazioni.

Dopo *Milano*, gli *SLL* di *Bologna*, *Torino* e *Roma* sono quelli che presentano le più consistenti registrazioni (tra 5.000 e 6.000, ovvero intorno al 5% del totale): solo i primi 4 *SLL*, pertanto, rappresentano congiuntamente ben il 27,22% del totale.

Dopo i primi 4 summenzionati, gli altri 6 centri inclusi tra i primi 10 presentano ben più modesti volumi realizzativi, con cifre, in termini assoluti, tra 1.400 e 2.600 brevetti ovvero, in termini relativi, tra l'1,30%

e il 2,40% del totale. Si precisa infine che, dei 611 *SLL* in cui è ripartito il territorio italiano, in ben 574 è stato registrato almeno un brevetto nella finestra temporale di osservazione, in solo 37 *SLL* non è stato presentato alcun brevetto, dando prova di una non singolare vivacità e capillarità dell'attività brevettuale.

Infine, per completezza, avendo a disposizione anche le date di registrazione dei singoli brevetti, una tabella contenente i volumi di brevetti per ciascun settore brevettuale *IPC*, con risoluzione temporale annuale:

		Sezione								
		A	B	C	D	E	F	G	H	Totale
Anno	1999	1.802	2.742	606	324	544	911	635	586	8.150
	2000	1.735	2.284	605	287	529	856	671	563	7.530
	2001	1.905	2.305	587	265	570	862	682	553	7.729
	2002	1.866	2.418	491	269	674	887	702	491	7.798
	2003	2.008	2.404	457	292	702	1.003	626	489	7.981
	2004	1.890	2.440	501	211	768	914	733	435	7.892
	2005	1.916	2.307	465	241	802	978	725	485	7.919
	2006	2.268	2.643	513	229	900	1.248	900	483	9.184
	2007	2.171	2.388	444	205	944	1.120	804	566	8.642
	2008	1.567	1.749	411	171	714	853	567	437	6.469
	2009	1.079	1.376	348	118	535	647	457	365	4.925
	2010	1.434	1.477	388	152	593	820	541	416	5.821
	2011	1.469	1.597	389	144	523	794	558	437	5.911
	2012	1.423	1.532	442	136	524	865	658	403	5.983
	2013	1.613	1.588	408	136	508	766	651	412	6.082
	Totale	26.146	31.250	7.055	3.180	9.830	13.524	9.910	7.121	

Tabella 3.7: Suddivisione dei brevetti per sezione *IPC* con risoluzione temporale annuale

III.4.2. Analisi delle caratteristiche dei SLL

Secondo quando riferito dall'*Istituto Nazionale di Statistica* e pubblicato sul sito *Istat.it*, nel 2011, in occasione del 15° censimento della popolazione, erano presenti in Italia 611 Sistemi Locali del Lavoro. Essendo i SLL indipendenti da confini amministrativi di province e regioni si ha che 56 *SLL* (ben il 9,2%) si collocano a cavallo di più regioni (di cui 2 – i Sistemi di Voghera (PV) e Melfi (PT) a cavallo di ben 3 regioni) e ben più numerosi (185, il 30,3% del totale) sono quelli che si collocano a cavallo di più province.

In termini di popolazione residente, il numero medio di residenti per ciascun *SLL* è di poco più di 97 mila abitanti; i primi tre *Sistemi* sono quelli di *Milano*, *Roma* e *Napoli*, con rispettivamente, circa 3,7 milioni, 3,5 milioni e 2,5 milioni di abitanti. Solo questi tre *Sistemi Locali* raccolgono, congiuntamente, il 16,3% della popolazione residente nazionale; se a questi aggiungiamo anche il sistema di *Torino*, che il quarto *sistema locale* per popolazione residente (contando più di un milione di abitanti), tale percentuale sale fino al 19,2%.

Sul fronte opposto, i *sistemi locali* di piccole dimensioni si concentrano, come facilmente intuibile, prevalentemente nelle aree interne del Paese ed in particolare nelle regioni montane: il più piccolo *sistema locale* è quello di *Canazei* (*Provincia autonoma di Trento*), con appena 3.138 abitanti nel 2011, seguito a breve distanza dal sistema di *Valtournenche* (*Valle d'Aosta*), con poco meno di 3,5 mila abitanti, e quello di *Visso* (*Macerata*) con 3.542 abitanti.

Osservando l'estensione dei vari *sistemi* (i quali presentano una dimensione media di 495,2 km^2) si nota che il *sistema locale* di *Roma*, con oltre 3.800 km^2, è il più esteso, (e il solo comune di *Roma* contribuisce per oltre un terzo della superficie) seguito, nella graduatoria, dal sistema di *Bologna*, con poco più di 2.500 km^2, e da quello di *Torino* (con 2.467 km^2). In fondo alla graduatoria si collocano i sistemi locali isolani di *Capri* (10,5 km^2), *Forio* (21,6 km^2) e *Ischia* (25,0 km^2).

Sotto, invece, il profilo numerico è la *Sicilia* la regione con il maggior numero di *sistemi locali* (ben 71), seguita dalla *Lombardia* (51 SLL) e dalla *Toscana* (48 *SLL*); il *Molise* e la *Valle d'Aosta*, ambedue con 5 sistemi, sono le regioni con il minor numero di partizioni.

CAPITOLO IV

APPLICAZIONE DELLA RELATEDNESS THEORY

Nel corso del presente capitolo verranno discusse le metodologie applicative e le evidenze risultanti derivanti dall'applicazione, ai due dataset (quello dei dati brevettuali fornito dal *Politecnico di Torino* e quello relativo alla popolazione dei *Sistemi Locali del Lavoro* scaricabile dal sito *Istat.it*), già studiati e descritti nel precedente capitolo, le ipotesi, le metriche e gli strumenti d'analisi tipici della *Relatedness Theory*.

Data la vastità della letteratura in merito lo scrivente ha scelto di strutturare il presente capitolo in quattro parti: in ciascuna di esse verrà sviluppato e discusso il calcolo di specifiche metriche applicabili ai *dataset* in oggetto o saranno effettuate specifiche analisi utili per effettuare considerazioni ai sensi della letteratura specifica, già ai precedenti capitoli discussa, sperando di poter dare, con i risultati di tali analisi, un piccolo contributo alla conoscenza del mondo, come la *Relatedness Theory* ci insegna.

IV.1. Costruzione dello Spazio dei Prodotti

Come è stato già trattato al *Capitolo 2*, per come definiti e calcolati i singoli elementi costituenti lo *Spazio dei Prodotti* (o *Proximity Matrix*), in un mondo in cui, per evidenti ragioni di praticità, non è possibile dividere in comparti stagni e ben definiti le conoscenze necessarie per realizzare ogni prodotto dell'ingegno umano, sono rappresentativi della "somiglianza" tra le conoscenze necessarie per realizzare tutte le possibili combinazioni di prodotti.

Come decritto in letteratura, la costruzione dello Spazio dei Prodotti è subordinato a una serie di *step*; la costruzione della matrice e l'esecuzione dei vari *step* sono stati realizzati con *Excel*.

Rispetto alla nomenclatura classica delle metriche riportate in letteratura, data le peculiarità del "prodotto" esaminato si sostituirà alla parola "prodotto" usata in letteratura la parola "tecnologia": si parlerà quindi di "Spazio delle Tecnologie" o "Tecnhnology Complexity Index".

La costruzione delle varie metriche sarà realizzata, come descritto in precedenza, adottando due criteri di catalogazione dei prodotti: lo standard *IPC* e la classificazione *NACE*. Nel seguito si riporteranno prima i risultati delle analisi adottando lo standard *IPC* e poi quelli adottando la classificazione *NACE*.

IV.1.1. Costruzione dello Spazio delle Tecnologie secondo standard IPC

Il primo passo per pervenire alla *Proximity Matrix* e la costruzione della *RCA Matrix*: indicando con n_C il numero dei *SLL* caratterizzati da una produzione non nulla di brevetti nell'arco temporale di osservazione, e con n_P il numero di classi brevettuali registrate nel medesimo succitato arco temporale, la *RCA Matrix* si presenta come una matrice di dimensioni n_C x n_P, che nel caso specifico è una matrice caratterizzata da 582 righe e 120 colonne.

Il generico elemento RCA_{ij} della *RCA Matrix*, conterrà l'indice *Revealed Comparative Advantage* RCA_{ij} (per come definito da *Balassa* e come descritto al *Capitolo 2*), se nel territorio *i* in riga è stato registrato, nel quindicennio di osservazione, almeno un brevetto contraddistinto dal codice *j* in colonna, o lo "zero" altrimenti. Dalla *RCA Matrix* è facilmente derivabile la M_{CP} *Matrix*, una matrice anch'essa di dimensioni 582 x 120, nella quale il generico elemento M_{CPij} della matrice è pari a "1" se il corrispondente elemento, alla medesima riga *i* e alla medesima colonna *j*, della *RCA Matrix* è maggiore di *1*, ed è pari a "0" altrimenti; la M_{CP} *Matrix*, nel caso di specie, risponde quindi alla domanda "In quale *SLL* si brevetta cosa?".

La *RCA Matrix* e la M_{CP} *Matrix*, nonostante le preziose informazioni in esse contenute, data la mole di dati che contengono, rendono difficile realizzare un'immediata analisi comparativa tra le diverse *classi* brevettuali o tra i vari *SLL*.

Di più agevole uso, per tale scopo, sono le due metriche *Diversità* $k_{C,0}$ e *Ubiquità* $k_{P,0}$, facilmente calcolabili per, rispettivamente, ciascun territorio e ciascun prodotto, a partire dalla M_{CP} *Matrix* semplicemente sommando, rispettivamente, tutti gli elementi di una riga e tutti gli elementi di una colonna. Si ottengono così facendo due vettori, uno, quello della *Diversità*, di 582 elementi e l'altro, quello dell'*Ubiquità*, di 120.

Nelle 3 tabelle che seguono si rappresentano i primi 10 *SLL* per valore della Diversità $k_{C,0}$ della loro produzione brevettuale (si omettono gli ultimi 10 *SLL* poiché in fondo alla classifica vi stanno 32 *SLL* aventi *Diversità* unitaria) e le prime 10 classi *IPC* più ubique (ottenute ordinando le classi per decrescenti di $k_{P,0}$) e le ultime 10 classi IPC per valore di *Ubiquità*.

Numero	*Codice SLL*	*SLL*	$k_{C,0}$	$k_{P,1}$
1°	508	VERONA	48	89
2°	303	VARESE	46	92
3°	536	VENEZIA	46	91
4°	315	BERGAMO	45	89
5°	313	MILANO	44	73
6°	710	GENOVA	44	78
7°	1208	POMEZIA	44	78
8°	1209	ROMA	44	74
9°	1517	NAPOLI	44	90
10°	350	LECCO	43	91
Valori medi primi 10 SLL			45	85

Tabella 4.1: Primi 10 SLL per *Diversità* $k_{C,0}$ e corrispondente *Ubiquità media* $k_{P,1}$ dei loro prodotti

Numero	*Codice IPC*	*Attività*	$k_{P,0}$	$k_{C,1}$
1°	E04	Costruzioni	248	20
2°	A01	Agricoltura, Silvicoltura, Allevamento, Caccia, Pesca	232	19
3°	A61	Scienze mediche e veterinarie, Igiene	209	17
4°	A47	Mobili, Apparecchi domestici, Macinacaffè, Macinini per spezie, Aspiratori	173	21
5°	E06	Porte, Finestre, Persiane o Avvolgibili in generale, Scale	169	23
6°	A23	Prodotti alimentari e loro derivati non coperti da altre classi	167	22
7°	A63	Sport, Giochi, Divertimenti	163	22
8°	F24	Riscaldamento e Ventilazione	163	22
9°	B60	Veicoli in generale	162	19
10°	F03	Macchine e motori con liquidi, Motori per propulsione non già previsti	158	24
Valori Medi per le prime 10 classi IPC			184	21

Tabella 4.2: Primi 10 prodotti per *Ubiquità* $K_{P,0}$ e *Diversità* media dei territori che li producono

Numero	*Codice IPC*	*Attività*	$k_{P,0}$	$k_{C,1}$
111°	G11	Archiviazione delle informazioni	19	28
112°	B04	Apparecchi o macchine centrifughe per eseguire processi fisici o chimici	16	34
113°	C30	Crescita dei cristalli	15	35
114°	D07	Funi e Cavi diversi da quelli elettrici	15	34
115°	B06	Generare e trasmettere vibrazioni meccaniche in generale	9	35
116°	B82	Nanotecnologie	8	38
117°	C13	Industria dello zucchero	8	40
118°	B81	Tecnologia per microstrutture	6	30
119°	C06	Esplosivi, Fiammiferi	6	33
120°	G12	Strumentazione di dettaglio	2	41
Valori Medi per le ultime 10 classi IPC			10	35

Tabella 4.3: Ultime 10 classi per *Ubiquità* $K_{P,0}$ e *Diversità* media dei territori che li producono

Dalla *Tabella 4.1* sopra riportata possiamo già desumere delle importanti considerazioni: in merito ai *SLL*, a fronte di una *Diversità Media* dei *SLL* $k_{C,0\text{ - Medio}}$ uguale a "17" e un'*Ubiquità Media Pesata* $k_{P,1\text{ -Medio}}$ pari a "124", tenendo conto che gli ultimi 32 SLL (come già in precedenza scritto non presentati nel documento) hanno una *Diversità Media* $k_{C,0\text{ – Media Ultimi 32 SLL}}$ uguale a "1" e un'*Ubiquità Media Pesata* $k_{P,1\text{ – Media Ultimi 32 SLL}}$ pari a "185", si può osservare che al crescere della *Diversità* $k_{C,0}$ dei SLL si tende a realizzare brevetti appartenenti a classi meno ubique (aventi cioè un $k_{P,1}$ inferiore). Avendo, infatti, una *Diversità Media* dei primi 10 SLL più diversificati $k_{C,0\text{ – Media Primi 10 SLL}}$ uguale a "45" e un'*Ubiquità Media Pesata* $k_{P,1\text{ – Media Primi 10 SLL}}$ pari a "85" si ha:

$$k_{C,0\text{ – Media Primi 10 SLL}} = 45 > k_{C,0\text{ - Medio}} = 17 > k_{C,0\text{ – Media Ultimi 32 SLL}} = 1$$

$$k_{P,1\text{ – Media Primi 10 SLL}} = 85 < k_{P,1\text{ -Medio}} = 124 < k_{P,1\text{ – Media Ultimi 32 SLL}} = 185$$

Tale evidenza rappresenta una conferma empirica di quanto previsto in letteratura, dandoci prova del fatto che il processo di diversificazione di un *SLL* si realizza verso tecnologie più "rare", ovvero tecnologie che, richiedendo la contemporanea presenza di diverse altre conoscenze, sono associate a brevetti realizzati in un ristretto numero di *SLL*, "rarità" che si manifesta in una minore *Ubiquità Media* $k_{P,1}$ delle tecnologie dei *SLL* più diversificati e una maggiore *Ubiquità Media* delle tecnologie realizzate nei *SLL* meno diversificati.

Considerazioni analoghe si possono fare in merito alle classi brevettuali: a fronte di una *Ubiquità* media delle classi brevettuali $k_{P,0\text{ -Medio}}$ uguale a "*82*" e una *Diversità media pesata* dei territori $k_{C,1\text{ - Medio}}$ pari a

"*28*", come si osserva dalla *Tabelle 4.2*, l'*Ubiquità media* delle 10 classi brevettuali *IPC* più ubique $k_{P,0\,–Medio\ Prime\ 10\ Classi\ IPC}$ è pari a "*184*" a fronte di una *Diversità media pesata* dei territori in cui si realizzano tali tecnologie $k_{C,1\,–\,Media\ Prime\ 10\ Classi\ IPC}$ uguale a "*21*", e, come desumibile dalla *Tabella 4.3*, l'*Ubiquità media* delle 10 classi brevettuali *IPC* meno ubique $k_{P,0\,–Media\ Ultime\ 10\ Classi\ IPC}$ è pari a "10" a fronte di una Diversità media pesata dei territori $k_{C,1\,–\,Media\ Ultime\ 10\ Classi\ IPC}$ in cui si realizzano tali tecnologie uguale a "*35*".

Si ha quindi che:

$$k_{P,0\,–\,Media\ Prime\ 10\ Classi\ IPC} = 184 > k_{P,0\,-\,Medio} = 82 > k_{P,0\,–\,Media\ Ultime\ 10\ Classi\ IPC} = 10$$

$$k_{C,1\,–\,Media\ Prime\ 10\ Classi\ IPC} = 21 < k_{C,1\,-Medio} = 28 < k_{C,1\,–\,Media\ Ultime\ 10\ Classi\ IPC} = 35$$

Anche dallo studio delle classi brevettuali si confermano le considerazioni prima fatte a partire dallo studio dei *SLL*: brevetti appartenenti a classi tecnologiche più ubique sono tendenzialmente realizzati in territori meno diversificati e viceversa, anche qui dando conferma empirica di quanto contenuto nella letteratura di riferimento.

Le due metriche *Diversità* e *Ubiquità*, per come definite, pur fornendoci un'immediata serie d'importanti indicazioni circa le caratteristiche, rispettivamente, dei *SLL* e delle *classi* brevettuali possono dar luogo a una serie di errate interpretazioni come già in precedenza descritto. I problemi interpretativi insiti nel calcolo delle due metriche saranno, come vedremo, risolti con il calcolo delle metriche *ECI* e T*CI*, al successivo paragrafo.

Calcolata la *MCP Matrix* il passo verso il calcolo della *Prossimità* $\Phi_{PP'}$ è breve: sfruttando le formule in letteratura, per ciascun *classe* brevettuale, possiamo agevolmente calcolare il valore della *prossimità* verso tutte le altre *classi* brevettuali, valore rappresentante, date due classi *p* e *p'*, la "somiglianza" tra i requisiti di *capacità* richiesti per realizzare brevetti appartenenti alle due classi stesse.

