



**Politecnico
di Torino**

POLITECNICO DI TORINO

Collegio di Laurea Magistrale
in Ingegneria Gestionale

Tesi di Laurea Magistrale

Banda Larga Ultraveloce e Innovazione: Analisi dell'Impatto sulle Startup

Relatori:

Prof. Carlo Cambini

Prof. Lorien Sabatino

Candidato

Lorenzo Baldocchi

Anno Accademico 2020/2021



**Politecnico
di Torino**

Sommario

CAPITOLO 1 - INTRODUZIONE	1
1.1 Contesto	2
1.1.1 Breve descrizione della banda larga ultraveloce	2
1.1.2 Panoramica attuale	4
1.1.3 La rete di telecomunicazione italiana	7
1.2 Letteratura	13
1.2.1 Crescita economica	14
1.2.2 Definire innovazione	15
1.2.3 Guadagni generalizzati e nuove opportunità	17
1.2.4 Eterogeneità negli effetti e asset complementari	19
1.2.5 Legame con altre forme di conoscenza	21
1.2.6 Altre evidenze su ICT	23
1.2.7 L'importanza delle startup	24
1.2.8 Variabili significative per l'innovazione nelle imprese	26
RIFERIMENTI BIBLIOGRAFICI	29
CAPITOLO 2 - ANALISI DESCRITTIVA	33
2.1 Startup innovative	33
2.1.1 Suddivisione geografica	34
2.1.2 Suddivisione settoriali	40
2.1.3 Suddivisione per grado di urbanizzazione	42
2.1.4 Analisi crescita startup nel 2020	44
2.2 Brevetti	47
2.3 Diffusione UBB	49
2.3.1 Livello aggregato	50
2.3.2 Livello regionale	55
CAPITOLO 3 - MODELLO EMPIRICO	59
3.1 Ordinary Least Squares	59
3.1.1 Modello	59
3.1.2 Stima empirica	66
3.1.3 Event study	82
3.2 Instrumental Variables	88
3.2.1 Giustificazione teorica	88
3.2.2 Scelta dello strumento	90
3.2.3 Stima empirica	96

CONCLUSIONI.....	107
Sintesi dei risultati ottenuti.....	108

Capitolo 1 - Introduzione

Oggi giorno ci troviamo in una profonda crisi economica e molti stati, per rilanciarsi, hanno deciso di fare leva sul digitale investendo nella banda larga ultraveloce. Durante i periodi di difficoltà, le amministrazioni pubbliche sono ansiose di dare un impulso attivo agli investimenti nelle reti NGA, poiché il dispiegamento può aumentare la produttività nel lungo periodo e promuovere la ripresa economica nel breve (Ganuza & Vieceus [2011]).

Il progresso tecnico e la crescita di intere epoche sono conseguenza di nuove tecnologie: dopo il motore a vapore, i semiconduttori e i computer, le reti UBB hanno ora il potenziale di alterare la società e apportare miglioramenti che inducano guadagni di produttività generalizzati all'intera economia. Tutto ciò è socialmente importante se i benefici generati da una maggiore velocità superano i costi di implementazione. Più è alta la velocità che un paese vuole raggiungere, più sono alti i costi di investimento. Nonostante l'impatto economico positivo sia stato dimostrato, non è però scontato che sia socialmente desiderabile. Scopo di questa tesi è proprio studiare uno dei molti effetti della diffusione della banda larga ultraveloce: l'impatto sull'innovazione, misurando la creazione di nuove startup in Italia nell'arco temporale 2013-2019. Sebbene la letteratura abbia ampiamente trattato l'impatto della banda larga sulle principali variabili macro e microeconomiche, le ricerche diminuiscono se ci focalizziamo sulle nuove reti NGA e non si sono ancora affermate conclusioni empiriche relative all'impatto delle nuove reti sull'innovazione. Approfondire ulteriormente questa tematica sarà probabilmente oggetto di ricerche future poiché l'innovazione ha il potenziale di guidare l'economia e sostenere la competitività a lungo termine.

La tesi è strutturata come segue. Lo scopo del primo paragrafo dell'introduzione è fornire un approfondimento del contesto attuale, descrivendo le principali tecnologie UBB, la diffusione attuale delle reti e infine l'infrastruttura di telecomunicazione italiana. Nel secondo paragrafo si effettua una review della letteratura. L'oggetto dei contributi presi in esame ruota attorno alle reti UBB e l'innovazione. Oltre a questi, sono trattati anche contributi esistenti degli impatti della banda larga di base e le più generiche tecnologie ICT. Nel secondo capitolo si presenta un'analisi descrittiva dei dati riguardanti le startup innovative, le registrazioni brevettuali e il dispiegamento delle reti a banda larga ultraveloce in Italia. Infine, nel terzo e conclusivo capitolo, sono analizzate le strategie empiriche utilizzate e commentati i risultati ottenuti dal modello statistico.

1.1 Contesto

1.1.1 Breve descrizione della banda larga ultraveloce

Le reti a banda larga sono in grado di gestire contemporaneamente numerosi segnali separati: consentono una velocità di trasmissione nettamente superiore a quella supportata dai modem tradizionali mantenendo un flusso costante di informazioni pari o superiore a 1,5Mbit/s.

L'ampiezza di banda, utilizzata dai mezzi e sistemi di telecomunicazione, è maggiore rispetto alle reti precedenti in modo da trasmettere e ricevere simultaneamente un'elevata quantità di dati a una velocità superiore. La maggior dimensione esprime indirettamente la capacità di far viaggiare contemporaneamente più segnali utilizzando lo stesso cavo. La velocità di trasmissione decade con la distanza, a causa della degradazione del SNR¹ per effetto dell'attenuazione del segnale utile e dell'incremento del rumore introdotto dal canale stesso. Avere una banda di trasmissione più larga consente dunque una riduzione di questo problema in ricezione, ovvero una maggiore velocità di connessione. I canali utilizzati sono la fibra ottica, DSL (*digital subscriber line*), satellitare, power line, modem via cavo o senza cavo (*wireless*). A causa del crescente aumento della domanda, gli operatori di telecomunicazioni hanno dovuto aggiornare le loro reti in rame e iniziare a distribuire le cosiddette reti a banda ultralarga "*Next Generation Access*" (NGA), basate sulla tecnologia a fibra ottica. Le caratteristiche della fibra, bassa perdita e ampia larghezza di banda, la rendono la candidata ideale per soddisfare le sfide di capacità per ora e per il futuro. Nel caso in cui tale connessione riesca a superare i 30Mbit/s, fino ad arrivare ad 1 Gbit/s, siamo in presenza della cosiddetta "banda ultra-larga": questa tecnologia consente un ulteriore aumento della capacità e l'adozione di servizi e applicazioni completamente nuovi sul lato della domanda. Un cavo unico può contenere fino a 7 fibre ottiche, ma spesso due di queste vengono sostituite da due fili di materiale elastomerico, il cui scopo è di irrobustirlo meccanicamente. Lungo la tratta, ogni tot km, sono presenti amplificatori e rigeneratori per ovviare ai problemi di attenuazione e dispersione del segnale.