La rappresentazione di tutti le possibili combinazioni di prossimità tra prodotti dà luogo a una matrice di dimensioni n_P x n_P (120 x 120 nel nostro caso), quadrata, simmetrica e con la diagonale composta unicamente da "1" in quanto non possono che essere algebricamente identiche le capacità richieste per realizzare due prodotti identici.

Tale matrice, battezzata anche come *Proximity Matrix*, è più nota in letteratura (quando rappresentata mediante grafo) come *Spazio delle Tecnologie*. Nell'"Allegato B: *Proximity Matrix*", nell' *Immagine B.1.* si riporta la *Proximity Matrix* delle classi brevettuali a 3 *digit* secondo standard *IPC*, in Italia, nel quindicennio 1999 – 2013.

Come visto negli esempi d'uso precedenti, lo *Spazio dei Prodotti*, combinato alle metriche *Diversità* e *Ubiquità*, è in grado di darci informazioni straordinarie circa lo stato delle *conoscenze* produttive necessarie per realizzare un prodotto (in questo caso un brevetto) e circa la loro distribuzione spaziale.

IV.1.2. Costruzione dello Spazio delle Tecnologie adottando la Classificazione NACE

Il già calcolato *Spazio delle Tecnologie* è ovviamente strettamente dipendente dai criteri adottati per classificare i differenti prodotti. Risulta quindi interessante ripercorrere i precedenti calcoli e su questi fare nuove considerazioni adottando, come criterio di classificazione dei brevetti la classificazione *NACE.*

Soprassedendo questa volta alla descrizione dei vari passaggi, in quanto questi sarebbero sostanzialmente identici a quelli sopra descritti, si sottolinea, come unica differenza, che n_P è pari in questo caso a *26*.

Anche con la classificazione dei brevetti secondo *NACE* perveniamo ai due vettori *Diversità*, di 582 elementi, e *Ubiquità*, di 26, e, seguendo l'approccio già adottato con la precedente classificazione si riportano nelle tabelle che seguono, rispettivamente, i primi 10 *SLL* ottenuti ordinando per valori decrescenti di $k_{C,0}$ il vettore *Diversità* (si omette anche qui la classifica degli ultimi 10 *SLL* poiché "in fondo" al vettore *Diversità* vi sono 38 *SLL* con *Diversità* unitaria), e, dato il più esiguo numero di divisioni brevettuali, la classifica delle divisioni *NACE* ottenuta ordinando per valori decrescenti di $k_{P,0}$ il vettore *Ubiquità*.

Numero	*Codice SLL*	*SLL*	$k_{C,0}$	$k_{P,1}$
1°	1108	ANCONA	15	129
2°	422	ROVERETO	13	119
3°	536	VENEZIA	13	138
4°	909	MONTECATINI-TERME	13	126
5°	1009	PERUGIA	13	123
6°	1102	FANO	13	129
7°	1301	AVEZZANO	13	144
8°	1914	PALERMO	13	131
9°	120	SAVIGLIANO	12	130
10°	131	BIELLA	12	133
Valori Medi Primi 10 SLL			13	130

Tabella 4.4: Primi 10 SLL per *Diversità* $k_{C,0}$ e *Ubiquità* media $k_{P,1}$ dei prodotti che in essi si realizzano

Come già fatto con la classificazione dei brevetti secondo standard *IPC* in merito ai *SLL*, con la classificazione *NACE*, tenendo conto che:

- la *Diversità Media* dei *SLL* $k_{C,0\text{ - }Medio}$ è uguale a "6" e l'*Ubiquità Media Pesata* $k_{P,1\text{ -}Medio}$ è pari a "170",

- come già accennato, "in fondo" al vettore Diversità vi stanno 38 *SLL* con *Diversità media* $k_{C,0\ –\ Media\ Ultimi\ 38\ SLL}$ pari a "1" e *Ubiquità Media pesata* dei loro prodotti $k_{P,1\ –\ Media\ Ultimi\ 38\ SLL}$ pari a "*211*",
- come desumibile dalla *Tabella 4.4* la *Diversità media* dei primi 10 *SLL* più diversificati $k_{C,0\ –\ Media\ Primi\ 10\ SLL}$ è uguale a "15" e l'*Ubiquità Media pesata* dei loro prodotti $k_{P,1\ –\ Media\ Primi\ 10\ SLL}$ è pari a "*130*",

si può anche qui asserire che:

$$k_{C,0\ –\ Media\ Primi\ 10\ SLL} = 15 > k_{C,0\ -\ Medio} = 6 > k_{C,0\ –\ Media\ Ultimi\ 38\ SLL} = 1$$

$$k_{P,1\ –\ Media\ Primi\ 10\ SLL} = 130 < k_{P,1\ -Medio} = 170 < k_{P,1\ –\ Media\ Ultimi\ 38\ SLL} = 211$$

Anche qui, si ha evidenza empirica di quanto previsto in letteratura, dimostrando che il processo di diversificazione di un *SLL* si realizza verso tecnologie meno ubique come dimostra la crescente *Ubiquità Media* delle divisioni brevettuali al diminuire della *Diversità* dei *SLL*.

Analoghe considerazioni a quelle già fatte si possono fare anche per le divisioni brevettuali: osservando la *Tabella 4.5* e non riportando, data l'eseguo numero di divisioni brevettuali coinvolte, alcuna media, salta immediatamente all'occhio osservando la colonna $k_{P,0}$ e $k_{C,1}$ che, pur oscillando $k_{C,1}$ tra 6 e 9, la moda dello stesso nella prima metà della classifica (dal 1° al 13° valore) è pari a "7" mentre la moda nella seconda metà della classifica (dal 14° al 26° valore) è pari a "8".

Anche qui come in precedenza, si può quindi asserire che la diversificazione non è un processo scontato e il suo verificarsi in pochi, specifici luoghi si riflette in una minore *Ubiquità media* dei prodotti realizzati nei *SLL* più diversificati.

Numero	*Codice NACE*	*Attività*	$k_{P,0}$	$k_{C,1}$
1°	32	Altre Industrie manifatturiere	265	6
2°	43	Lavori di costruzione specializzati	255	7
3°	28	Fabbricazione di macchinari e apparecchiature n.c.a.	240	6
4°	25	Fabbricazione di prodotti in metallo, esclusi macchinari e attrezzature	203	7
5°	27	Fabbricazione di apparecchiature elettriche	179	7
6°	10	Industrie Alimentari	170	7
7°	20	Fabbricazione di prodotti chimici	167	7
8°	30	Fabbricazione di altri mezzi di trasporto	165	7
9°	29	Fabbricazione di autoveicoli, rimorchi e semirimorchi	159	6

10°	26	Fabbricazione di computer e prodotti di elettronica e ottica	156	6
11°	23	Fabbricazione di altri prodotti della lavorazione di minerali non metalliferi	142	7
12°	31	Fabbricazione di mobili	131	8
13°	22	Fabbricazione di articoli in gomma e materie plastiche	126	7
14°	42	Ingegneria civile	116	8
15°	14	Confezione di articoli di abbigliamento	105	8
16°	62	Programmazione, consulenza informatica e attività connesse	92	9
17°	15	Confezione di articoli in pelle e simili	91	8
18°	18	Stampa e riproduzione su supporti registrati	80	8
19°	11	Produzione di bevande	78	8
20°	21	Fabbricazione di prodotti farmaceutici di base e di preparati farmaceutici	74	8
21°	24	Attività metallurgiche	72	9
22°	13	Industrie tessili	71	8
23°	16	Industria del legno e dei prodotti in sughero	59	8
24°	12	Industria del tabacco	50	9
25°	19	fabbricazione di coke e prodotti derivati dalla raffinazione del petrolio	45	9
26°	17	Fabbricazione di carta e prodotti di carta	39	9

Tabella 4.5: Classifica delle divisioni brevettuali *NACE* ordinate per valori decrescenti di $k_{P,0}$

Passando per la M_{CP} *Matrix* e calcolando successivamente i vari valori delle *Prossimità* $\Phi_{PP'}$, possiamo rappresentare tutte le possibili combinazioni di *Prossimità* tra brevetti tramite una matrice quadrata simmetrica 26 x 26, che rappresenta lo *Proximity Matrix* delle classi brevettuali secondo *NACE*, in Italia, nel quindicennio 1999 – 2013. Tale *Proximity Matrix* è raffigurata nell' *Immagine B.2.* nell'"Allegato B: *Proximity Matrix*".

La rappresentazione della *Proximity Matrix* mediante grafo costituisce lo *Spazio delle Tecnologie*; lo *Spazio delle Tecnologie* per le divisioni brevettuali *NACE* è raffigurata nell'*Immagine 4.1.* che segue.

Nel grafo nell'*Immagine 4.1.*, costruito grazie al software *R* e con l'algoritmo *Minimum Spanning Tree*, la dimensione di ciascun nodo è proporzionale al corrispondente valore del *Tecnhology Complexity Index TCI* (calcolati nei paragrafi che seguono) della divisone *NACE* a esso associata, lo spessore delle linee tra i nodi costituenti la "rete" è proporzionale alla *prossimità* tra le varie divisoni brevettuali, e, infine, il colore di ciascun nodo è indicativo della "Sezione NACE" cui è a associata la corrispondente divisone (indicata dentro al nodo).

Infine, per ragioni legate alla bontà di rappresentazione del grafo, si è scelto di eliminare dalla rappresentazione le "linee" associate a una *Prossimità* inferiore a "0,20"; tale valore è stato scelto per due motivazioni:

- Una *prossimità* di "0,20" rappresenta il minimo valore di *Prossimità* tale per cui tutti i nodi risultino legati ad almeno un altro nodo;
- È un valore di *prossimità* abbastanza modesto e tale per cui si possono ritenere modeste anche le *prossimità* tra divisioni inferiori a tale valore.

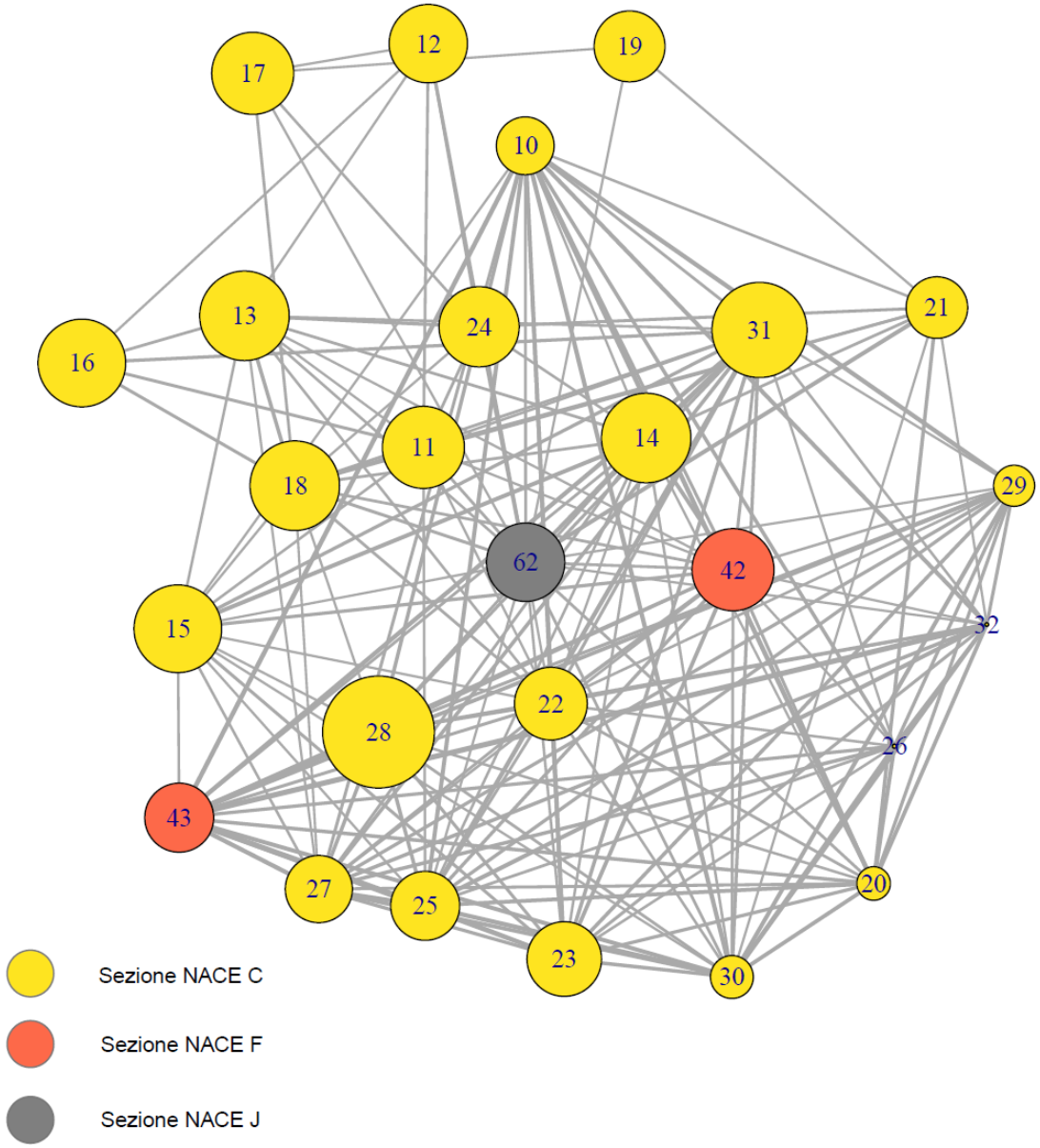

Immagine 4.1.: Rappresentazione dello *Spazio delle Tecnologie* costruito adottando la *Classificazione NACE* e utilizzando i volumi brevettuali, in Italia, nel quindicennio 1999-2013.

IV.2. Calcolo dei vettori ECI, TCI e OV e relative considerazioni

Come descritto nella sezione relativa alla letteratura scientifica in merito, l'impego delle metriche *Ubiquità* e *Diversità* presentano le controindicazioni d'uso già in precedenza descritte. Allo stesso modo lo *Spazio delle Tecnologie*, per quanto ricco delle preziose informazioni circa il grado di "somiglianza" tra le conoscenze produttive richieste per realizzare tutte le possibili combinazioni di categorie brevettuali, non è però in grado di darci informazioni circa la "vicinanza" di un territorio, dato il suo corrente mix tecnologico, alla produzione di categorie brevettuali che non già produce.

Per colmare i succitati limiti, nel presente paragrafo saranno calcolate le metriche *ECI*, *TCI* e *OV* seguendo le indicazioni e la metodologia specificata in letteratura; analogamente a quanto fatto al precedente paragrafo, tali metriche saranno calcolate adottando entrambi i criteri di catalogazione introdotti, lo standard *IPC* e la classificazione *NACE*.

Il processo di mutua correzione tra *Diversità* e *Ubiquità* è un processo ricorsivo come descritto al *Paragrafo 2.2.2.*; prendendo inizialmente in considerazione, per semplicità, la sola metrica *Diversità* tale processo ricorsivo alla *n-esima* iterazione arriva a convergenza quando si verifica la condizione $k_{C,N} = k_{C,N-2} = 1$. Sotto tale condizioni è possibile calcolare la matrice quadrata M_{CC}, di dimensioni n_C x n_C (di cui $k_{C,N}$ è l'autovettore associato al più elevato autovalore ma che, essendo un vettore di soli "uno", non ha contenuto informativo), il cui autovettore associato al secondo più elevato autovalore rappresenta il vettore *Diversità* "corretto" che cattura la più grande varianza del sistema.

Da tale vettore, come visto in letteratura, possiamo calcolare la metrica *Economic Complexity Index ECI*, metrica espressa su una scala lineare a intervallo, i cui valori, per singolo territorio, sono privi di significato in termini assoluti ma che ci permettono invece di costruire un ordinamento tra territori in termini di "complessità economica", ovvero in termini di *Diversità* del territorio stesso, tenendo però anche conto dell'*Ubiquità* dei diversi prodotti realizzati dal territorio e della *Diversità* degli altri territori che producono tali prodotti.

Replicando tali considerazioni per l'*Ubiquità*, è parimenti possibile, dalla M_{CP} *Matrix*, calcolare una matrice quadrata (battezzata M_{PP}) di dimensioni n_P x n_P, da cui ricavare la metrica *Technology Complexity Index TCI*, anch'essa espressa su una scala lineare d'intervallo, grazie al quale si può costruire un ordinamento tra tecnologie in termini di "complessità" delle stesse, ovvero in termini di *Ubiquità* della tecnologia, tenendo conto della *Diversità* dei territori in cui tale tecnologia si realizza, e dell'*Ubiquità* delle tecnologie realizzate in tali territori.

Infine, partendo dalle già calcolate matrici M_{CP} *Matrix* e *Spazio dei Tecnologie* si è già discusso come sia possibile calcolare la metrica "distanza" indicativa della distanza di una tecnologia *p* dal mix tecnologico esistente di un territorio *c*, metrica che, calcolata rispetto a ogni tecnologia e per ogni territorio, ci permette di costruire una matrice, di dimensioni n_C x n_P, contenete tutte le "distanze" da cui, a suo volta derivare, note le *distanze* e note le *complessità TCI* delle tecnologie raggiungibili, una misura delle *Opportunità* insite nel mix tecnologico corrente di un territorio, data la sua posizione nello Spazio delle Tecnologie.

Tale misura, battezzata in letteratura *Opportunity Value OV*, è, come *ECI* e *TCI* espressa in una scala lineare di intervallo, e, per quanto priva di significato in terminati assoluti, permette, invece, di costruire un ordinamento tra i territori, sotto il profilo delle *Opportunità* future degli stessi.