Fondamentalmente, tre architetture possono essere distribuite per la distribuzione della fibra:

- Architettura punto-punto: le singole fibre vanno dalla centrale locale ad ogni casa. Anche conosciuta come Fiber To The Home (FTTH), è caratterizzato dalla presenza di moltissime fibre, ciò comporta alti costi di installazione ma fornisce la massima capacità e la maggiore flessibilità per aggiornare i servizi per i clienti. Le velocità di trasmissione raggiungibili sono pari a 1 Gbit/s.
- Architettura a stella attiva: una singola fibra porta tutto il traffico a un nodo attivo (cabinet) in prossimità degli utenti finali, il collegamento con le abitazioni è poi effettuato utilizzando la linea telefonica in rame. È necessario solo un singolo feeder e un certo numero di brevi fibre verso gli utenti finali, perciò i costi sono bassi. D'altro canto, il nodo attivo ha bisogno di alimentazione e manutenzione. La velocità complessiva è elevata se la tratta secondaria è ridotta. Le tecnologie DSL hanno infatti un miglior funzionamento sulle brevi distanze ed il calo di capacità sarà limitato. Il nodo attivo può essere situato in tre posizioni diverse.

¹ Signal to Noise Ratio o (S/N nell'uso italiano): grandezza numerica che mette in relazione la potenza del segnale utile rispetto a quella del rumore in un qualsiasi sistema di acquisizione, elaborazione o trasmissione dell'informazione. In fase di progetto di un sistema di telecomunicazioni, il primo obiettivo è far pervenire al ricevitore un SNR sufficientemente elevato.

- FTTN (Fiber To The Node), il collegamento in fibra ottica arriva a una cabina esterna, distante diversi chilometri dalla sede del cliente.
- FTTC (Fiber To The Cabinet), il collegamento arriva a una cabina esterna molto vicina alla sede dell'utente o al successivo armadio distributore, tipicamente entro 300 metri.
- FTTB (Fiber To The Building), il collegamento raggiunge il limite del palazzo o seminterrato dell'edificio del cliente e l'ultimo tratto sarà coperto dai cavi in rame o da una connessione wireless.
- Architettura a stella passiva: il nodo è sostituito da un divisore di potenza ottico passivo che alimenta la ramificazione agli utenti finali. La natura completamente passiva dell'impianto esterno evita i costi di alimentazione e manutenzione che si verificherebbero con apparecchiature attive sul campo. Questa topologia è quindi diventata molto popolare ed è ampiamente conosciuta come Passive Optical Network (PON).

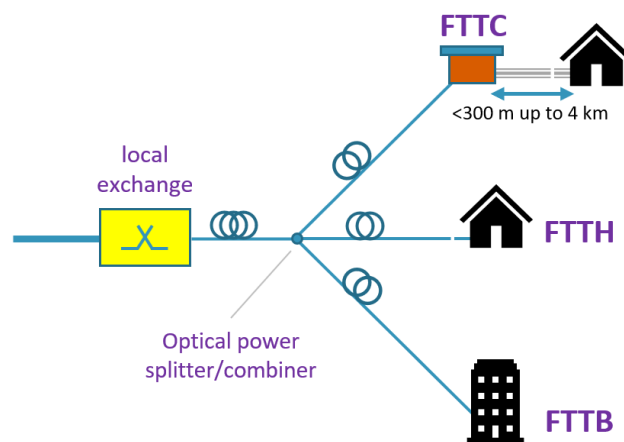


Figura 1-1 Architettura a stella passiva

Di seguito nella [Figura 1-2](#), sono rappresentate le quattro tecnologie più comunemente utilizzate con le relative proporzioni tra cavi in fibra ottica e in rame.

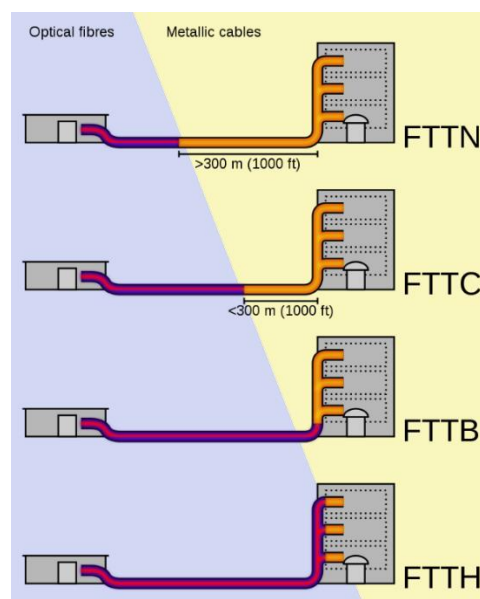


FIGURA 1-2 TECNOLOGIE A BANDA LARGA MAGGIORMENTE UTILIZZATE

1.1.2 Panoramica attuale

Sebbene l'Italia sia la terza economia dell'Unione Europea, negli ultimi venti anni è stata caratterizzata dalla stagnazione della crescita della produttività, ben prima quindi della crisi finanziaria del 2008 e della pandemia COVID-19. Pellegrino e Zingales hanno identificato la malattia dell'economia italiana nell'incapacità delle imprese di trarre pieno vantaggio dalle nuove tecnologie digitali. L'ICT rappresenta infatti una fonte vitale di competitività e oggi, nel pieno della rivoluzione digitale, l'Italia non può farsi trovare impreparata una seconda volta. L'accesso a una connessione a banda larga ultraveloce e affidabile è fondamentale poiché i principali servizi sociali ed economici sono forniti online. La trasformazione digitale apre nuove opportunità e stimola lo sviluppo di tecnologie nuove e affidabili. Un'infrastruttura moderna e robusta fornisce la copertura necessaria per beneficiare di tutto ciò. Per di più la crisi pandemica ha provocato un significativo aumento di domanda di banda e alzato l'asticella della velocità di trasmissione minima che i consumatori sono disposti ad accettare.

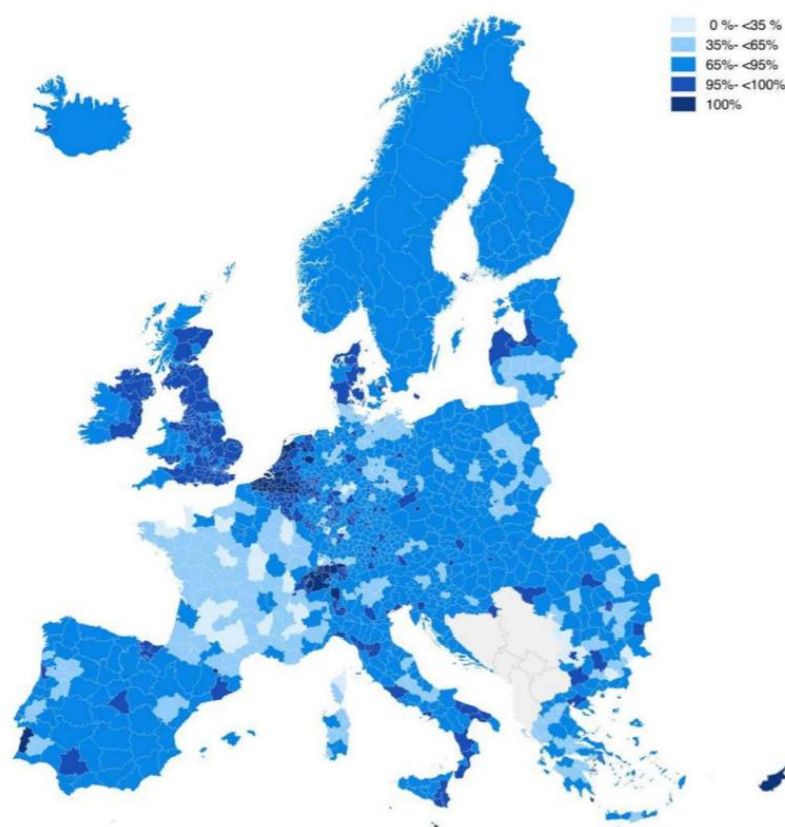
Dopo la caduta dei mercati, la Commissione europea è corsa ai ripari approvando un nuovo strumento per la ripresa, il Recovery Fund. Nel complesso la quota italiana è di circa 209 miliardi ripartiti (81,4 miliardi in contributi a fondo perduto e 127,4 miliardi in prestiti). Tale strumento rappresenta una vera e propria sfida ed una concreta opportunità per rilanciare l'economia del nostro paese. 40,8 miliardi sono stati delineati per gli investimenti nel digitale. Con tali risorse, si punta all'acquisto di beni digitali da parte di 60 milioni di imprese all'anno entro il 2026 (+20% rispetto a oggi) e a coprire con velocità di 1 gigabit/secondo 2,6 milioni di unità abitative, cioè il 30% di quelle ancora in digital divide². In passato, la diffusione di internet tra la popolazione ha portato a un aumento delle disuguaglianze, migliorando le prospettive di coloro che sono già in posizione privilegiata e sfavorendo le opportunità di avanzamento ai meno privilegiati. Le determinanti principali di questo problema, per cui oggi si cerca una soluzione, sono legate alle disuguaglianze che emergono a partire dall'accesso fino all'uso di internet.