IV.2.1. Calcolo di ECI, PCI e OV adottando lo standard IPC

La M_{CP} già calcolata adottando lo standard *IPC* è una matrice di dimensioni *582 x 120*. Come visto al *Capitolo 2*, da questa sono facilmente calcolabili sia la matrice M_{CC} che la Matrice M_{PP}, le quali, presentano dimensioni, rispettivamente, *582 x 582* e *120 x 120*.

Il calcolo di *ECI* e *TCI* è stato realizzato tramite il software di analisi statistiche *R*, grazie al quale, importando le matrici M_{CC} e M_{PP} sono di facile calcolo i vettori *ECI* e *TCI*. Nel seguito i due *script* adottati per il calcolo delle due metriche.

Partendo dalla matrice M_{CC} importata in *R*, in formato *csv*, tramite il file "Mcc.csv" otteniamo, tramite il seguente *script*, il vettore *ECI* dentro il file "ECI.csv":

```
Mcc <- read.csv("Mcc.csv", header = FALSE, sep = ";")
        Mcc <- as.matrix(Mcc)
        ev <- eigen(Mcc)
                autovalori_Mcc <- ev$values
                autovettori_Mcc <- ev$vectors
                        secondo_autovalori_Mcc <- c(autovalori_Mcc[2])
                        secondo_autovettore_Mcc <- autovettori_Mcc[,2]
                media_secondo_autovettore_Mcc <- mean(secondo_autovettore_Mcc)
                sd_secondo_autovettore_Mcc <- sd(secondo_autovettore_Mcc)
ECI <- (secondo_autovettore_Mcc-media_secondo_autovettore_Mcc)/sd_secondo_autovettore_Mcc
write.table(ECI, file = "Eci.csv", sep = ",")
```

Partendo, invece, dalla matrice M_{PP} importata in *R*, in formato *csv*, tramite il file "Mpp.csv" otteniamo, tramite il seguente *script*, il vettore *TCI* dentro il file "TCI.csv":

```
Mpp <- read.csv("Mpp.csv", header=FALSE, sep = ";")
```

```
Mpp <- as.matrix(Mpp)
        Mpp <- as.matrix(Mpp)
        ev <- eigen(Mpp)
                autovalori_Mpp <- ev$values
                autovettori_Mpp <- ev$vectors
                        secondo_autovalore_Mpp <- c(autovalori_Mpp[2])
                        secondo_autovettore_Mpp <- c(autovettori_Mpp[,2])
        media_secondo_autovettore_Mpp <- mean(secondo_autovettore_Mpp)
        sd_secondo_autovettore_Mpp <- sd(secondo_autovettore_Mpp)
TCI <- (secondo_autovettore_Mpp- media_secondo_autovettore_Mpp)/sd_secondo_autovettore_Mpp
write.table(TCI, "TCI.csv", sep = ",")
```

Il software *R* ci restituisce due vettori *ECI* e *TCI*, rispettivamente di 582 e 120 elementi, dei quali vengono, di ciascuno, riportati nel seguito solo i due estremi superiore e inferiore, ottenuti dopo aver ordinato gli elementi dei vettori in ordine decrescente, di modo da evidenziare, sia dei territori che delle tecnologie, sia quelli più complessi che quelli meno complessi.

Si noti che nelle tabelle che seguono, in rifermento ai vettori *TCI* presentati nel presente paragrafo secondo standard *IPC* e nel successivo secondo classificazione *NACE*, saranno anche indicate le corrispettive posizioni nell'ordinamento ottenuto con l'altro criterio di classificazione di modo da poter così realizzare un confronto tra i due criteri al *Paragrafo 4.2.3.*

Inoltre, per entrambi i vettori saranno specificate le corrispondenti *Diversità* $k_{C,0}$ dei *SLL* e *Ubiquità Medie* $k_{P,1}$ delle loro tecnologie per le tabelle contenenti gli *ECI,* e le *Ubiquità* $k_{P,0}$ delle tecnologie e l'Ubiquità media $k_{C,1}$ dei territori che li producono per le tecnologie contenute per le tabelle contenenti i valori *TCI*. Si presentano, nelle due tabelle che seguono, i primi 10 *SLL* e gli ultimi 10 *SLL* ottenuti ordinando il vettore *ECI*, costruito adottando lo standard *IPC*, per valori decrescenti dello stesso:

Posizione$_{IPC}$	*Codice SLL*	*SLL*	*ECI*	*Diversità* $k_{C,0}$	*Ubiquità Media* $k_{P,1}$	*Posizione*$_{NACE}$
1°	301	BUSTO ARSIZIO	4,40	41	74	12°
2°	337	VIGEVANO	3,62	32	78	21°
3°	111	NOVARA	3,52	38	77	109°
4°	948	PRATO	3,51	28	74	59°
5°	822	IMOLA	3,48	32	79	146°
6°	510	ARZIGNANO	3,24	38	80	13°
7°	322	CHIARI	3,17	41	86	19°
8°	517	VICENZA	3,17	36	84	201°
9°	536	VENEZIA	3,11	46	91	99°
10°	512	BASS. DEL GRAPPA	3,09	39	89	42°

Tabella 4.6: Classifica dei primi 10 *SLL* per valore della metrica *ECI* adottando lo standard *IPC*

Posizione$_{IPC}$	*Codice SLL*	*SLL*	*ECI*	*Diversità $k_{C,0}$*	*Ubiquità Media $k_{P,1}$*	*Posizione$_{NACE}$*
573°	1502	MONDRAGONE	-1,37	13	142	384°
574°	1637	MAGLIE	-1,39	15	150	435°
575°	1003	CASTIGLIONE DEL LAGO	-1,50	11	159	561°
576°	702	IMPERIA	-1,53	20	136	563°
577°	1957	GIARRE	-1,53	17	141	417°
578°	1514	CASTEL. DI STABIA	-1,63	19	136	574°
579°	1006	GUALDO TADINO	-1,65	14	152	265°
580°	1315	GUARDIAGRELE	-1,76	19	140	418°
581°	1903	MARSALA	-1,79	19	130	582°
582°	918	CECINA	-1,81	21	136	528°

Tabella 4.7: Classifica degli ultimi 10 *SLL* per valore della metrica *ECI* adottando lo standard *IPC*

È opportuno ribadire che la metrica *Economic Complexity Index ECI*, per come calcolata, non ha alcun significato in termini assoluti. Per un *SLL*, infatti, un valore di *ECI* pari a "+4", altro non significa che a quello specifico *SLL* è associato all'interno dell'autovettore $\boldsymbol{K}$ che calcoliamo a valle del processo iterativo di correzione della diversità del *SLL* stesso, un valore la cui differenza rispetto alla media dei valori dell'autovettore stesso è pari a 4 volte la deviazione standard dei valori dell'autovettore.

Pur non contenendo alcun significato specifico in termini assoluti, grazie a tale metrica possiamo però creare un ordinamento tra *SLL*, dal più economicamente complesso a quello meno economicamente complesso. Dalle *Tabelle 4.6* e *4.7* si può osservare, a riprova di quanto suggerito dalla letteratura, che *SLL* maggiormente complessi tendono ad essere più diversificati e tendono a realizzare tecnologie meno ubique. Adottando lo standard *IPC* come criterio di classificazione dei brevetti i tre *SLL* tecnologicamente più complessi sono "Busto Arsizio", "Vigevano" e "Novara". Dalla M_{CP} *Matrix*, si nota come in tali 3 *SLL* si brevettano la maggior parte delle tecnologie che si trovano nella classifica delle prime 10 più complesse classi brevettuali *IPC*; più in particolare (a margine della classe si riporta la posizione nella classifica del *Technology Complexity Index* della *Tabella 4.8*):

- a *Busto Arsizio* si brevettano le classi B24 (1°), B26(2°), D05 (7°), G03(9°), D03(10°);
- a *Vigevano* si brevettano le classi B24 (1°), B26 (2°), F26 (3°), B08 (4°), B30 (5°), B32 (6°), D05 (7°), G03 (9°), D03 (10°);
- a *Novara* si brevettano le classi B24 (1°), F26 (3°), B08 (4°), B30 (5°), B32 (6°), D05 (7°), B02 (8°), D03 (10°).

Nei primi 3 *SLL* si brevettano quindi la maggior parte delle tecnologie che si trovano tra le prime 10 classi *IPC* e, ancor più nel dettaglio, Busto Arsizio vanta ad esempio un *Reaveled Comperative Advantge RCA* pari a "+6" nella classe brevettuale *D05* ("Cucito; Ricamo; Tufting" - la 7° maggiormente complessa) e un *RCA* pari a "+4,56" nella classe *D03* ("Tessitura" – la 10° maggiormente complessa)

Nelle due tabelle che seguono, invece, le prime 10 classi brevettuali e le ultime 10 classi ottenute ordinando il vettore *TCI*, costruito adottando lo standard *IPC*, per valori decrescenti dello stesso:

Posizione	*Codice IPC*	*Descrizione*	*TCI*	$k_{P,0}$	$k_{C,1}$
1°	B24	Molatura, Lucidatura	1,28	80	29
2°	B26	Utensili da taglio a mano, Taglio, Rescissione	1,08	91	28
3°	F26	Essiccazione	0,94	49	32
4°	B08	Pulizia	0,90	90	28
5°	B30	Presse	0,88	74	30
6°	B32	Prodotti stratificati	0,86	99	30
7°	D05	*Cucire, Ricamo, "Tufting"*	0,83	47	32
8°	B02	Schiacciamento, Polverizzazione, Disintegrazione, Macinazione	0,77	61	32
9°	G03	Fotografia, Cinematografia, Elettrografia, Olografia	0,75	54	32
10°	D03	Tessitura	0,71	39	30

Tabella 4.8: Classifica delle prime 10 classi brevettuali *IPC* per valore della metrica *TCI*

Posizione	*Codice IPC*	*Descrizione*	*TCI*	$k_{P,0}$	$k_{C,1}$
111°	G06	Informatica, Calcolare, Conteggio	-1,29	126	23
112°	H04	Tecniche di comunicazione elettrica	-1,30	119	23
113°	G01	Misurare, Testare	-1,57	126	22
114°	F16	Elementi di ingegneria, Produzione e mantenimento di macchine o impianti, Isolamento termico	-1,57	152	21
115°	A47	Mobilia, Articoli domestici, Macinacaffè, Macinini per spezie, Aspiratori	-1,59	173	21
116°	A23	Prodotti alimentari o per il loro trattamento	-1,81	167	22
117°	B60	Veicoli in generale	-2,97	162	19
118°	E04	Costruzione	-3,43	248	20
119°	A01	Agricoltura, Silvicoltura, Allevamento, Caccia, Intrappolamento, Pesca	-3,52	232	19
120°	A61	Scienze mediche e veterinarie, Igiene	-6,25	209	17

Tabella 4.9: Classifica delle ultime 10 classi brevettuali *IPC* per valore della metrica *TCI*

Anche qui, a conferma di quanto previsto in letteratura, dalle *Tabelle 4.8* e *4.9* si può notare come tecnologie più complesse tendono ad essere meno ubique di quelle meno tecnologicamente complesse (un più basso valore di $k_{P,0}$) e gli *SLL* nella quale si brevettano classi *IPC* più complesse tendono ad avere una *Diversità media* più elevata (un più elevato valore di $k_{C,1}$). Quanto fin qui detto si può anche osservare dai due *Grafici 4.1* e *4.2*.

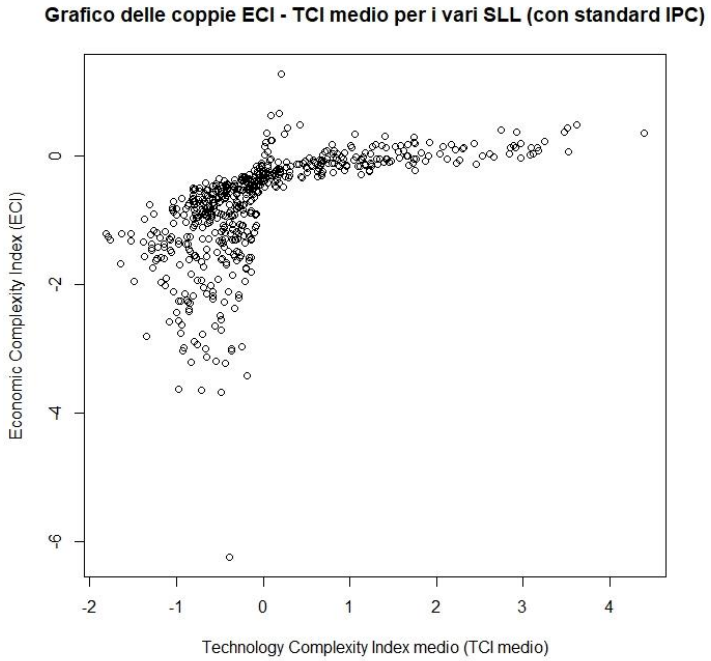

Grafico 4.1: Grafico delle coppie *ECI – TCI medio* per i vari SLL (con *Standard IPC*)

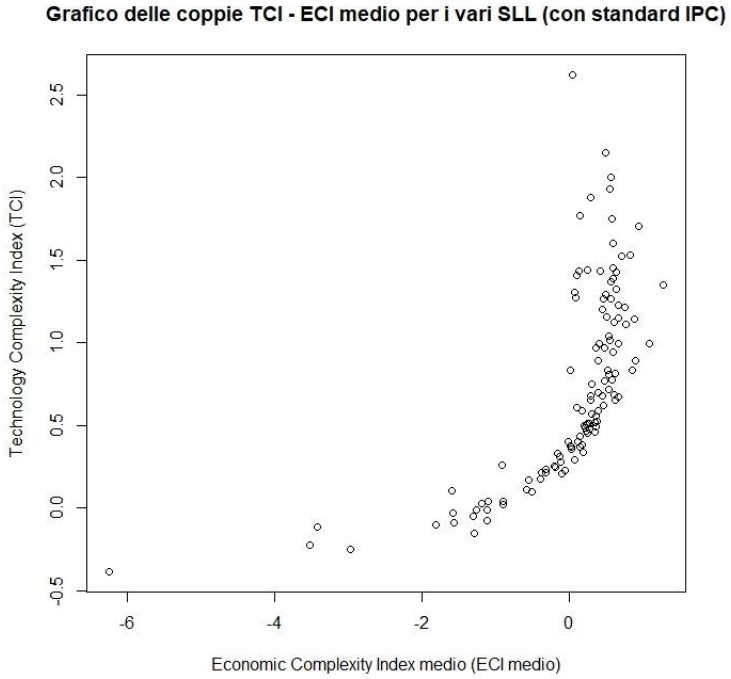

Grafico 4.2: Grafico delle coppie *TCI – ECI medio* per i vari SLL (con *Standard IPC*)

Nelle due tabelle che seguono, infine, i due estremi del vettore *Opportunity Value OV*, così ottenuti dopo aver riordinato il vettore per valori decrescenti della variabile stessa.

Posizione	*Codice SLL*	*SLL*	*Opportunity Value*	*Posizione$_{ECI}$*
1°	919	LIVORNO	4,70	214°
2°	1309	TERAMO	4,32	447°
3°	1612	BARI	4,17	218°
4°	1009	PERUGIA	4,05	151°
5°	1604	FOGGIA	3,63	570°
6°	121	ASTI	3,54	361°
7°	1005	FOLIGNO	3,54	328°
8°	1914	PALERMO	2,93	165°
9°	2016	CAGLIARI	2,89	138°
10°	1518	NOLA	2,83	572°

Tabella 4.10: Classifica dei primi 10 *SLL* per valore della metrica *OV* adottando lo standard *IPC*

Posizione	*Codice SLL*	*SLL*	*Opportunity Value*	*Posizione$_{ECI}$*
573°	838	RIMINI	-1,70	34°
574°	1202	CIVITA CASTELLANA	-1,71	66°
575°	131	BIELLA	-1,96	15°
576°	510	ARZIGNANO	-2,09	6°
577°	304	COMO	-2,19	13°
578°	315	BERGAMO	-2,32	17°
579°	822	IMOLA	-2,39	5°
580°	803	PIACENZA	-2,88	25°
581°	337	VIGEVANO	-3,15	2°
582°	301	BUSTO ARSIZIO	-3,32	1°

Tabella 4.11: Classifica degli ultimi 10 *SLL* per valore della metrica *OV* adottando lo standard *IPC*

Come si può desumere dalle *Tabelle 4.10* e *4.11* i *SLL* caratterizzati da un maggior *Opportunity Value* sono quelli caratterizzati da un minor valore dell'*ECI* (addirittura il *SLL* che presenta il più basso *Opportunity Value* è il *SLL* di "Busto Arsizio" che è il *Sistema Locale* più tecnologicamente complesso.

Una spiegazione di tale fenomeno si può trovare nella formula stessa dell'*Opportunity Value*: per come tale metrica è infatti definita in letteratura il termine "$(1 - M_{CP})$" fa sì che il contributo, nella sommatoria, alla metrica OV_C per ciascun *SLL* sia nullo quando il termine M_{CP} è pari a "1", ovvero quando in quel *SLL* quel prodotto si realizzi già.

Per tale motivo, essendo i *SLL* più tecnologicamente complessi anche, come abbiamo già empiricamente evidenziato, tendenzialmente più diversificati (presentando quindi un maggior numero di tecnologie *p* per il quale il *SLL* c già presenta un valore $M_{CP} = 1$), si ha che il numero di termini, nella sommatoria, che contribuiscono al valore OV_C è più basso nei *SLL* che presentano *ECI* più alto e viceversa.

Oltre a questo, si deve inoltre tener conto del fatto che i *SLL* più tecnologicamente complessi tendono già a realizzare i brevetti più tecnologicamente complessi (che presentano ovvero un più elevato valore di TCI_P): per tale motivo nei *SLL* più tecnologicamente complessi i termini della sommatoria, oltre a essere tendenzialmente inferiori in numero, sono anche numeri tendenzialmente più piccoli poiché caratterizzati da più bessi valori di PCI_P (poiché quelli già più tecnologicamente complessi tali *SLL* li esportano già).

Per questi due motivi se ne deduce che *Sistemi Locali del Lavoro* meno tecnologicamente complessi non possono che avere maggiori valori della metrica *Opportunity Value*. Quanto fin qui detto si può anche osservare dal successivo *Grafico 4.3.*

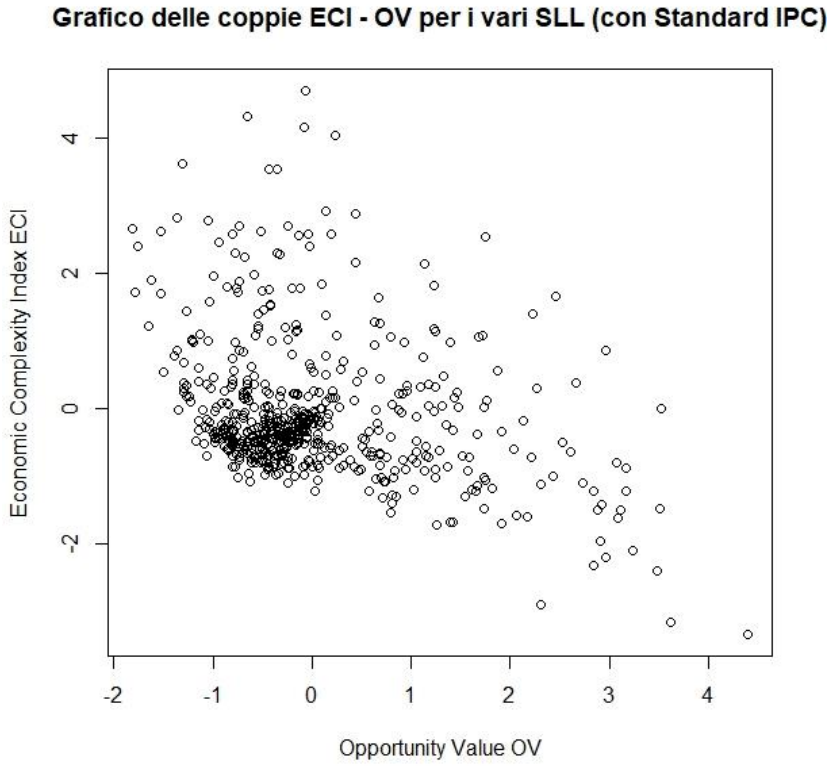

Grafico 4.3: Grafico delle coppie *ECI – OV* per i vari *SLL* (con *Standard IPC*)

IV.2.2. Calcolo di ECI, PCI e OV adottando la classificazione NACE

Così come fatto con lo *Spazio delle Tecnologie* si riportano tutte le tabelle relative alle metriche del precedente paragrafo calcolate adottando lo *Standard IPC*, impiegando invece la classificazione *NACE*.

Nelle due tabelle che seguono, le *Tabella 4.12* e *4.13*, si riportano i primi 10 *SLL* e gli ultimi 10 *SLL* ottenuti ordinando per valori decrescenti il vettore *ECI*, a sua volta costruito adottando la classificazione *NACE*.

Posizione$_{NACE}$	*Codice SLL*	*SLL*	*ECI*	*Diversità $k_{C,0}$*	*Ubiquità Media $k_{P,1}$*	*Posizione$_{IPC}$*
1°	909	MONTECATINI-TERME	3,54	13	126	38°
2°	1102	FANO	3,35	13	129	128°
3°	304	COMO	3,34	11	128	13°
4°	940	SINALUNGA	2,99	8	126	114°
5°	422	ROVERETO	2,75	13	119	123°
6°	1114	MACERATA	2,74	11	129	79°
7°	833	FORLÌ	2,68	11	131	67°
8°	1103	PERGOLA	2,55	7	125	273°
9°	131	BIELLA	2,54	12	133	15°
10°	1104	PESARO	2,50	9	142	40°

Tabella 4.12: Classifica dei primi 10 *SLL* per valore della metrica *ECI* adottando la classificazione *NACE*

Posizione$_{NACE}$	*Codice SLL*	*SLL*	*ECI*	*Diversità $k_{C,0}$*	*Ubiquità Media $k_{P,1}$*	*Posizione$_{IPC}$*
573°	905	CASTELNUOVO DI GARFAGNANA	-1,68	5	197	221°
574°	1514	CASTELLAMMARE DI STABIA	-1,71	6	198	578°
575°	1606	MANFREDONIA	-1,71	5	187	552°
576°	1949	LEONFORTE	-1,73	6	173	562°
577°	1503	PIEDIMONTE MATESE	-1,75	6	155	515°
578°	1619	RUTIGLIANO	-1,80	6	203	565°
579°	1907	BAGHERIA	-1,81	6	186	571°
580°	1517	NAPOLI	-1,87	11	143	44°
581°	1608	SAN GIOVANNI ROTONDO	-2,01	4	187	514°
582°	1903	MARSALA	-2,04	9	172	581°

Tabella 4.13: Classifica degli ultimi 10 *SLL* per valore della metrica *ECI* adottando la classificazione *NACE*

Anche qui, come con lo standard *IPC*, si può osservare come *SLL* maggiormente complessi tendono ad essere più diversificati e tendono a realizzare tecnologie meno ubique; con la classificazione *NACE* come criterio di classificazione dei brevetti, questa volta i tre *SLL* tecnologicamente più complessi sono "Montecatini-Terme", "Fano" e "Como". Anche qui, nei primi 3 *SLL* si brevettano la maggior parte delle tecnologie che si trovano nella classifica delle prime 10 più complesse divisioni brevettuali *NACE*; più in particolare, queste divisioni brevettuali (del quale si riporta a margine anche la posizione nella classifica del *Technology Complexity Index* della *Tabella 4.14*) sono:

- a *Montecatini-Terme* si brevettano le divisioni *28* (1°), *14* (3°), *13* (4°), *18* (5°), *15* (7°), *11* (8°), *17* (9°), *42* (10°);
- a *Fano* si brevettano le divisioni *28* (1°), *31* (2°), *14* (3°), *13* (4°), *18* (5°), *16* (6°), *17* (9°);
- a *Como* si brevettano le divisioni *28* (1°), *31* (2°), *14* (3°), *13* (4°), *18* (5°), *16* (6°), *17* (9°);

Scendendo ancor più nel dettaglio, *Montecatini-Terme* vanta ad esempio un *Reaveled Comperative Advantge RCA* pari a "+6,82" nella divisione brevettuale *11* ("Produzione di bevande" - la 8° maggiormente complessa), un *RCA* pari a "+3,54" nella divisone *17* ("Fabbricazione di carta e prodotti di carta" – la 9° maggiormente complessa) e un *RCA* pari a "+2,75" nella divisone *15* ("Confezione di articoli in pelle e simili" – la 7° maggiormente complessa).

Nella tabella che segue, invece, la classifica delle divisioni brevettuali ottenuta ordinando il vettore *TCI*, costruito adottando la classificazione *NACE*, per valori decrescenti della metrica stessa.

Prodotto	*TCI*	*Attività*	*Ubiquità* $k_{P,0}$	*Diversità media* $k_{C,1}$
28	1,63	Fabbricazione di macchinari e apparecchiature n.c.a.	240	6
31	0,99	Fabbricazione di mobili	131	8
14	0,79	Confezione di articoli di abbigliamento	105	8
13	0,79	Industrie tessili	71	8
18	0,76	Stampa e riproduzione su supporti registrati	80	8
16	0,73	Industria del legno e dei prodotti in sughero	59	8
15	0,68	Confezione di articoli in pelle e simili	91	8
11	0,51	Produzione di bevande	78	8
17	0,49	Fabbricazione di carta e prodotti di carta	39	9
42	0,47	Ingegneria civile	116	8
24	0,42	Attività metallurgiche	72	9
62	0,34	Programmazione, consulenza informatica e attività connesse	92	9

12	0,34	Industria del tabacco	50	9
23	0,21	Fabbricazione di altri prodotti della lavorazione di minerali non metalliferi	142	7
22	0,11	Fabbricazione di articoli in gomma e materie plastiche	126	7
19	0,07	fabbricazione di coke e prodotti derivati dalla raffinazione del petrolio	45	9
43	0,01	Lavori di costruzione specializzati	255	7
25	0,00	Fabbricazione di prodotti in metallo, esclusi macchinari e attrezzature	203	7
27	-0,10	Fabbricazione di apparecchiature elettriche	179	7
21	-0,28	Fabbricazione di prodotti farmaceutici di base e di preparati farmaceutici	74	8
10	-0,41	Industrie Alimentari	170	7
30	-0,99	Fabbricazione di altri mezzi di trasporto	165	7
29	-1,10	Fabbricazione di autoveicoli , rimorchi e semirimorchi	159	6
20	-1,38	Fabbricazione di prodotti chimici	167	7
32	-2,46	Altre Industrie manifatturiere	265	6
26	-2,63	Fabbricazione di computer e prodotti di elettronica e ottica	156	6

Tabella 4.14: Codici NACE cui afferiscono i brevetti oggetto di analisi, ordinati per valori decrescenti di TCI

Dalla *Tabella 4.14* si può notare come tecnologie più complesse (fatta eccezione per la "28") tendono ad essere meno ubique di quelle meno tecnologicamente complesse e i *SLL* nella quale si brevettano divisioni *NACE* più complesse tendono anche qui ad avere una *Diversità media* più elevata. Quanto fin qui detto si può anche osservare dai due *Grafici 4.4* e *4.5*.

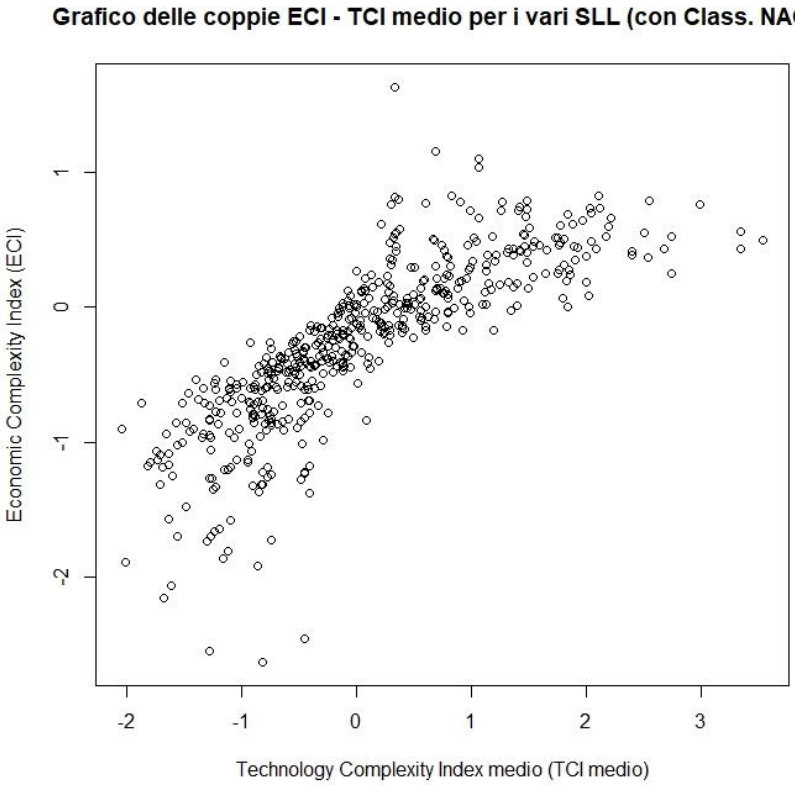

Grafico 4.4: Grafico delle coppie *ECI – TCI medio* per i vari SLL (con *Class. NACE*)

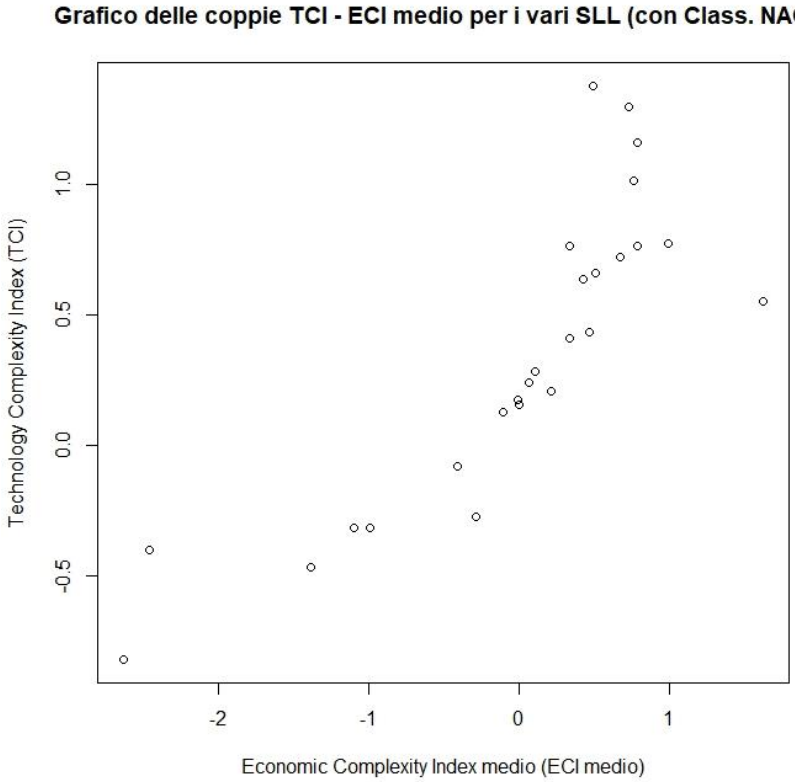

Grafico 4.5: Grafico delle coppie *TCI – ECI medio* per i vari SLL (con *Class. NACE*)

Nella tabella che segue, infine, i due estremi del vettore *Opportunity Value*, così ottenuti dopo aver riordinato lo stesso per valori decrescenti della variabile stessa.

Posizione	*Codice SLL*	*SLL*	*Opportunity Value*	*Poisizione$_{ECI}$*
1°	1517	NAPOLI	2,61	580°
2°	919	LIVORNO	2,43	556°
3°	1301	AVEZZANO	2,38	420°
4°	1903	MARSALA	2,28	582°
5°	1914	PALERMO	2,03	389°
6°	1501	CASERTA	1,91	513°
7°	1543	SALERNO	1,83	516°
8°	1612	BARI	1,77	549°
9°	1925	MESSINA	1,61	353°
10°	1518	NOLA	1,59	512°

Tabella 4.15: Primi 10 *SLL* per valore della metrica *OV,* calcolata adottando la classificazione *NACE*

Posizione	*Codice SLL*	*SLL*	*Opportunity Value*	*Poisizione$_{ECI}$*
573°	402	BOLZANO/BOZEN	-1,33	28°
574°	1104	PESARO	-1,36	10°
575°	525	CASTELFRANCO VENETO	-1,60	36°
576°	833	FORLÌ	-1,60	7°
577°	131	BIELLA	-1,73	9°
578°	304	COMO	-1,84	3°
579°	537	CITTADELLA	-1,86	24°
580°	1114	MACERATA	-1,91	6°
581°	1102	FANO	-2,25	2°
582°	909	MONTECATINI-TERME	-2,32	1°

Tabella 4.16: Ultimi 10 SLL per valore della metrica *OV,* calcolata adottando la classificazione *NACE*

Anche qui, come dalle due tabelle precedenti, i *SLL* caratterizzati da un maggior *Opportunity Value* sono quelli caratterizzati da un minor valore dell'*ECI* (e anche qui, come nella precedente classificazione, il *SLL* che presenta il più basso *Opportunity Value* è il *SLL* di "Montecatini-Terme" che è il *Sistema Locale* più tecnologicamente complesso) per i motivi già citati al paragrafo precedente. Quanto fin qui detto si può anche osservare dal successivo *Grafico 4.6.*

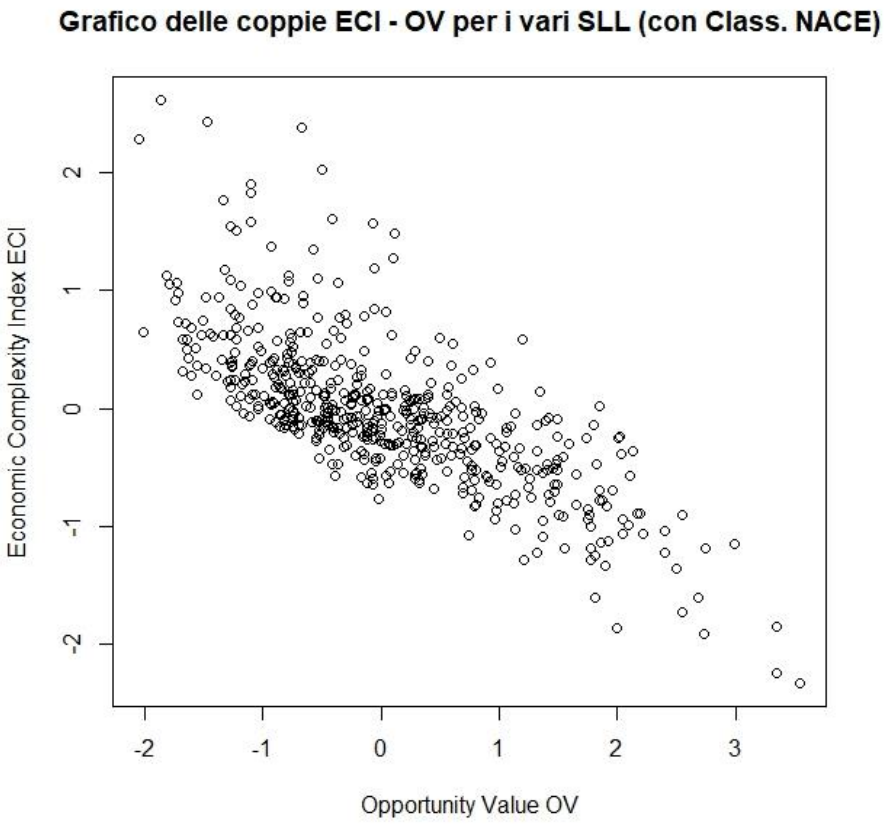

Grafico 4.6: Grafico delle coppie *ECI – OV* per i vari *SLL* (con *Class. NACE*)

IV.2.3. Confronto tra le classificazioni IPC e NACE e relative considerazioni

Avendo realizzato le analisi presentate nei precedenti paragrafi con entrambi i criteri di classificazione, lo *standard IPC* e la classificazione *NACE*, vale la pena ora fare un confronto tra i risultati ottenuti adottando entrambi i criteri.