Negli scorsi anni, l'Europa nella sua Agenda Digitale (DAE), aveva già sottolineato l'importanza strategica delle infrastrutture e dei servizi a banda larga per lo sviluppo economico e fissato ambiziosi obiettivi di copertura e adozione. Il DAE "cerca di garantire che, entro il 2020, (i) tutti gli europei abbiano accesso a velocità di rete molto più elevate, superiori a 30Mbit/s e (ii) che il 50% o più delle famiglie europee siano abbonate a connessioni Internet superiori a 100Mbit/s".

La classe politica si è da tempo resa conto dell'importanza della connettività all'interno della società: non solo permette al lavoro di continuare ma offre anche la possibilità di comunicare con gli altri e usufruire appieno dei prodotti e servizi offerti dai mercati moderni. Anche durante l'attuale pandemia, l'Europa ha lanciato diverse misure nel settore del digitale con lo scopo di minimizzare il contagio, sostenere il sistema sanitario e rafforzare l'intera infrastruttura di telecomunicazioni (ad esempio, attraverso il monitoraggio del traffico Internet in ogni paese in modo da essere in grado di bilanciare i problemi di capacità). La fornitura di risorse educative online e i servizi pubblici digitali sono poi stati sviluppati col fine di promuovere l'inclusione digitale. Allo stesso tempo, è stato accelerato il sostegno alla digitalizzazione delle imprese incentivando il commercio elettronico, lo smart working e la formazione online. Anche la cybersecurity e la lotta contro

² Il "digital divide" indica il problema di accesso ai mezzi di informazione e comunicazione da parte di determinate aree geografiche o fasce di popolazione. La disparità genera un'esclusione dai vantaggi della società digitale con danni socioeconomici e culturali per chi ne è colpito.

le fake news è stata una priorità. I rapporti DESI 2020 dell'Unione Europea, basati su dati raccolti nel 2019, valutano lo stato della società digitale prima della pandemia. La banda larga risulta disponibile per tutte le famiglie, se si considerano tutte le principali tecnologie. L'accesso è fornito principalmente da tecnologie fisse, che sono rimaste stabili al 97%. Tra queste, xDSL ha la maggiore copertura (91%). La diffusione della banda larga delle zone rurali rimane un punto impegnativo, poiché il 10% delle famiglie non è coperto da alcuna rete fissa e il 41% non ha accesso ad alcuna tecnologia NGA. La disponibilità della banda mobile è aumentata di 2 punti percentuali, anche se questa è utilizzata come una tecnologia complementare, piuttosto che sostituta delle tecnologie fisse. La diffusione globale della banda larga è aumentata solo marginalmente dal 2011, dal 95% al 97%. La copertura rurale è migliorata dall'80% nel 2011 al 90% nel 2019. Le tecnologie di accesso di nuova generazione (NGA), in grado di fornire velocità di download di almeno 30Mbit/s, hanno continuato a crescere, raggiungendo l'86% nel 2019 (Figura 1-3), rispetto al 48% del 2011. La copertura NGA è migliorata significativamente nelle aree rurali, con un aumento di 50 punti percentuali in 8 anni: nel 2011, si attestava al 9% delle famiglie, mentre nel 2019 al 59%.



**FIGURA 1-3 PERCENTUALE DI FAMIGLIE CON ACCESSO A RETI NGA AL TERMINE DEL 2019 –
FONTE: DESI 2020**

La statistica successiva (Figura 1-4) mostra il numero di connessioni fisse a banda larga dal 2011 al 2019, suddivise per tecnologia. L'ADSL è ancora la tecnologia leader con 58 milioni di connessioni nel 2019, anche se la quota dal 2011 è diminuita dal 71,72% al 29,74%. L'FTTP/B ha avuto una crescita costante (annualmente tra il 18 e 30%) passando da 7 a 41 milioni dell'intervallo di tempo considerato. Una crescita simile si è verificata per

la VDSL che è diventata la seconda tecnologia più utilizzata raggiungendo 49 milioni di connessioni. L'analisi della diffusione della banda larga ultraveloce in Italia sarà affrontata nel capitolo 2.