In merito alle metriche *Economic Complexity Index ECI*, *Technology Complexity Index TCII* e *Opportunity Value OV*, si può, in virtù di quanto fin qui detto, asserire che i risultati mediamente ottenuto con i due distinti criteri sono pressoché simili, avendo dedotto le medesime considerazioni a partire dai risultati ottenuti: si sono, infatti, potute fare le medesime considerazioni in merito a tutte e tre le variabili adottando entrambi i succitati criteri.

La sostanziale differenza tra i risultati ottenuti si riscontra in realtà nel posizionamento dei diversi *Sistemi Locali del Lavoro* all'interno della classifica dell'*Economic Complexity Index ECI*: tale diverso posizionamento puntuale dei *SLL* non riflette però il posizionamento medio dei *SLL*, i quali, invece, presentano in media posizionamenti simili.

Come desumibile dall'ultima colonna delle *Tabelle 4.6*, *4.7*, *4.12* e *4.13*, in media i *SLL* che si posizionano nella classifica dei primi 10 *SLL* più tecnologicamente complessi adottando lo standard *IPC* si posizionano anche nella prima metà della corrispondente classifica costruita adottando la classificazione *NACE* e allo stesso modo i primi 10 *SLL* maggiormente complessi adottando la classificazione *NACE* si posizionano nella prima metà della corrispondente classifica adottando lo standard *IPC*. Le medesime considerazioni, adottando entrambi i succitati criteri, si possono fare in merito alle classifiche degli ultimi 10 *SLL* meno tecnologicamente complessi.

Il fatto che, quantomeno in media, i SLL si ripartiscono nelle medesime porzioni della classifica dell'*ECI* adottando due distinti criteri di classificazione ci dimostra la bontà delle considerazioni alle quali siamo fin qui giunti, oltre che la resilienza della *Relatedness Theory*.

La diversa allocazione puntuale dei *Sistemi Locali del Lavoro* si ritiene invece essere imputabile alla diversa granularità (numero medio e deviazione standard del numero di brevetti in ciascuna categoria dei diversi criteri) e all'allocazione delle varie classi brevettuali nelle diverse divisioni *NACE* realizzate tramite l'apposita, già in precedenza citata, tabella di concordanza; tali diversi risultati puntuali, però, si ritiene non minino le considerazioni finora fatte in merito alle varie metriche, considerazioni che si è visto riflettono quando suggerito dalla letteratura di riferimento.

IV.3. Analisi Statistica degli effetti di scala super-lineari

Come già descritto al *Capitolo 2*, è possibile trovare in letteratura, parlando di *Relatedness Theory* e *Complessità Economica*, evidenze empiriche della relazione "super-lineare" esistente tra la *Complessità* delle attività economiche e la concentrazione demografica nelle diverse aree urbane di un paese[5], dando prova che, almeno con riferimento alle attività brevettuali negli *Stati Uniti*, più un'attività economica è complessa, tanto più ampia è la sua tendenza a concentrarsi nelle grandi città, e dimostrando altresì, implicitamente, che è più probabile che le "conoscenze" si diffondano e diano vita a nuovi prodotti in grandi aree urbane piuttosto che nei piccoli centri.

Ci proponiamo, avendo a disposizione sia, nel *database* dei dati brevettuali, la localizzazione geografica (in termini di *SLL*) dei brevetti, sia, disponibile nel sito *istat.it*, un *database* con una rilevazione demografica dei *SLL* all'interno del quindicennio di raccolta dei dati brevettuali, di studiare la presenza di eventuali effetti "super-lineari" della conoscenza (la cui *proxy* è il volume di brevetti) sulla popolazione italiana all'interno del quindicennio 1999 – 2013.

Assumendo, per ipotesi, l'esistenza di una generica "legge di scala" del tipo " $y \sim x^{\beta}$ ", con $\beta > 0$, nella quale x rappresenta la *popolazione* di un centro urbano e y rappresenta il *numero di brevetti* nel centro urbano stesso, è facile notare che, a meno di un fattore positivo e costante che moltiplica *x*, essendo $\beta > 0$ e non potendo essere *x* negativo, si ha che il comportamento di *y* è condizionato dal valore assunto da β: più in particolare *y è* crescente a tassi decrescenti se $\beta < 1$, linearmente crescente se $\beta = 1$, crescente a tassi crescenti se $\beta > 1$.

Essendo l'obiettivo della nostra analisi riuscire a effettuare una stima di β all'interno del campione di brevetti a nostra disposizione, per raggiungere tale obiettivo operiamo nel modo seguente: ipotizzando che il fenomeno sia regolato dalla funzione

$$y = A\,x^{\beta}\,, \qquad A > 0, \beta > 0$$

con *x* e *y*, rispettivamente, la popolazione di un *SLL* e il corrispondente volume di brevetti, facendo il logaritmo naturale di ambo i membri otteniamo:

$$\ln y = \ln A + \beta \ln x$$

[5] Si veda Balland, P. A., Jara-Figueroa, C., Petralia, S. G., Steijn, M. P., Rigby, D. L., & Hidalgo, C. A. (2020). Complex economic activities concentrate in large cities. Nature Human Behaviour, 4(3), 248-254.

che, noti *y* e *x*, (e quindi noti i loro logaritmi) è un'equazione nelle due incognite *lnA* e β. Avendo a disposizione non la singola coppia (*y, x*) di campioni ma, invece, una coppia di vettori (***Y***, ***X***) contenenti ciascuno n_C campioni, il problema della stima di *lnA* e β si risolve tramite una regressione *OLS* nelle due variabili succitate.

IV.3.1 Analisi degli effetti di scala tra le diverse sezioni dello standard IPC

Fissato il vettore ***Y*** della popolazione dei *SLL* risulta interessante stimare il coefficiente β, non solo adottando come vettore ***X*** il vettore con il totale dei brevetti realizzati in ciascun *SLL* (il cui regressore ribattezzeremo come β_{TOT}) ma anche i vettori con i volumi di brevetti realizzati dai singoli *SLL* nelle specifiche *sezioni* (da A ad H) brevettuali dello standard *IPC* di modo da valutare l'impatto della singola *sezione* tecnologica (mediante i coefficienti β_A, … , β_H) sulla concentrazione demografica[6].

Si riporta nel seguito lo script di *R* grazie al quale si è effettuata la regressione, nel caso specifico dei brevetti della sezione brevettuale "A"; partendo da un file di nome "BrevA" con estensione ".txt" contenente le due colonne "N_Brev_A", contenente il volume di brevetti nel quindicennio considerato, e "PopSLL", contenente la popolazione del corrispondente *SLL*, in un opportuno script su *R* applichiamo il seguente codice:

```
Brev_A <- read.table("BrevA.txt", header = TRUE)
        log_Brev_A <- log(Brev_A)
        fmA <- lm(N_Brev_A ~ Pop_SLL, data = log_Brev_A )
```

R ci restituisce in "fmA" le stime di "lnA" e "β_A". Ripetendo tale *iter* sia per le 8 sezioni brevettuali dello standard *IPC* sia per il totale dei brevetti nel loro complesso, otteniamo 9 stime di β (una per ciascuna sezione e una per la totalità dei brevetti); tralasciando le considerazioni in merito ai vari generici regressori stimati "lnX" (in quanto parametri non portatori di interesse per la nostra analisi) si riportano, nella tabella che segue, i risultati delle stime dei 9 citati regressori "β", con le corrispondenti "misure di bontà" della regressione.

[6] Laddove il numero di brevetti realizzati da un *SLL* in una specifica sezione brevettuale sia pari a zero, per non incorrere nel problema legato al dover effettuare il logaritmo di "zero", si è convenuto di sostituire, prima del logaritmo, tale "zero" con un "uno", in quanto l'esiguità di tale numero è tale da non inficiare, comunque, la validità dei risultati della regressione.
Un'altra via praticabile sarebbe potuta essere la rimozione delle righe dei vettori ***Y*** e ***X*** corrispondenti a produzioni brevettuali nulle di modo da rimuovere alla radice il problema del dover eseguire il logaritmo dello "zero". L'esercizio di tale pratica restituirebbe però un dominio di valori che non corrisponde a quello desiderato in quanto così facendo si "taglierebbe" via quella "coda" di valori, grazie anche ai quali la superlinearità si manifesta: tale pratica risulterebbe quindi, in definitiva, non corretta.

Sezione$_{IPC}$	*Descrizione*	β	*Standard Error*	R^2	*PCI$_{Medio}$*
A	Necessità Umane	1,25	0,04098	0,62	-0,76
B	Operazioni, Trasporti	1,28	0,04611	0,57	0,27
C	Chimica, Metallurgia	0,97	0,0332	0,57	0,30
D	Tessuti, Carta	0,53	0,03583	0,27	0,54
E	Costruzioni Fisse	1,04	0,03788	0,57	-0,81
F	Ingegneria meccanica; Illuminazione; Riscaldamento; Armi; Distruzione	1,09	0,04103	0,55	0,08
G	Fisica	1,03	0,03587	0,59	-0,09
H	Elettricità	0,89	0,03451	0,53	-0,56
TOT	*Tutte le classi fin qui citate*	1,41	0,04296	0,65	≈ 0

Tabella 4.17: Stime del regressore β e relative misure di bontà

Come già descritto al *Capitolo 2*, data la "legge di scala" $y \sim x^{\beta}$ adottata per spiegare il fenomeno oggetto di studio, il volume brevettuale in funzione delle popolazione dei *SLL* sarà, rispettivamente, crescente a tassi crescenti, a tasso fisso, o a tassi decrescenti se, rispettivamente, $\beta > 1$, $\beta = 1$ e $\beta < 1$.

Sulla base di ciò e di quanto ottenuto dalle regressioni (e rappresentato nella soprastante *Tabella 4.17*) possiamo trarre diverse considerazioni. Se consideriamo le singole *Sezioni* brevettuali si può osservare che:

- $\beta > 1$ per le sezioni "A" (Necessità Umane), "B" (Operazioni, Trasporti); "F" (Ingegneria meccanica, Illuminazione, Riscaldamento, Armi, Distruzione);
- $\beta \approx 1$ per le sezioni "C" (Necessità Umane), "E" (Costruzioni Fisse); "G" (Fisica);
- $\beta < 1$ per le sezioni "D" (Tessuti, Carta), "H" (Elettricità);

Dalla medesima tabella, osservando il valore della complessità tecnologica *TCI* di ciascuna sezione possiamo asserire che non sembra esservi un legame significativo tra il valore di β e il valore di *TCI*. Si osserva infatti come *TCI* positivi e negativi siano presenti sia tra le sezioni con $\beta > 1$, sia tra quelle con $\beta \approx 1$, che tra quelle con $\beta < 1$.

Osservando anche i grafici a dispersione presenti nell' "*Allegato A*", delle considerazioni specifiche si posso fare in merito alla stima del regressore per la sezione "D" e alla stima del regressore per la totalità dei brevetti; la sezione "D", come si può osservare dai *Grafici A.7* e *A.8* presenta una nube di punti caratterizzati da elevato volume brevettuale in corrispondenza di bassi volumi demografici. Da tale accentramento si ipotizza ne consegua, nel modello di regressione, una così bassa stima di β.

Dai *Grafici A.17* e *A.18* si osserva invece come, impiegando la totalità dei brevetti, il modello sembra descrivere maggiormente la varianza dei dati (maggior R^2 rispetto a tutte le altre regressioni) e la stima del regressore assume in tale caso il maggiore tra i valori in tabella (β_{TOT} = 1,41), valore che non è, peraltro, la media dei valori stimati per le varie sezioni.

Da tali valori si può quindi trarre la deduzione che applicare il modello di regressione descritto a una "partizione" del dataset "indebolisce" la bontà del modello stesso e "sminuisce" il valore del regressore stimato in quanto l'introduzione esogena di un arbitrario criterio di classificazione fa sì che non si tenga conto, nella regressione, della possibilità che un certo volume brevettuale in un certo SLL possa essere dovuto e legato alla produzione di brevetti in un'altra delle categorie brevettuali.

Nel grafico che segue l'insieme delle coppie ("Numero di Brevetti" - "Popolazione del SLL") con la stima della curva di super-linearità per tutte le sezioni e per la totalità dei brevetti.

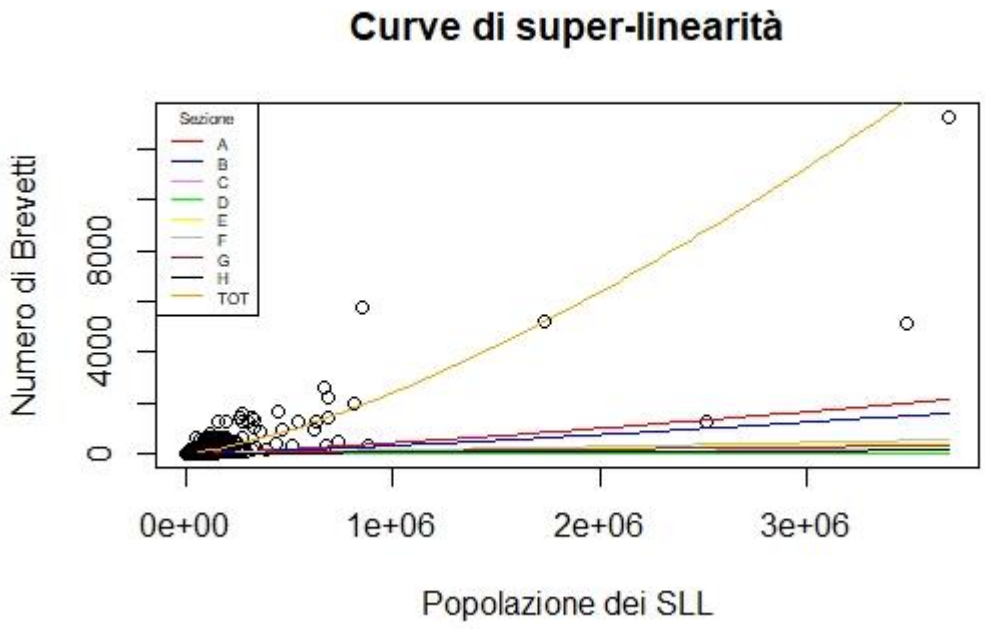

Grafico 4.7: Grafico a Dispersione rappresentativo delle coppie "*Numero di Brevetti – Popolazione dei SLL*" per i brevetti appartenenti alla Sezione "H" dello Standard *IPC* con rappresentazione delle curve super-lineari di regressione stimate.

IV.4 Analisi degli effetti di scala tra i gli SLL ricadenti in diversi quartili della metrica ECI

Oltre all'analisi dei diversi effetti di scala manifestatisi in base alle diverse sezioni dello standard *IPC* considerate, un ulteriore interessante conducibile analisi è quella di studiare gli effetti di super-linearità tra gli *SLL* ricadenti in diversi quartili della metrica *ECI*. Dal vettore *ECI*, di dimensioni 582 x 1, già calcolato e discusso ai paragrafi precedenti di questo capitolo, è possibile dividere gli *SLL* in 4 gruppi in base al quartile della metrica *ECI* in cui essi ricadono.

Come nel paragrafo precedente, indicando con Y_{NQ} il vettore della popolazione dei *SLL* ricadenti nel *n*-esimo quartile e con X_{NQ} il vettore dei corrispondenti brevetti di ciascun *SLL*, adottando il precedente modello di regressione e l'associato *script*, possiamo ricavare il valore del regressore β, per ciascuna porzione del vettore *ECI*. Si riportano, nella tabella seguente, i risultati di tali analisi (con le corrispondenti misure di bontà):

Quartile	β	*p-Value*	*Popolazione Media*	R^2
1Q	0,97	*p-Value* < 2*10^(-16)	234.973	0,69
2Q	1,35	*p-Value* < 2*10^(-16)	64.738	0,63
3Q	1,11	*p-Value* < 2*10^(-16)	40.981	0,42
4Q	0,75	*p-Value* < 6*10^(-16)	65.979	0,61

Tabella 4.18: Stime del regressore β per i *SLL* ricadenti nei differenti quartili della metrica ECI

Così come, al termine del paragrafo precedente, si è cercato un legame tra la misura dell'"effetto-scala" (il regressore β) e la complessità tecnologica della sezione brevettuale (la metrica *TCI*) - senza osservare un legame significativo tra le due metriche - analogamente non sembra manifestarsi, come si può desumere dai valori rappresentati nella *Tabella 4.18* un legame tra la stima di β e la complessità tecnologica dei territori, essendo inferiori all'unità le stime di β per i brevetti realizzati nei *SLL* presenti nel 1° e nel 4° quartile del vettore *ECI* e superiori all'unità quelle dei brevetti realizzati nei *SLL* presenti nel 2° e nel 3° quartile, non manifestandosi quindi un chiaro ordinamento nei valori di β che ripercorre l'ordinamento della metrica *ECI*.

CAPITOLO V

CONCLUSIONI

Come specificato nel capitolo introduttivo, obiettivo del presente elaborato è la ricerca delle ipotesi applicative della *Relatedness Theory* a un database di dati brevettuali (rappresentativo della distribuzione spaziale della conoscenza in Italia) e l'applicazione, al medesimo, delle metodologie di calcolo tipiche di tale teoria al fine di calcolarne le metriche e trarne le opportune conclusioni.

L'applicazione della citata teoria al database dei volumi brevettuali (diffusamente trattato e studiato al *Capitolo 3*), per quanto apparentemente distante dall'originaria applicazione della *Relatedness Theory* (la quale, come già accennato, nasce al fine di poter confrontare le "complessità" delle produzioni di beni e servizi realizzati nei vari paesi del mondo) ben si presta, in realtà, allo studio della distribuzione spaziale delle conoscenze (di cui i brevetti sono una *proxy*) in quanto:

- le unità territoriali (i *Sistemi Locali del Lavoro*) per il quale tali volumi brevettuali sono specificati rappresentano, per loro stessa definizione, porzioni di territorio "autonome" dal punto di vista economico, lavorativo e sociale, mutando così le caratteristiche di "autonomia" delle porzioni spaziali (gli stati) impiegate nell'applicazione originaria;
- i brevetti, essendo catalogabili in specifiche categorie grazie a sistemi di classificazione (lo standard *IPC* e la classificazione *NACE*) esclusivi ed esaustivi, sono dei "prodotti" dell'ingegno umano ai quali si possono mutuare le considerazioni fatte nell'applicazione originaria ai prodotti e ai servizi realizzati da ciascuna porzione spaziale.

In virtù di quanto sopra si reputano come consistenti le ipotesi per l'applicazione della *Relatedness Theory* al *database* in esame (le cui modalità applicative sono discusse in questo elaborato) e analogamente consistenti si possono ritenere i risultati ottenuti.

Per quanto già diffusamente esposto e discusso nei vari paragrafi del *Capitolo 4*, esponiamo in maniera compatta, in questo capitolo, i risultati e le considerazioni cui siamo giunti.

Le *Proximity Matrix* (o gli *Spazi delle Tecnologie* se quest'ultime sono rappresentate tramite grafi) rappresentano sicuramente degli strumenti eccezionali per capire lo stato della diffusione della

conoscenza (sia essa produttiva come nell'applicazione originaria o tecnologica come in questo elaborato), nonché per intuire, coniugandole con altre metriche, le nuove conoscenze che, in futuro, tali territori tenderanno ad acquisire, rappresentando pertanto, a parere dello scrivente, uno straordinario strumento di ausilio per chi, sia a livello governativo che a livello imprenditoriale, è chiamato a fare scelte d'investimento.

La possibilità, infatti, per un imprenditore, di trovare già in uno specifico territorio le conoscenze necessarie per intraprendere un'attività rappresenta, già nel breve termine, una non indifferente opportunità di contrazione dei costi legati al reperimento di tali conoscenze e quindi una maggiore potenziale profittabilità.

Analogamente, per chi è chiamato a fare delle decisioni di spesa a livello governativo, lo *Spazio delle Tecnologie* può evitare, nel breve periodo, grossolani errori d'investimento (quali la realizzazione delle infrastrutture necessarie a realizzare un certo prodotto, per cui sono necessarie specifiche conoscenze, quando le conoscenze necessarie per realizzarlo non sono presenti nel territorio in oggetto – ovvero, in breve, una "cattedrale in un deserto") e similarmente può, nel medio e nel lungo periodo, contribuire a scelte d'investimento che tengano conto del principio di equità sociale mirando alla redistribuzione della conoscenza (foraggiando e stimolando gli *spillover* delle conoscenze legate a tecnologie con maggiormente complesse, cioè con maggior *TCI*) verso territori che aspettano il giusto stimolo per veder sbocciare le loro potenzialità (valutate mediante la metrica *Opportunity Value*) con le intuibili conseguenze reddituali e sociali che ne conseguono.

In tal senso la *Relatedness Theory* si presenta come uno strumento eccezionale, ancora poco sfruttato, per chiunque sia chiamato a fare decisioni d'investimento.

I risultati cui siamo pervenuti, così come le relative considerazioni, si possono considerare congrui con quanto esposto in letteratura; con entrambi i criteri di classificazione adottati, infatti, abbiamo potuto appurare come, in media:

- la *Diversità* $k_{C,0}$ delle conoscenze presenti in un territorio e l'*Ubiquità media* $k_{P,1}$ delle stesse sono due metriche i cui valori sono inversamente proporzionali in quanto la realizzazione di nuovi prodotti che richiedono apporti di variegate conoscenze si realizza più facilmente dove già tali conoscenze esistono e tanto più difficilmente, in media, all'aumentare del numero di conoscenze che mancano, in un territorio, per realizzarli;

- l'*Ubiquità* $k_{P,0}$ di una certa categoria brevettuale e la *Diversità media* $k_{C,1}$ dei territori in cui tale tecnologia si realizza, sono parimenti anch'esse metriche i cui valori sono inversamente proporzionali in quanto le tecnologie più ubique sono quelle più facilmente realizzabili anche nei territori nella quale non sono presenti molte, diverse, conoscenze tecnologiche e che, in quanto tali, difficilmente possono realizzare prodotti che richiedono conoscenze diversificate non già disponendo delle stesse.

- tanto maggiore è la *Complessità Economica ECI* di un territorio tanto più elevata è la *Diversità* $k_{C,0}$ delle conoscenze in esso presenti e tanto più alta è, in virtù di quanto detto ai punti precedenti, bassa l'*Ubiquità media* $k_{P,1}$ delle conoscenze stesse.

- tanto maggiore è la *Complessità Tecnologica TCI* di un prodotto tanto più bassa è l'*Ubiquità* $k_{P,0}$ dello stesso e tanto più alta è, per quanto detto ai punti precedenti, la *Diversità media* $k_{C,1}$ dei territori che la realizzano.

- la *Complessità Economica ECI* di un territorio e la *Complessità Tecnologica media* TCI_{Media} delle conoscenze in esso presenti (e, parimenti, la *Complessità Tecnologica TCI* di una categoria brevettuale e la *Complessità Economica media* ECI_{Medio} dei territori in questa si realizza) non possono che essere, per quanto fin qui detto, direttamente proporzionali.

- Le *Opportunità* (misurate tramite la metrica *Opportunity Value*) insite in un territorio sono inversamente proporzionali alla sua *Complessità Economica ECI*, poiché i territori meno economicamente complessi possono ambire a "produrre" molte conoscenze più tecnologicamente complesse (cioè con elevato *PCI*) che non già producono e che, non essendo ancora "prodotte", rendono il valore della metrica *ECI* più basso.

Si sottolinea come le considerazioni sopra riportate, fatte grazie allo studio e all'elaborazione del *database* fornito dall'ateneo e al calcolo delle specifiche metriche, sono in linea con quanto suggerito dalla letteratura di riferimento.

Tali considerazioni sono state inoltre dedotte adottando due diversi criteri di classificazione dei brevetti, motivo per cui si evidenzia la bontà dei risultati ottenuti.

Anche per ciò che concerne un eventuale effetto di super-linearità dei volumi brevettuali rispetto alla popolazione dei territori in cui questi si realizzano si può, per quanto visto al *Capitolo 4*, ritenere che all'aumentare della popolazione di un territorio i volumi brevettuali in questo realizzati tendono ad

aumentare più che proporzionalmente (per quanto tale effetto sia legato alla natura specifica della tecnologia brevettuale), risultato, anche questo, che trova riscontro in letteratura.

Si sottolinea, infine, come i risultati sopra specificati siano risultati che, per quanto validi in media, possono puntualmente (cioè per specifico territorio o specifica tecnologia) ampiamente differire cambiando criterio di classificazione come desumibile dalle elaborazioni presentate.

Come descritto nel capitolo introduttivo, alla base dell'esponenziale accrescimento della conoscenza umana cui si è potuto assistere negli ultimi secoli, vi è la maggior propensione, che le società moderne hanno, nello sfruttare in modo coordinato i singoli, modesti apporti alla conoscenza dell'umanità derivanti da tutti membri che la compongono anziché affidarsi, come in passato, agli sparuti lampi di genialità di singoli individui isolati.

Allo stesso modo, perseguendo la medesima filosofia, e dato il diverso e originale contesto applicativo della *Relatedness Theory* rispetto alla sua originale applicazione, si spera, anche con questo elaborato, di dare un piccolo, modesto contribuito alla conoscenza dell'umanità.

ALLEGATO A – GRAFICI A DISPERSIONE DELLE CURVE DI SUPER-LINEARITÀ

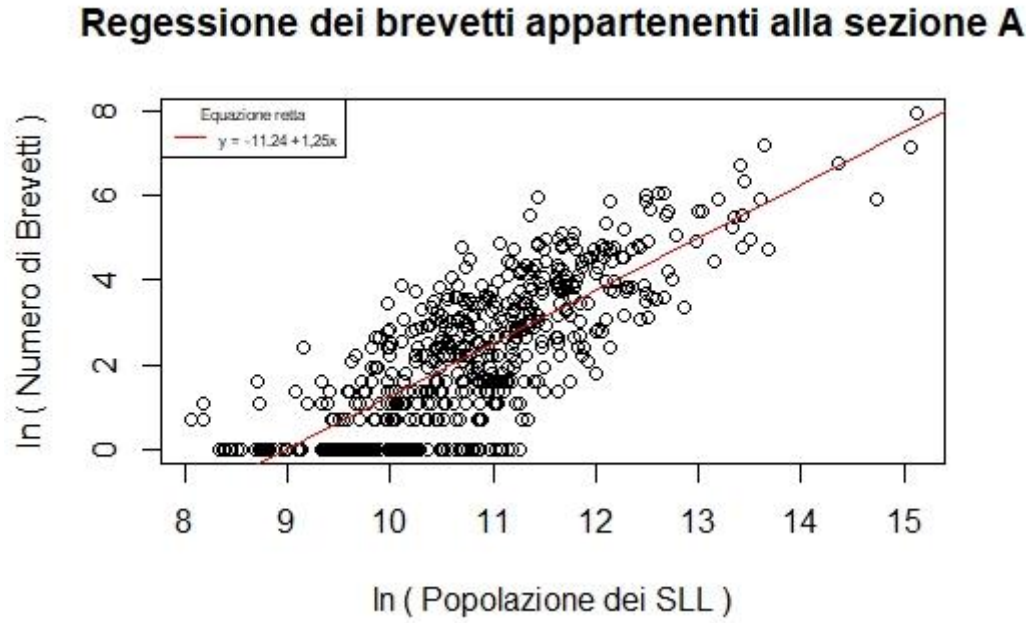

Grafico A.1: Grafico a Dispersione rappresentativo delle coppie "ln(*Numero di Brevetti*) – ln(*Popolazione dei SLL*)" per i brevetti appartenenti alla Sezione "A" dello Standard *IPC* con rappresentazione della retta di regressione stimata (in rosso).

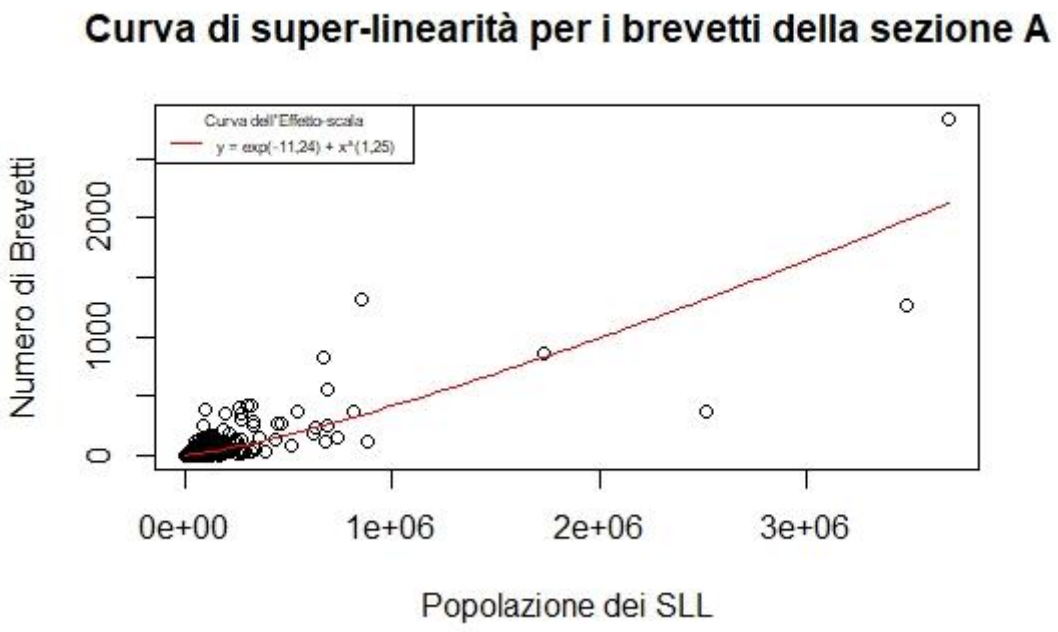

Grafico A.2: Grafico a Dispersione rappresentativo delle coppie "Numero di Brevetti – Popolazione dei SLL" per i brevetti appartenenti alla Sezione "A" dello Standard *IPC* con rappresentazione della curva super-lineare di regressione stimata (in rosso).

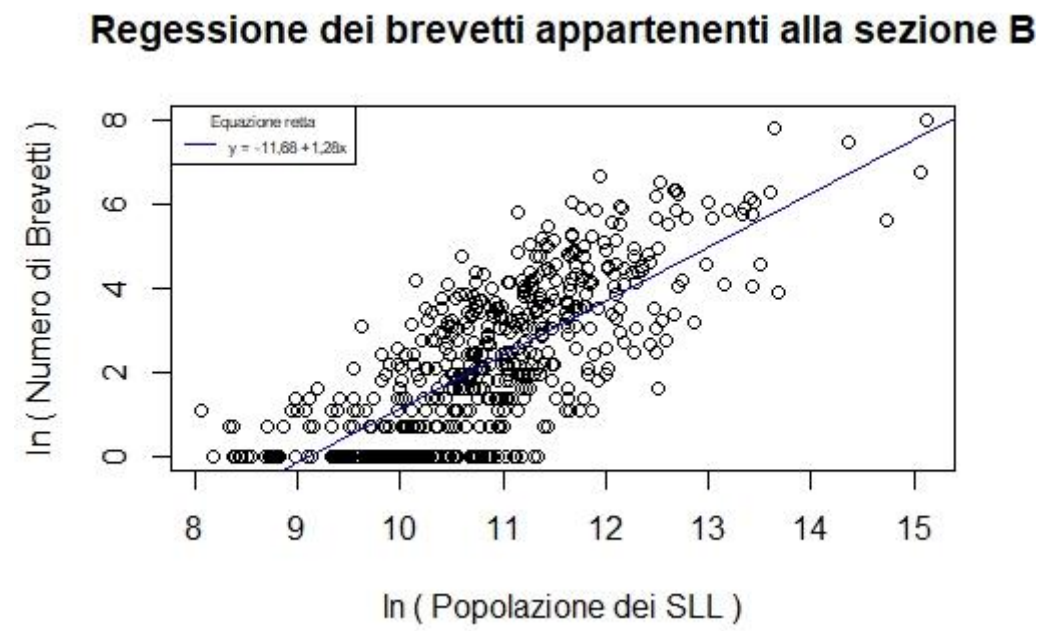

Grafico A.3: Grafico a Dispersione rappresentativo delle coppie "ln(*Numero di Brevetti*) – ln(*Popolazione dei SLL*)" per i brevetti appartenenti alla Sezione "B" dello Standard *IPC* con rappresentazione della retta di regressione stimata (in blu).

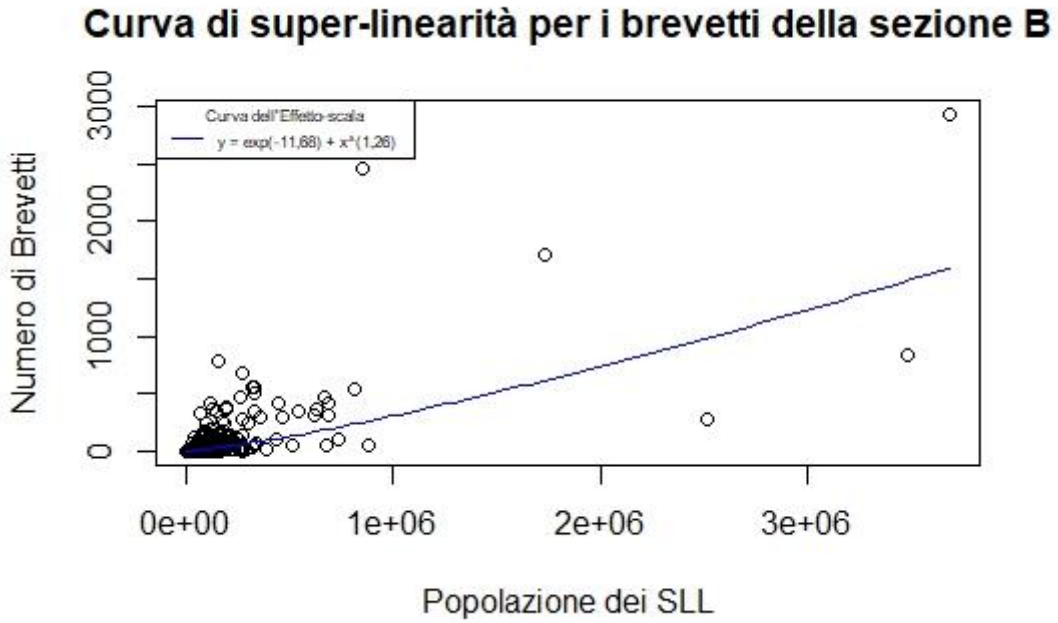

Grafico A.4: Grafico a Dispersione rappresentativo delle coppie "*Numero di Brevetti* – *Popolazione dei SLL*" per i brevetti appartenenti alla Sezione "A" dello Standard *IPC* con rappresentazione della curva super-lineare di regressione stimata (in blu).