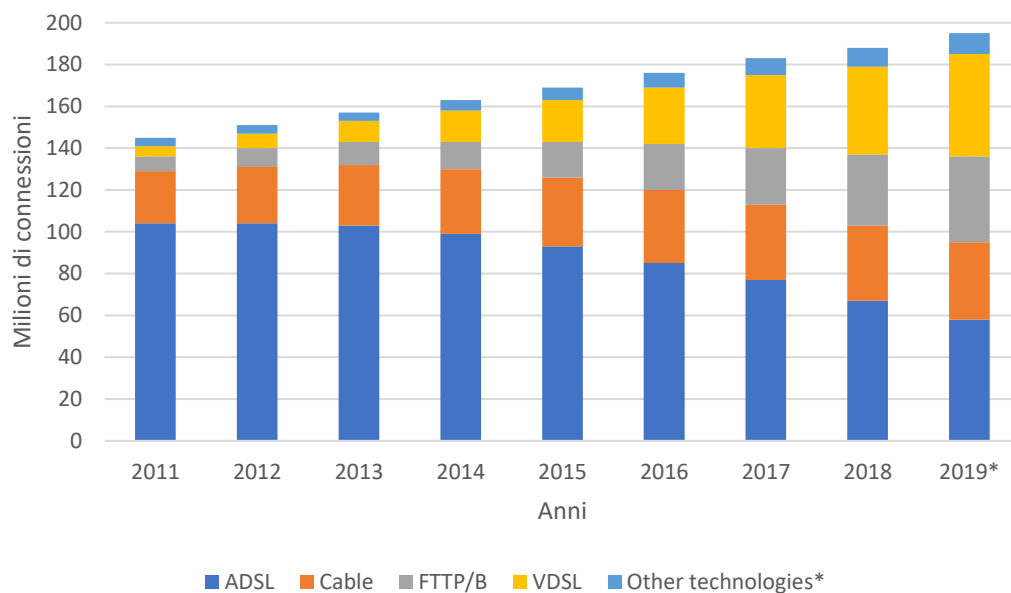


FIGURA 1-4 CONNESSIONI FISSE IN EUROPA, SUDDIVISE PER TECNOLOGIA – FONTE: IDATE DIGIWORLD

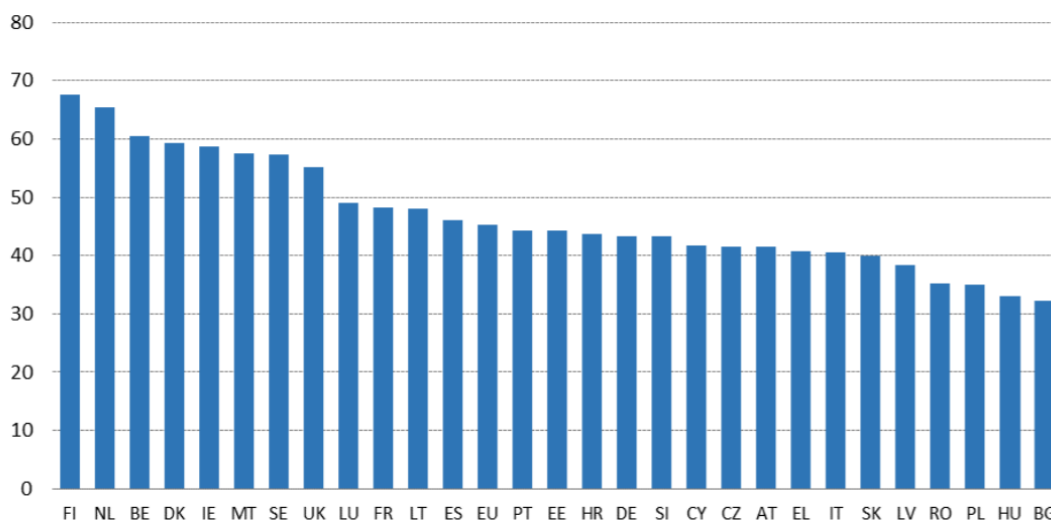


FIGURA 1-5 INTEGRAZIONE TECNOLOGIE DIGITALI NEGLI STATI MEMBRI DELL'UNIONE EUROPEA – FONTE: DESI 2020

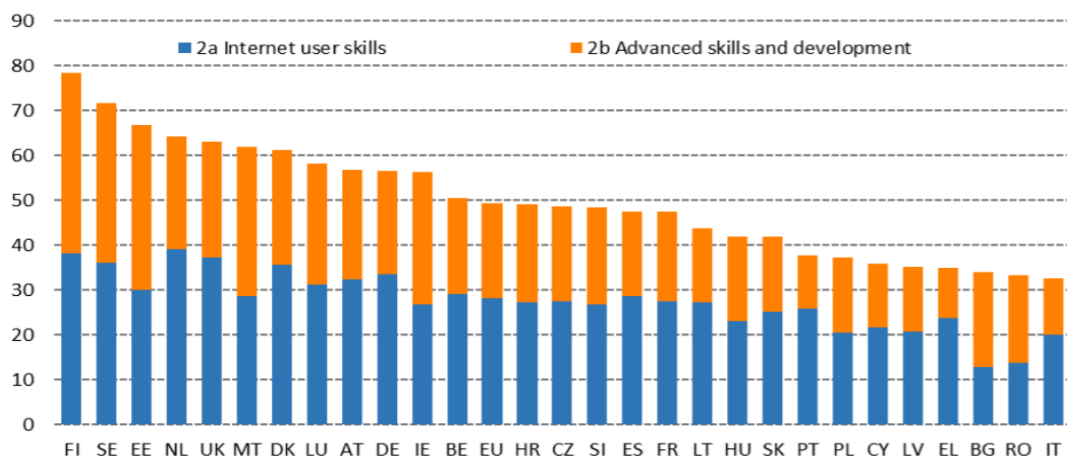


FIGURA 1-6 CAPITALE UMANO NEGLI STATI MEMBRI DELL'UE – FONTE: DESI 2020

Si è verificata un'eterogenea integrazione delle tecnologie digitali nelle attività aziendali (cioè la condivisione elettronica delle informazioni, dei big data, dei social media, del cloud e del commercio online) tra gli stati membri dell'Unione Europea. L'Italia si colloca al 23° posto su 28 paesi nell'indice DESI (Figura 1-5). Le principali cause sono l'alto livello di frammentazione dell'industria insieme al basso capitale umano digitale (Figura 1-6), in cui l'Italia si colloca addirittura all'ultimo posto. Le competenze ICT di base e avanzate sostengono le società e ciò rappresenta un drammatico freno alla crescita dell'integrazione digitale italiana poiché c'è una stretta complementarità tra tecnologie dell'informazione, capitale umano e cambiamento organizzativo (Giuri et al. [2008]).

1.1.3 La rete di telecomunicazione italiana

Lo sviluppo delle reti a banda larga ultraveloce in Italia è iniziato nel 2015. Il Governo, con l'obiettivo di contribuire a ridurre il gap infrastrutturale e di mercato esistente, creare condizioni più favorevoli allo sviluppo, ha approvato il quadro nazionale di riferimento per le iniziative pubbliche a sostegno della diffusione delle reti. La Strategia, al fine di soddisfare gli obiettivi fissati dall'Agenda Digitale Europea per il 2020, prevedeva la copertura dell'85% della popolazione con infrastrutture in grado di veicolare servizi a velocità pari o superiori a 100Mbit/s, garantendo al contempo alla restante quota di popolazione l'accesso ad internet ad almeno 30Mbit/s. La Strategia è inoltre coerente con gli obiettivi del 2025 espressi dalla Gigabit Society - Comunicazione COM (2016) in quanto diffonde la fibra ottica in modo capillare sul territorio e rende disponibili connessioni FTTH alle sedi della PA (scuole, sedi della sanità etc.) e alle aree produttive. La situazione, al termine del 2019, è la seguente: circa il 55% dei comuni italiani ha accesso a servizi UBB con velocità di connessione superiore a 30Mbit/s, pari a circa il 90% della popolazione italiana. Una piccola frazione di comuni (6%) pari al 37% della popolazione ha accesso alle più avanzate connessioni Fiber-to-the-Home, che consentono una velocità massima di 1 Gigabit al secondo (Gbps). Le connessioni ADSL sono nella loro fase matura, in quanto quasi tutta la popolazione ha accesso a una connessione ADSL di base con velocità massima di 7 (Mbit/s).

La principale differenza tra le due tecnologie consiste nella struttura dell'ultimo miglio. Le connessioni ADSL utilizzano la preesistente linea telefonica in rame per collegare i CO³ ai consumatori finali. Poiché la lunghezza di tale linea influisce significativamente sulle prestazioni della connessione, la letteratura ha spesso l'ultimo miglio come una variabile esogena che influisce sulla diffusione della banda larga ADSL. La principale innovazione imposta dall'UBB risiede proprio nella natura di questa tratta. Il collegamento avviene con almeno una porzione di fibra ottica, il che implica una velocità di connessione significativamente maggiore. In Italia, sono utilizzate quattro possibili configurazioni, a seconda della proporzione di fibra ottica utilizzata nell'ultimo miglio (Figura 1-2). La lunghezza della porzione dell'ultimo miglio è inversamente proporzionale alle prestazioni di connessione in tutte le tecnologie. Quindi minore è la distanza tra il comune e l'OLT più vicino, maggiore è la copertura UBB dell'area comunale.