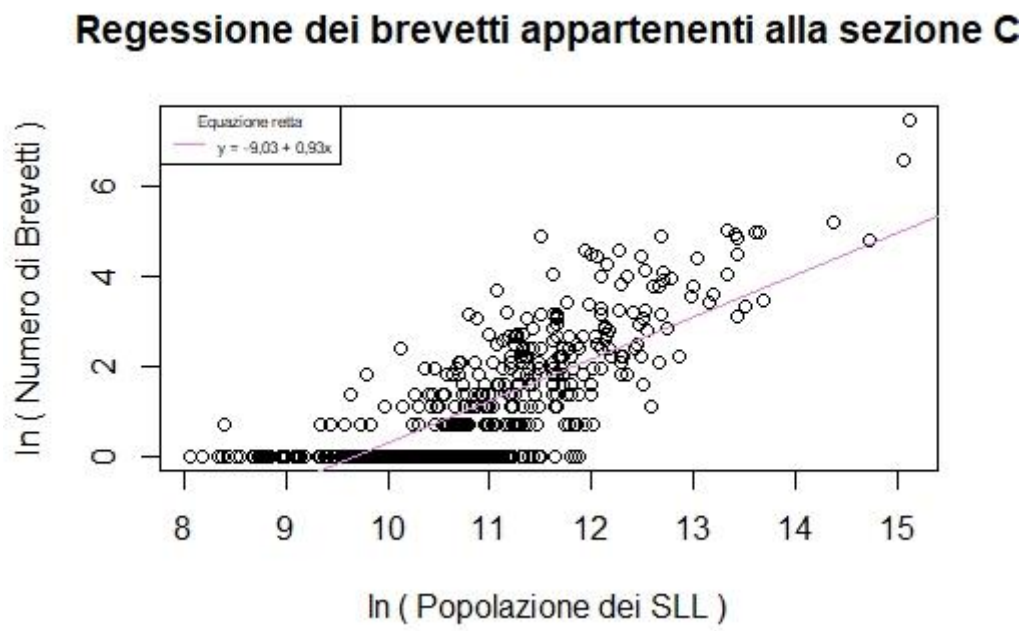

Grafico A.5: Grafico a Dispersione rappresentativo delle coppie "ln(*Numero di Brevetti*) – ln(*Popolazione dei SLL*)" per i brevetti appartenenti alla Sezione "C" dello Standard *IPC* con rappresentazione della retta di regressione stimata (in viola).

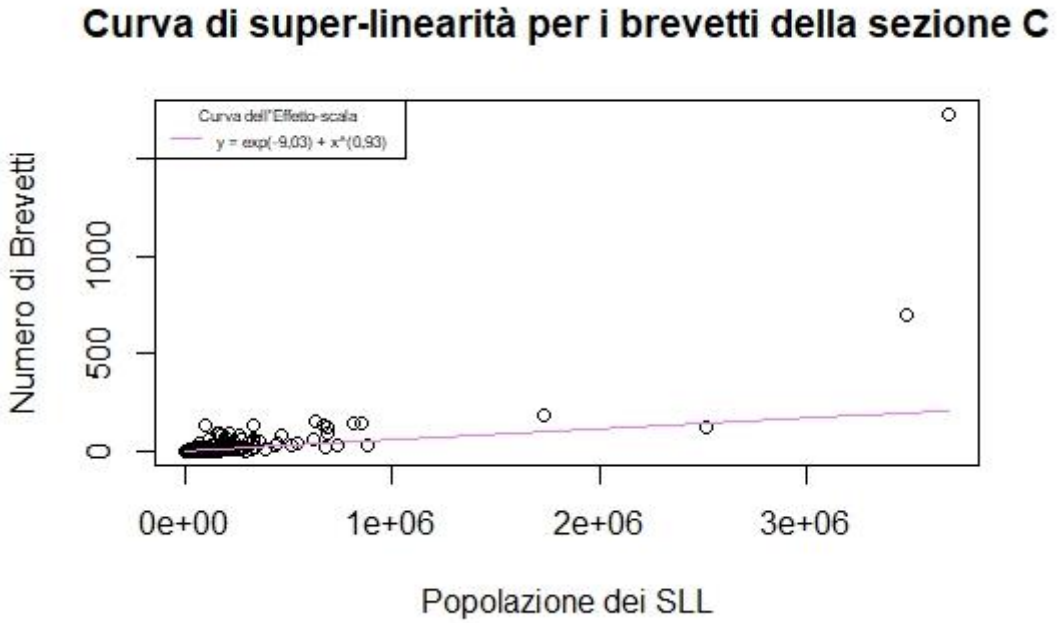

Grafico A.6: Grafico a Dispersione rappresentativo delle coppie "*Numero di Brevetti – Popolazione dei SLL*" per i brevetti appartenenti alla Sezione "C" dello Standard *IPC* con rappresentazione della curva super-lineare di regressione stimata (in viola).

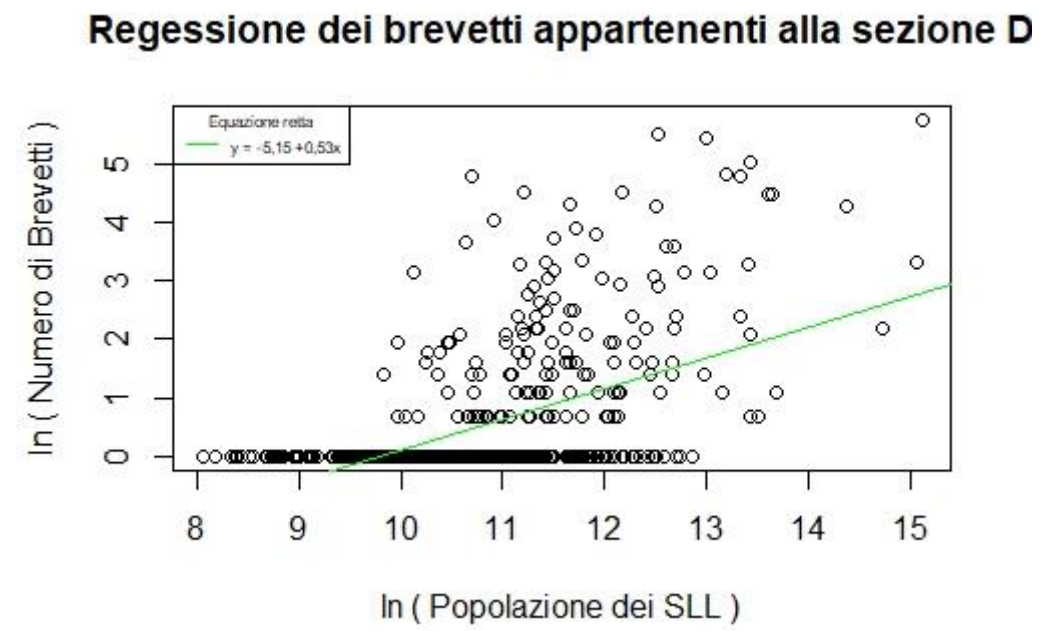

Grafico A.7: Grafico a Dispersione rappresentativo delle coppie "ln(*Numero di Brevetti*) – ln(*Popolazione dei SLL*)" per i brevetti appartenenti alla Sezione "D" dello Standard *IPC* con rappresentazione della retta di regressione stimata (in verde).

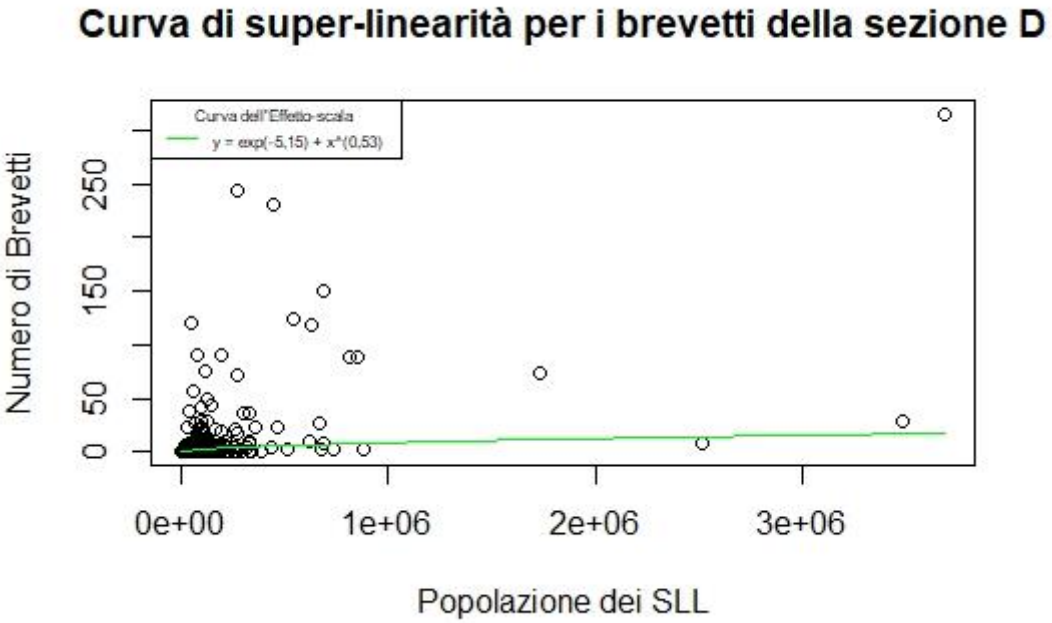

Grafico A.8: Grafico a Dispersione rappresentativo delle coppie "*Numero di Brevetti – Popolazione dei SLL*" per i brevetti appartenenti alla Sezione "D" dello Standard *IPC* con rappresentazione della curva super-lineare di regressione stimata (in verde).

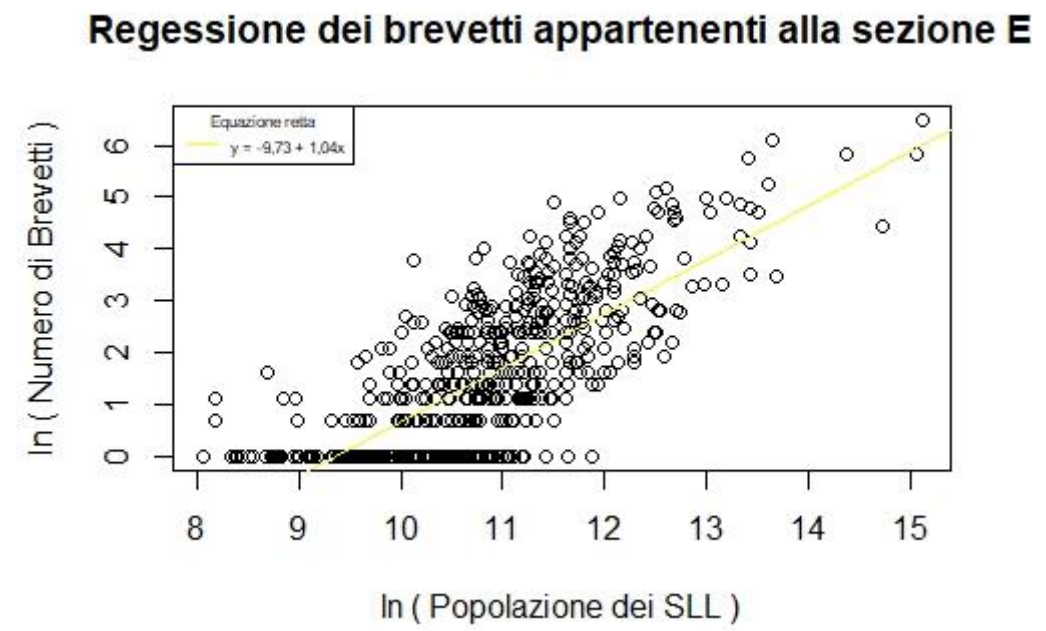

Grafico A.9: Grafico a Dispersione rappresentativo delle coppie "ln(*Numero di Brevetti*) – ln(*Popolazione dei SLL*)" per i brevetti appartenenti alla Sezione "E" dello Standard *IPC* con rappresentazione della retta di regressione stimata (in giallo).

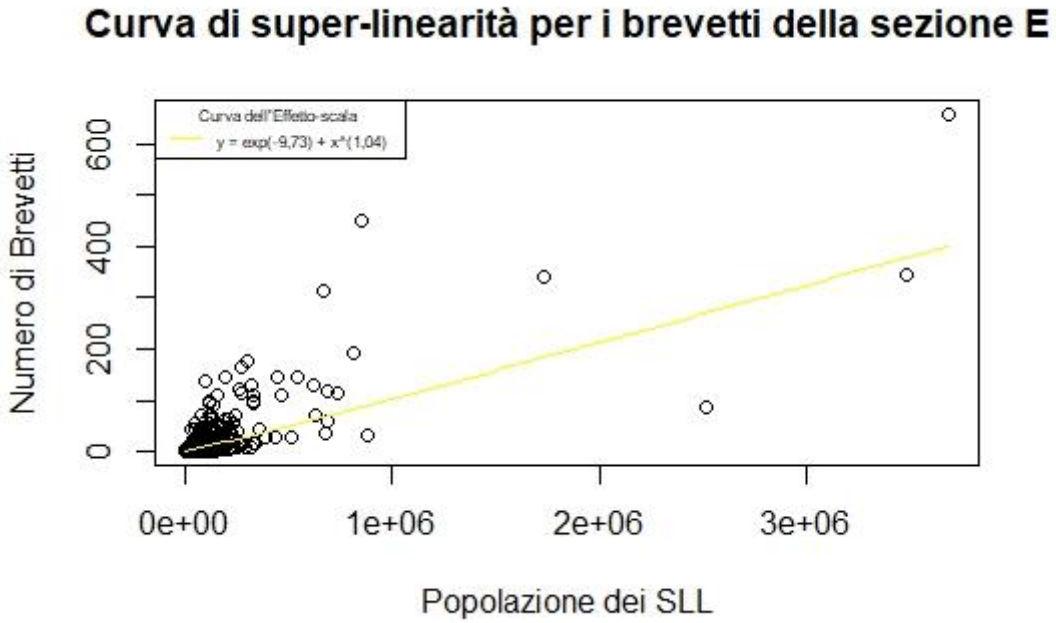

Grafico A.10: Grafico a Dispersione rappresentativo delle coppie "*Numero di Brevetti* – *Popolazione dei SLL*" per i brevetti appartenenti alla Sezione "E" dello Standard *IPC* con rappresentazione della curva super-lineare di regressione stimata (in giallo).

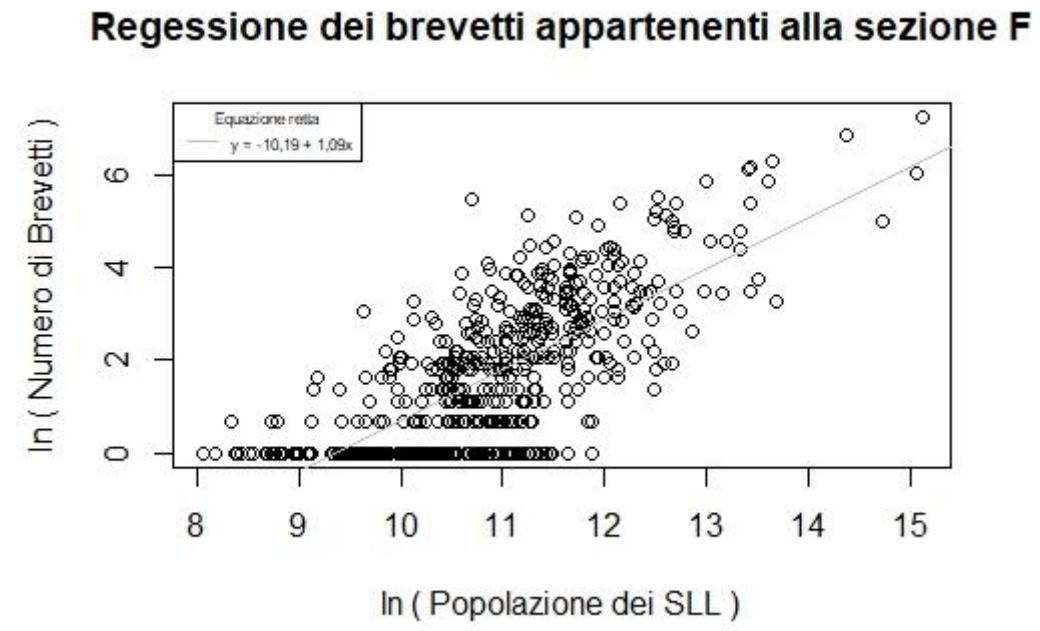

Grafico A.11: Grafico a Dispersione rappresentativo delle coppie "ln(*Numero di Brevetti*) – ln(*Popolazione dei SLL*)" per i brevetti appartenenti alla Sezione "F" dello Standard *IPC* con rappresentazione della retta di regressione stimata (in grigio).