Il primo tassello della strategia italiana, la cui attuazione è stata affidata al Ministero dello Sviluppo Economico, è rappresentato dal cosiddetto “Piano Aree Bianche” e riguarda le aree a fallimento di mercato presenti sull'intero territorio nazionale. L'intervento pubblico è stato ritenuto necessario per correggere le disuguaglianze sociali e geografiche generate dall'assenza di iniziativa privata da parte degli operatori e consentire una maggior coesione sociale e territoriale. Si tratta di interventi realizzati secondo il modello diretto: l'infrastruttura pubblica viene messa a disposizione degli operatori del mercato TLC⁴ per l'offerta dei servizi agli utenti finali, mediante la cessione di diritti d'uso su quote delle infrastrutture o sulle singole fibre ottiche in modalità IRU (Indefeasible Right of Use). Questo è un tipo di accordo contrattuale permanente di locazione di telecomunicazioni in cui il cliente acquista il diritto di utilizzare una certa quantità di capacità, per un determinato numero di anni. I contratti IRU sono quasi sempre a lungo termine, di solito durano da 20 a 30 anni. Un proprietario può utilizzare incondizionatamente ed esclusivamente la capacità pertinente della rete. L'acquirente si obbliga a pagare una parte dei costi operativi e i costi di manutenzione del cavo. Nelle aree bianche, la rete di proprietà dello Stato sarà gestita da Open Fiber in concessione per 20 anni. Principalmente in Italia ci sono due reti distinte che forniscono servizi UBB ai consumatori locali. La prima è di proprietà dell'incumbent delle telecomunicazioni Telecom Italia Mobile (TIM), e fa leva sulle preesistenti infrastrutture di telecomunicazione utilizzate per la telefonia vocale e le connessioni ADSL. TIM ha iniziato a investire in infrastrutture UBB in tutta Italia dal 2015, e copre la stragrande maggioranza dei comuni con un mix di connessioni FTTH e fibra ottica-rame (FTTx). A livello nazionale, la trasmissione dei dati avviene attraverso l'Optical Packet Backbone (OPB), che costituisce il livello a monte dell'infrastruttura conosciuta col nome di “Kaleidon” (Figura 1-7). L'OPB “tradizionale” è stato sviluppato tra il 2001 e il 2005, era composto da 32 nodi che gestivano il traffico a livello nazionale per fornire connessioni internet ADSL in tutta Italia (Figura 1-10, Figura 1-12); essi rappresentano i nodi principali dell'infrastruttura nazionale e hanno lo scopo di reindirizzare il traffico verso i central office (CO). All'interno di quest'ultimi devono essere installati i terminali di linea ottica (OLT), dispositivi specifici per convertire i segnali ottici. Quindi, solo i CO dotati di OLT servono effettivamente come endpoint dell'operatore di telecomunicazione (Figura 1-9, Figura 1-12).

³ I central office (CO) rappresentano il livello intermedio dell'infrastruttura di telecomunicazioni. Il loro scopo è convertire il segnale ottico e reindirizzarlo ai consumatori finali. La distanza con l'ubicazione del cittadino è definita “ultimo miglio”. Nella pagina seguente è descritta con maggior precisione l'infrastruttura nazionale e il ruolo specifico dei CO.

⁴ Il mercato TLC indica il settore delle telecomunicazioni. Può essere separato in due grandi aree di business: da un lato i servizi di telecomunicazione, offerti dagli operatori; dall'altro gli investimenti in apparecchiature e sistemi, che comprendono l'acquisto di terminali da parte dei consumatori finali.

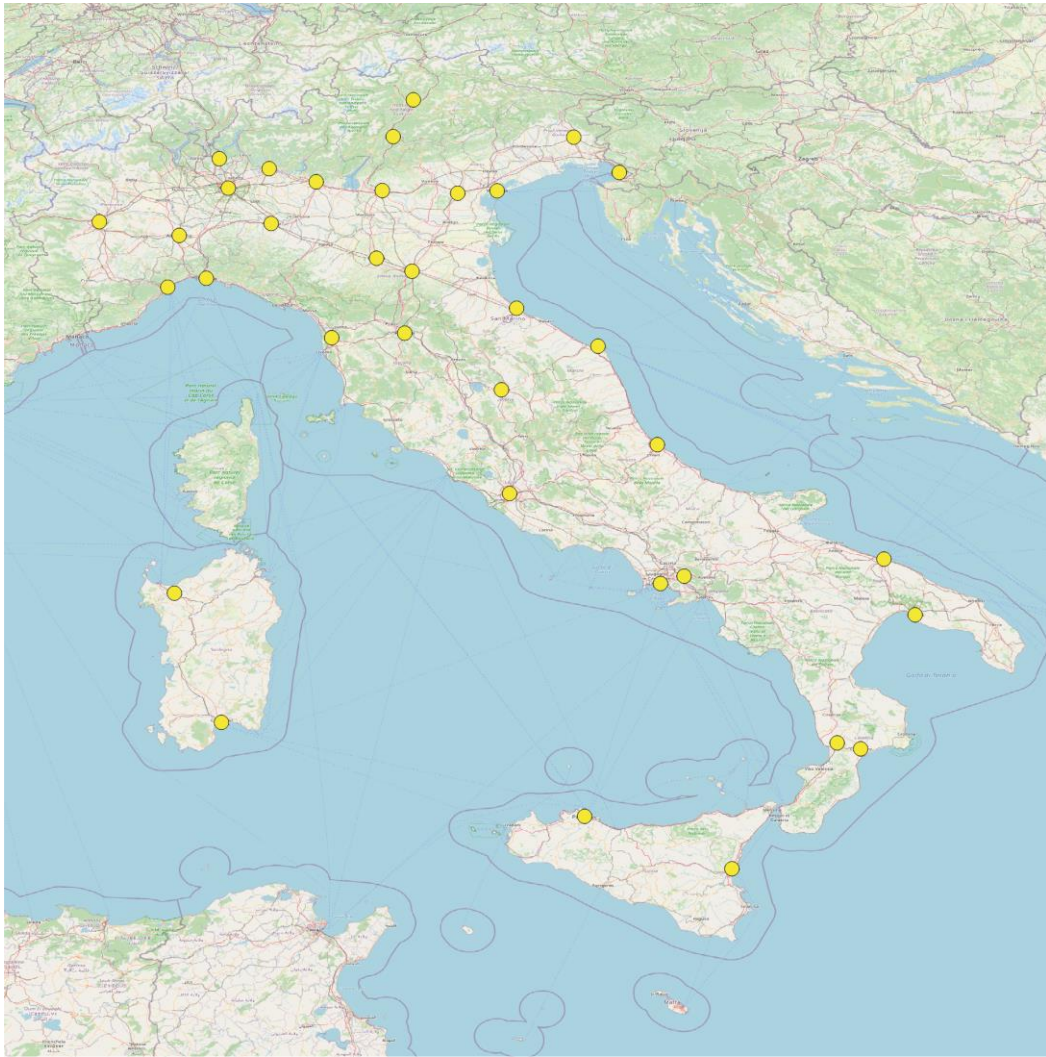


FIGURA 1-7 POSIZIONE OPB – FONTE TIM NETBOOK 2018

Il piano di sviluppo della nuova dorsale Kaleidon è stato approvato nel 2012 e concluso nel dicembre 2014, proprio prima dell'introduzione dell'UBB. L'aggiornamento dell'infrastruttura si è reso necessario per gestire la crescente quantità di dati di traffico e ha comportato il miglioramento dei 32 nodi "tradizionali" dell'OPB e l'introduzione di altri cinque nuovi nodi. Nonostante la serie di processi di innovazione che hanno portato all'attuale configurazione, tutti i nodi nazionali (sia OPB che CO) derivano da strutture preesistenti del sistema di telefonia vocale.

La seconda rete di telecomunicazioni è di proprietà di Open Fiber, operatore appartenente alla Cassa Depositi e Prestiti e ad Enel. Open Fiber è entrata nel mercato nel 2017 e da allora, ha investito in connessioni FTTH principalmente nelle grandi città e nei piccoli comuni vicini alle aree metropolitane; si è anche aggiudicata tre gare pubbliche indette dal Ministero dello Sviluppo Economico italiano per il dispiegamento di un'infrastruttura in fibra ottica nei piccoli comuni di tutta Italia. Open Fiber ha iniziato a distribuire le reti solo all'inizio del 2019 e alla fine dello stesso anno ha coperto 265 piccoli comuni, 112 dei quali sono anche (totalmente o parzialmente) coperti da TIM.

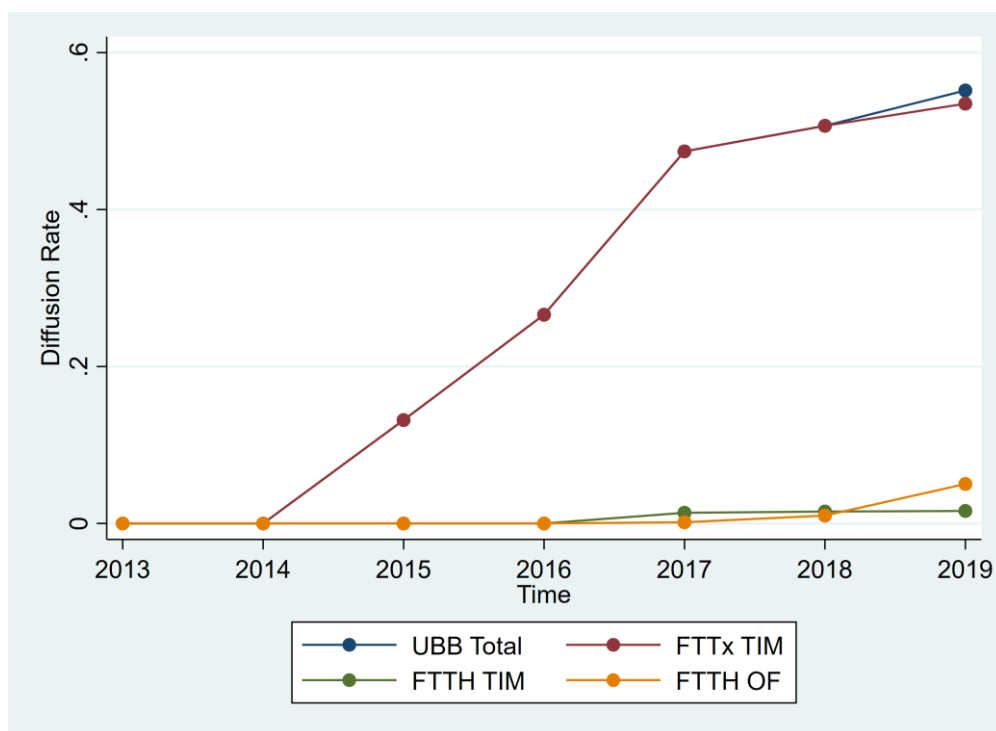


FIGURA 1-8 DIFFUSIONE TIM E OPEN FIBER

La [Figura 1-8](#) mostra il tasso di diffusione delle diverse tecnologie offerte dai due operatori. Il co-investimento può essere uno strumento chiave per mantenere la competizione infrastrutturale riducendo i rischi di duplicazione della rete e ottimizzando gli investimenti. A supporto di questa tesi, J. Bouckaert et al. [2010] hanno dimostrato che la concorrenza tra piattaforme incoraggia la penetrazione della banda larga, mentre la concorrenza intra-piattaforma basata sui servizi costituisce un ostacolo alla penetrazione. Promuovere la competizione intra-piattaforma potrebbe rappresentare un trampolino di lancio per i nuovi entranti ma ciò potrebbe addirittura dare incentivi negativi agli investimenti.

Al Piano Aree Bianche farà seguito una seconda fase della strategia, con interventi destinati alle aree grigie (caratterizzate dalla presenza nei prossimi tre anni di una sola rete UBB) e all'incentivazione della domanda di banda, attualmente in via di definizione e da sottoporre ad approvazione della Commissione Europea.