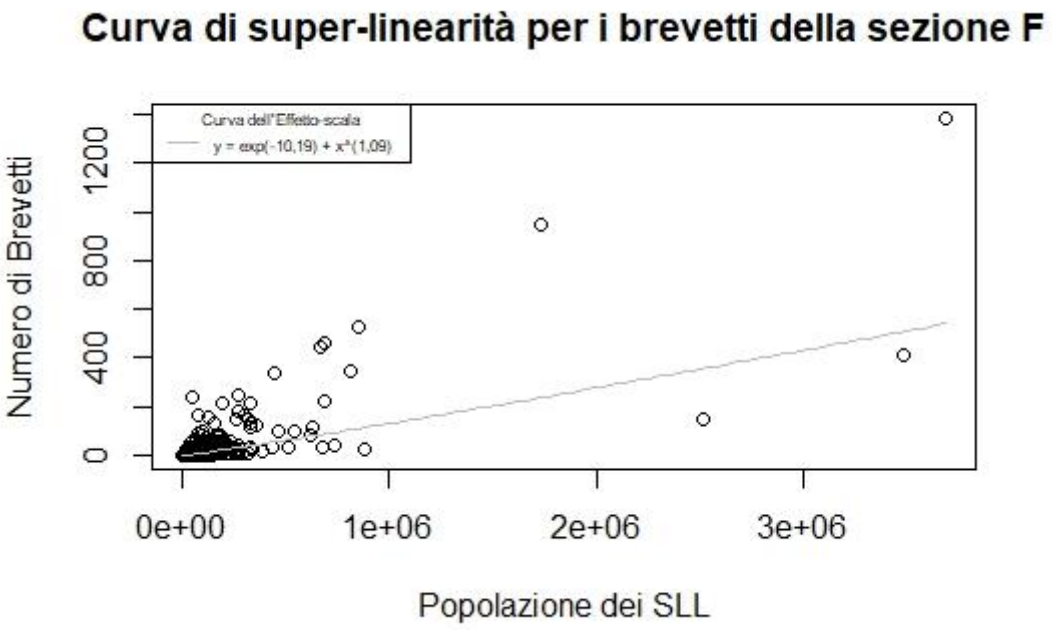

Grafico A.12: Grafico a Dispersione rappresentativo delle coppie "*Numero di Brevetti* – *Popolazione dei SLL*" per i brevetti appartenenti alla Sezione "F" dello Standard *IPC* con rappresentazione della curva super-lineare di regressione stimata (in grigio).

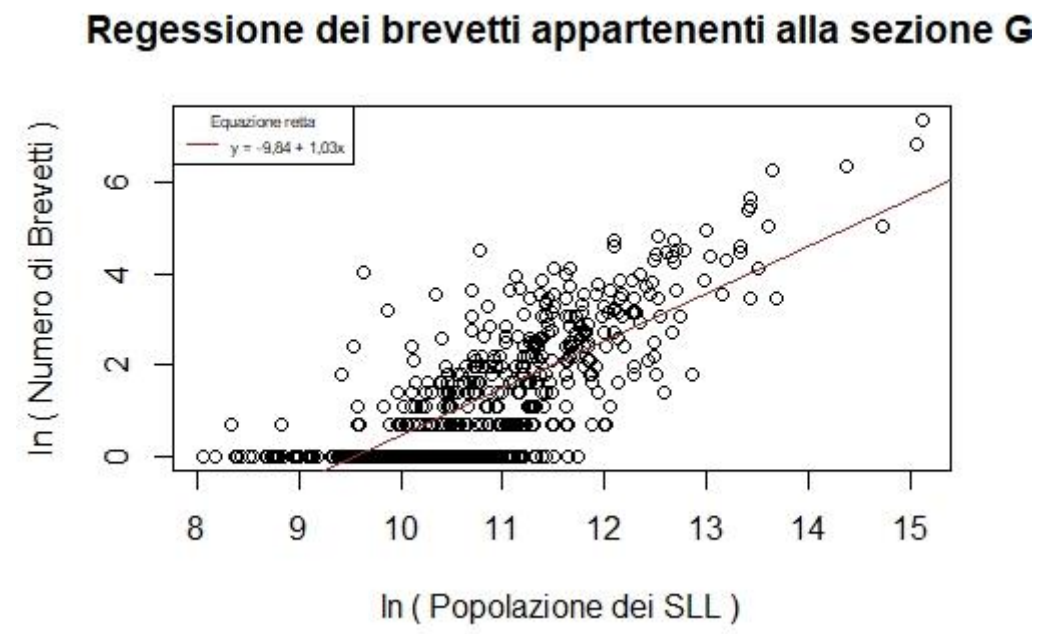

Grafico A.13: Grafico a Dispersione rappresentativo delle coppie "ln(*Numero di Brevetti*) – ln(*Popolazione dei SLL*)" per i brevetti appartenenti alla Sezione "G" dello Standard *IPC* con rappresentazione della retta di regressione stimata (in marrone).

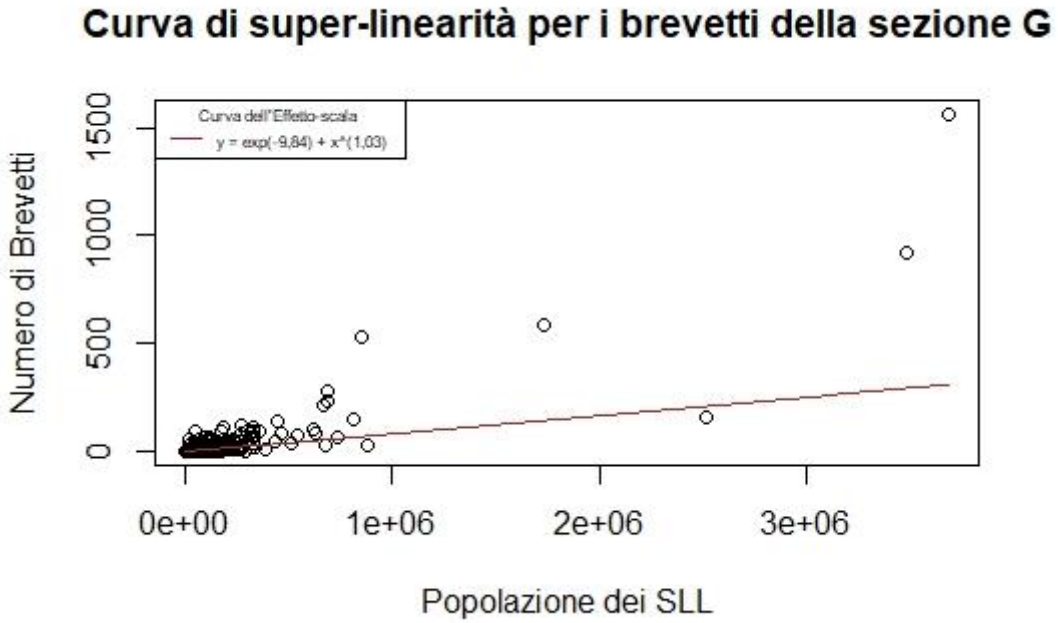

Grafico A.14: Grafico a Dispersione rappresentativo delle coppie "*Numero di Brevetti* – *Popolazione dei SLL*" per i brevetti appartenenti alla Sezione "G" dello Standard *IPC* con rappresentazione della curva super-lineare di regressione stimata (in marrone).

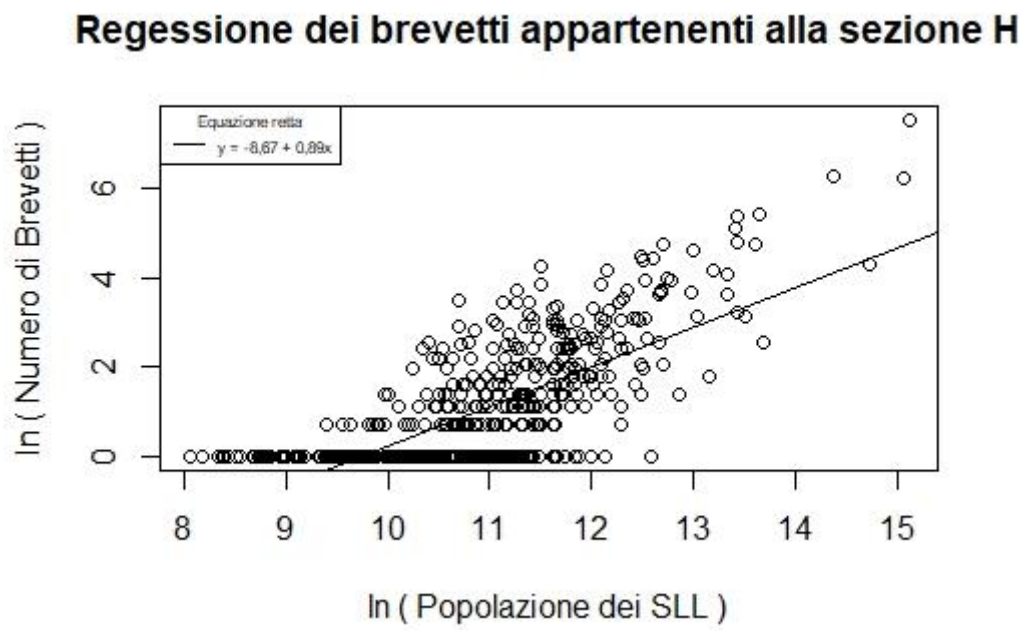

Grafico A.15: Grafico a Dispersione rappresentativo delle coppie "ln(*Numero di Brevetti*) – ln(*Popolazione dei SLL*)" per i brevetti appartenenti alla Sezione "H" dello Standard *IPC* con rappresentazione della retta di regressione stimata (in nero).

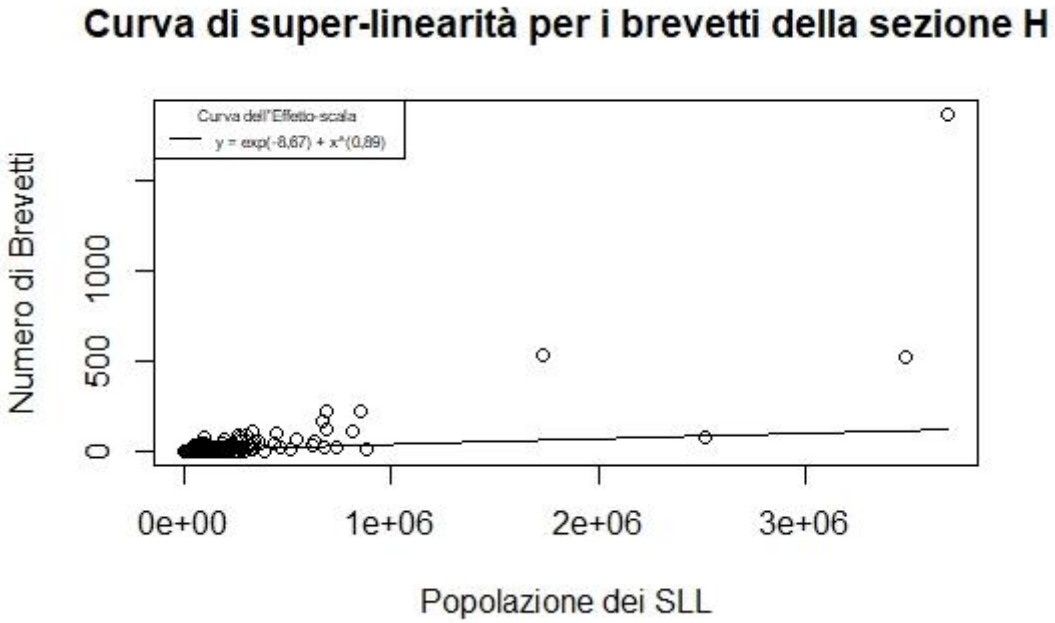

Grafico A.16: Grafico a Dispersione rappresentativo delle coppie "*Numero di Brevetti* – *Popolazione dei SLL*" per i brevetti appartenenti alla Sezione "H" dello Standard *IPC* con rappresentazione della curva super-lineare di regressione stimata (in nero).

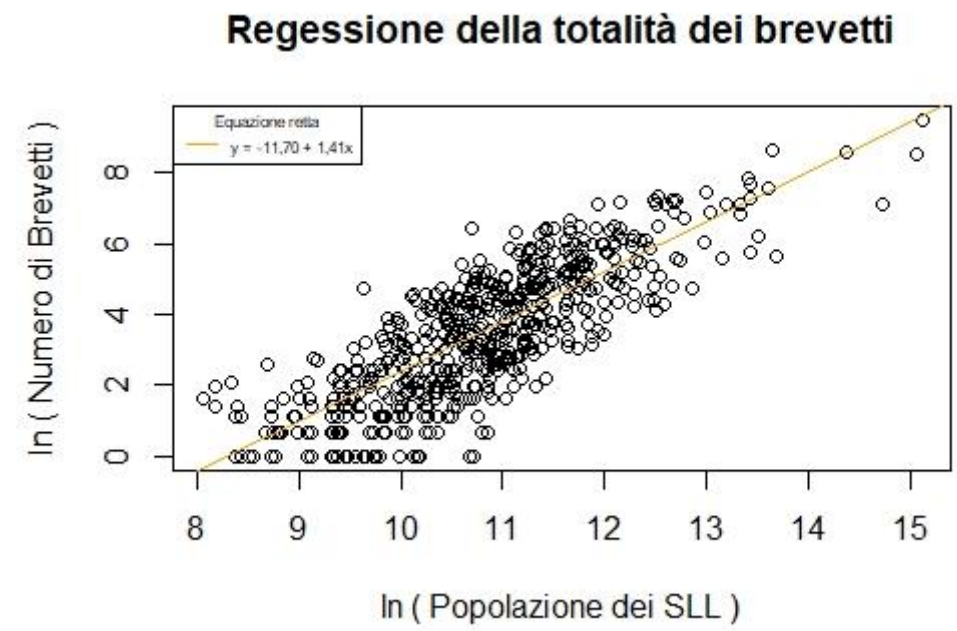

Grafico A.17: Grafico a Dispersione rappresentativo delle coppie "ln(*Numero di Brevetti*) – ln(*Popolazione dei SLL*)" per i brevetti appartenenti a tutte le sezioni dello Standard *IPC* con rappresentazione della retta di regressione stimata (in arancione).

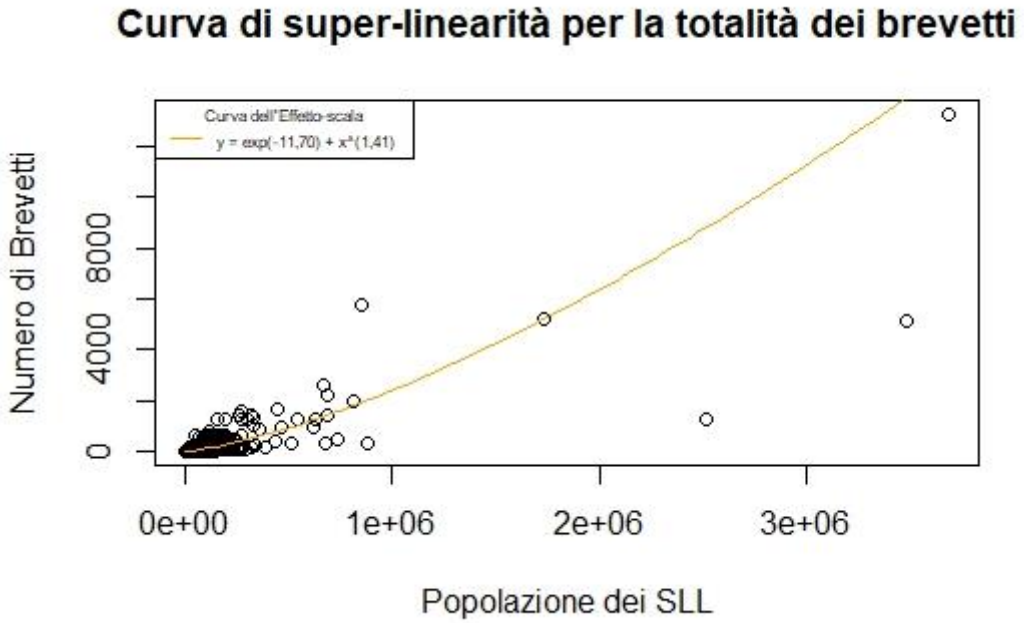

Grafico A.18: Grafico a Dispersione rappresentativo delle coppie "*Numero di Brevetti* – *Popolazione dei SLL*" per i brevetti appartenenti a tutte le sezioni dello Standard *IPC* con rappresentazione della curva super-lineare di regressione stimata (in arancione).

BIBLIOGRAFIA

The Product Space Conditions the Development of Nations - C.A.Hidalgo, B.Klinger, A.L.Barabási, R.Hausmann - www.sciencemag.org – 2007

The building blocks of economic complexity - Cesar A.Hidalgo and Ricardo Hausmann - Center for International Development and Harvard Kennedy School - 2009

Country diversification, product ubiquity, and economic divergence - Ricardo Hausmann, César A. Hidalgo - Center for International Development – 2010

The Network Structure of Economic Output - Ricardo Hausmann, César A. Hidalgo - - Journal of Economic Growth - 2011

Neighbors and the Evolution of the Comparative Advantage of Nations: Evidence of International Knowledge Diffusion? - Dany Bahar, Ricardo Hausmann, Cesar A. Hidalgo – Journal of International Economics - 2014

The Atlas of Economic Complexity - Ricardo Hausmann, Cesar A. Hidalgo, Sebastián Bustos, Michele Coscia, Sarah Chung, Juan Jimenez, Alexander Simoes, Muhammed A. Yıldırım, MIT Press - 2014

Complex Economic Activities Concentrate in Large Cities - Cristian Ignacio, Jara Figueroa, Mathieu Steijn, Cesar Hidalgo - SSRN Electronic Journal - 2018

The Principle of Relatedness - César A. Hidalgo, Pierre-Alexandre Balland, Ron Boschma, Mercedes Delgado, Maryann Feldman, Koen Frenken, Edward Glaeser, Canfei He, Dieter F. Kogler, Andrea Morrison, Frank Neffke, David Rigby, Scott Stern, Siqi Zheng, Shengjun Zhu - Springer Nature Switzerland AG - 2018

SITOGRAFIA

https://atlas.cid.harvard.edu/

https://www.wipo.int/classifications/ipc/en/

https://ec.europa.eu/eurostat/ramon/documents/IPC_NACE2_Version2_0_20150630.pdf