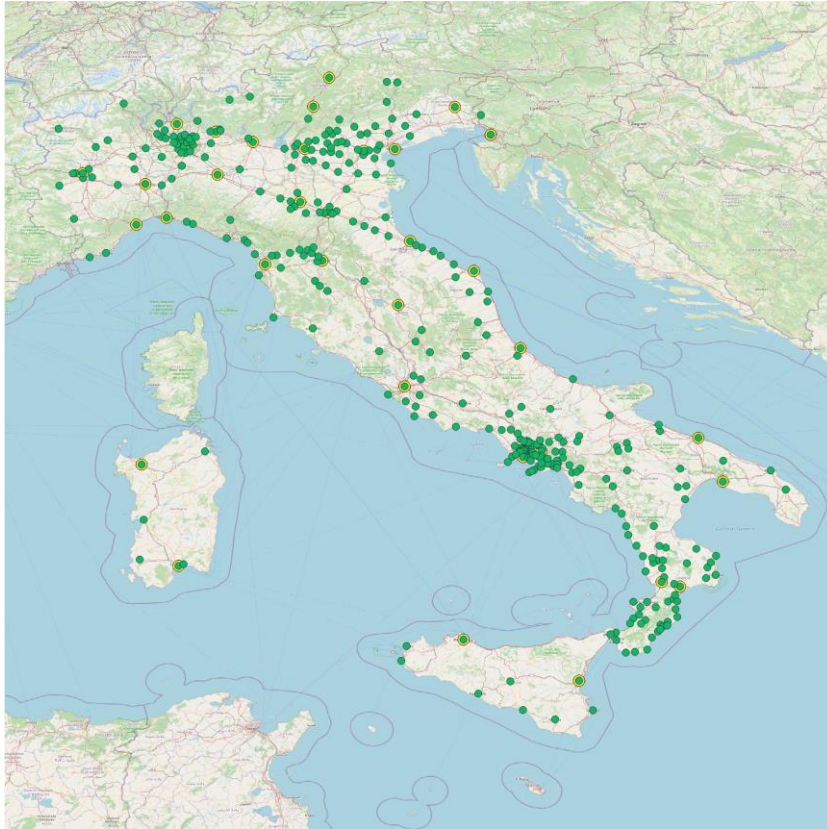


FIGURA 1-9 2015: OLT LOCATIONS

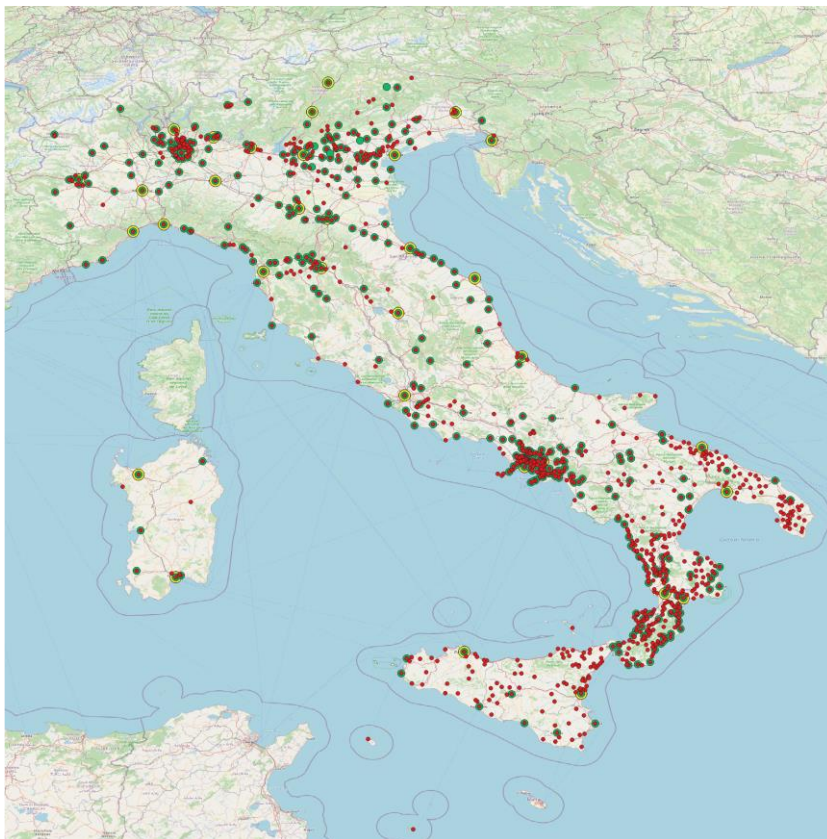


FIGURA 1-10 2015: UBB ACCESS



FIGURA 1-11 2019: OLT LOCATIONS

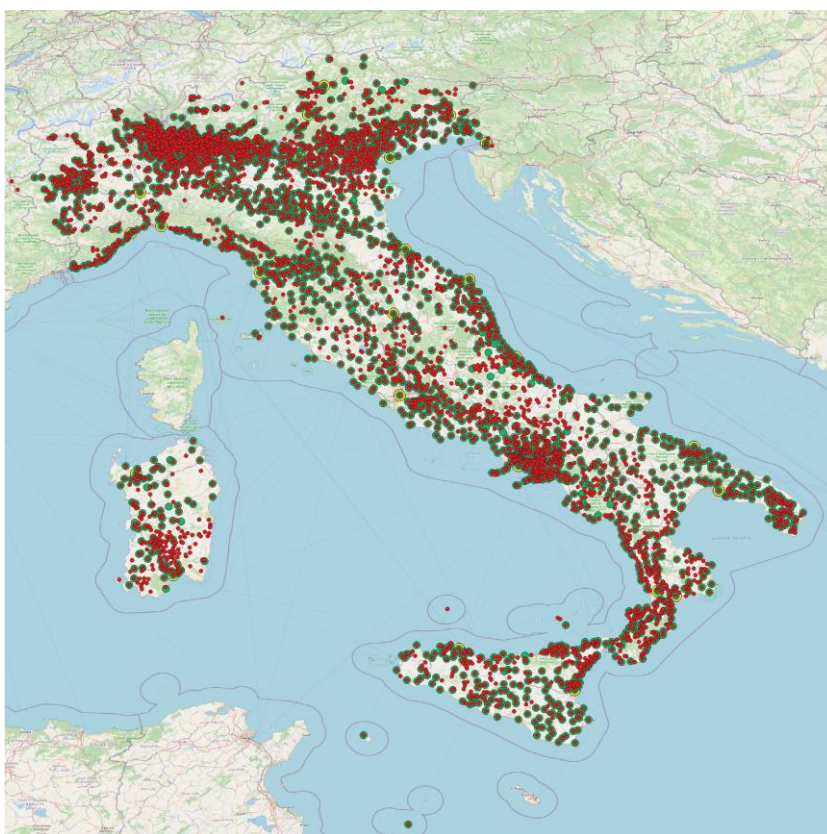


FIGURA 1-12 2019: UBB ACCESS

1.2 Letteratura

Le tecnologie ICT e l'innovazione sono strettamente interconnesse. Le nuove tecnologie possono aiutare le economie, le industrie e le imprese a crescere e ad avere successo, e il gran numero di iniziative governative per promuovere la diffusione della banda larga ultraveloce testimonia l'aspettativa riguardante l'impatto sociale indiretto che potrà anche andare oltre gli effetti economici. A causa dell'aumento della domanda, si è verificato un cambiamento di paradigma tecnologico: gli operatori di telecomunicazioni hanno avuto la necessità di aggiornare l'infrastruttura, iniziando a distribuire le cosiddette reti a banda larga "Next Generation Access" (NGA) basate sulla fibra ottica. Gli effetti immediati sono stati un massiccio aumento della capacità di banda e l'adozione di servizi completamente nuovi sul lato dei consumatori.

La diffusione di queste nuove reti è un'importante abilitatrice per lo sviluppo della società moderna, fondamentale per aumentare la produttività e la crescita nel lungo periodo. Poiché i benefici sono strettamente legati alle esternalità positive e l'adozione spontanea della tecnologia si sarebbe estesa solo fino al punto in cui il guadagno marginale uguaglia il costo di adozione (Battisti e Stoneman [2005]), è stato necessario l'intervento dei governi per aumentare il livello di equilibrio della diffusione. Con questo processo ("diffusione indotta"), molte policy hanno incentivato la velocità e/o il livello totale del dispiegamento, internalizzando direttamente o indirettamente le esternalità positive e/o negative (Diaz-Rainey [2009]). I costi derivanti sono sostanziali, poiché è stata necessaria un'intera infrastruttura che è in parte o totalmente basata sulla fibra. Alcuni studi evidenziano la velocità di reintegro

delle risorse finanziarie, come *The European House – Ambrosetti* che stima il contributo incrementale al PIL italiano derivante dal pieno dispiegamento delle reti in un

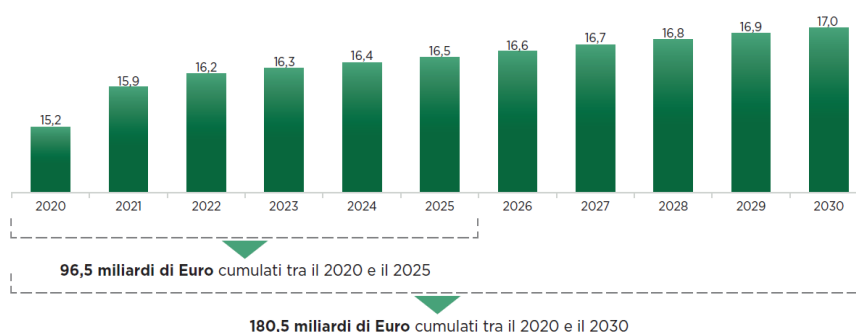


FIGURA 1-13 IMPATTO PREVISTO SUL PIL ITALIANO

aumento pari a 180,5 miliardi di € nei prossimi 10 anni (Figura 1-13). Canzian et al. [2019] hanno valutato l'impatto di una politica locale del 2012 volta alla diffusione della tecnologia ADSL2+ a banda larga ad alta velocità in aree rurali nella Provincia di Trento. È stata scelta questa provincia poiché aveva il più alto utilizzo di banda larga in Italia pari 57,4% delle famiglie aveva accesso a tale tecnologia, che era ben al di sopra della media nazionale del 48,6%. Gli effetti iniziarono a manifestarsi circa un anno dopo l'attivazione e al termine del 2014, si verificò complessivamente un aumento di circa il 14,8% dei ricavi e del 9,1% della produttività dei fattori.

Sebbene l'impatto positivo sia fuori discussione, investire in connessioni NGA è socialmente importante solo se i benefici generati da una maggiore velocità superano i costi di implementazione. Ancora oggi sono limitati gli studi riguardanti i benefici incrementali delle tecnologie ultraveloci e alcuni effetti sono stati analizzati solo marginalmente. Nello specifico, l'impatto della banda larga ultraveloce sull'innovazione è scarsamente trattato, è ancora un'area grigia che pochi studi empirici hanno cercato di chiarire. Per questo, nella

rassegna della letteratura saranno commentati diversi effetti su variabili economiche, in particolare l'innovazione e per cercare di approfondire dettagliatamente lo stato dell'arte, saranno presi in esame risultati ottenuti analizzando anche altre tecnologie digitali. Tuttavia, è importante ricordare che trattare l'ICT come una singola variabile è fuorviante, poiché gli effetti sono intricati ed è soprattutto l'uso di infrastrutture più avanzate a portare a effetti positivi sulla crescita che attendiamo dalle nuove tecnologie pervasive.

Il paragrafo è suddiviso in otto sezioni. Nella prima parte sono trattati l'impatto della diffusione della banda larga ultraveloce sull'intero sistema economico, i relativi rendimenti marginali e i problemi derivanti dal generalizzare alla banda larga ultraveloce. Dopo aver definito l'innovazione, si commenta perché le nuove reti possono essere considerate GPT e cosa ne deriva, il nuovo contesto dinamico che si viene a creare e chi potrebbe beneficiarne. Successivamente si analizza perché la banda larga ultraveloce potrebbe essere una delle responsabili di tassi eterogenei di crescita tra realtà diverse e la forte complementarità che vanta con altre forme di conoscenza. Infine, si riportano altri contributi esterni riguardanti il legame tra ICT e innovazione, il vitale contributo della nascita di startup alla crescita economica (e una panoramica sulla situazione attuale) e variabili che entrano in gioco e potrebbero impattare significativamente sugli effetti positivi delle reti NGA sull'innovazione.

1.2.1 Crescita economica

La prima dimensione per valutare l'impatto degli investimenti nella banda larga ultraveloce è la crescita economica. Mentre è generalmente prevista una relazione positiva, la portata di questa non è immediata e anche il potenziale impatto differenziato di soluzioni tecnologiche alternative (tra cui la velocità raggiungibile) non è previsto. Il paper di Briglauer & Gugler [2018] è uno dei primi studi completi che analizza i dati delle connessioni a banda larga ultraveloci. Utilizzando un set di dati sul dispiegamento negli stati membri dell'Unione Europea per l'arco temporale dal 2003 al 2015, stimano un piccolo ma significativo effetto sul PIL: un aumento dell'1% nell'adozione porta a una variazione incrementale di circa lo 0,004-0,005%. Tuttavia, i maggiori effetti di crescita sono stimati per l'adozione della banda larga di base, quindi la copertura parziale ma non totale della banda larga ultraveloce comporterà i maggiori benefici economici netti. Inoltre, rilevano guadagni marginali decrescenti con l'aumento della velocità, concordando con lo studio sviluppato dal *Copenhagen Economics* [2010] relativo all'economia danese. Quest'ultimo stima che con un aumento della velocità della banda larga da 5 a 10Mbit/s, il PIL crescerà dell'1,9%; mentre il guadagno è solo approssimativamente pari a 0,5% quando la velocità aumenta da 25 a 30Mbit/s. Questi rendimenti marginali decrescenti sono coerenti con la disponibilità a pagare dei consumatori dal lato della domanda. Utilizzando i dati di un sondaggio statunitense somministrato nel 2009, Rosston et al. [2010] stimano un modello di utilità delle preferenze delle famiglie per i servizi Internet a banda larga, offrendo tre diversi gradi di velocità (lento, veloce e molto veloce). I risultati mostrano che la famiglia rappresentativa ha un'alta *willingness to pay* per un servizio ad alta velocità, ma un basso incremento marginale per un servizio ad altissima velocità. Una famiglia rappresentativa degli Stati Uniti era disposta a pagare un premio relativamente più alto per un aggiornamento della velocità della banda larga da un servizio "lento" a "veloce" (circa 25 dollari al mese), ma solo un piccolo premio aggiuntivo (di circa 3 dollari) per un

aggiornamento da veloce a un "molto veloce". Sempre su questo tema, Grzybowski et al. [2018] osservano che i consumatori apprezzano le connessioni FTTH significativamente di più delle connessioni DSL con velocità di 1-8Mbit/s e che la valutazione di connessioni molto stia aumentando nel tempo.

Investire in connessioni è giustificato se la variazione dei benefici generati supera i costi necessari per il dispiegamento e quindi è fondamentale soffermarsi sugli effetti delle differenze di velocità. Come espresso in precedenza, più è alta la capacità che un paese vuole raggiungere, più investimenti sono necessari. Rohman & Bohlin [2012] analizzando 33 paesi OECD, ricavano che raddoppiare la velocità della banda larga contribuirà alla crescita del PIL dello 0,3%. Gruber et al. [2014] concludono uno studio sulla diffusione nell'Unione Europea sostenendo che il beneficio economico della banda larga ad alta velocità, in termini di crescita del PIL, supera sempre il costo dell'investimento, anche sotto la tecnologia più performante. I paper citati successivamente, ad eccezione di Hasbi [2017], Briglauer & Gugler [2018], Grzybowska et al. [2018], Canzian et al. [2019], Abrardi & Cambini [2019] e Cambini & Sabatino [2020] analizzano l'impatto della banda larga. Generalizzare alla banda larga superveloce i risultati ottenuti sulla banda larga di base presenta problemi a causa dei loro diversi profili di costo e prestazione. In letteratura sono tuttavia presenti risultati contrastanti sull'impatto dell'alta velocità e probabilmente ciò è dovuto alla strategia empirica impiegata: alcuni studiosi hanno utilizzato misure della sua disponibilità (cioè lo sviluppo della rete), piuttosto che della sua adozione. I bassi tassi di adozione implicano un sostanziale eccesso di capacità sul lato dell'offerta. Di conseguenza, gli studi empirici che impiegano misure di investimento in banda larga (disponibilità) sottostimano gli effetti economici reali e i guadagni di benessere effettivi (Czernich [2014]).

1.2.2 Definire innovazione

L'innovazione è "l'atto o il processo di introdurre nuove idee, dispositivi o metodi" oppure può essere definita come "lo sfruttamento economico di un'invenzione" (Roberts [1987]). In termini semplici, la società passa dall'invenzione all'innovazione quando un'invenzione viene commercializzata e acquistata. Ciò significa che un produttore è in grado di sviluppare un artefatto che darà ai clienti un'utilità superiore al costo di produzione, nonché offrirlo a un determinato prezzo e in modo tale che i clienti siano effettivamente in grado di riconoscere tale valore (M. Cantamessa, F. Montagna [2016]). L'innovazione è poi seguita dalla diffusione, che è il processo con cui il mercato adotta progressivamente la nuova tecnologia e la rende mainstream. La crescita economica e l'innovazione sono fortemente intrecciate. Robert Solow [1981], ha dimostrato che la crescita del PIL non può essere spiegata semplicemente guardando la crescita dei fattori di produzione (cioè, capitale e lavoro) ma sono fattori determinanti la crescita della produttività, il progresso della conoscenza e della tecnologia.

Come misura dell'innovazione si sono considerate la nascita di startup innovative e le registrazioni brevettuali dal 2013 al 2019. Alcuni studiosi hanno già dimostrato l'impatto positivo della UBB sulla creazione di nuove imprese: la disponibilità dell'infrastruttura basata sulla fibra ha un effetto significativo nelle aree dove è disponibile il capitale umano dell'alta istruzione (McCoy et al. [2016]) e aumenta del 2,7% le aziende che operano all'interno di un mercato (Hasbi [2017]).

Tuttavia, l'obiettivo di questa tesi è restringere il campo focalizzandosi su quel particolare tipo di imprese che cercano di affermare la propria proposta innovativa. Il rischio e l'incertezza, che sono inerenti al ciclo di vita delle startup, portano a una serie di effetti che hanno il potenziale desiderabile di stimolare il processo di innovazione. Prima di tutto, il rischio di fallimento rende il processo selettivo e competitivo. Gli individui e le aziende impegnate faranno di tutto per proporre soluzioni nuove, alternative e migliorate, creando così un'arena sperimentale molto ampia in cui le soluzioni più promettenti possono alla fine avere successo. La società ne beneficia poiché, a causa della stessa incertezza, non sarebbe possibile scegliere la soluzione migliore ex-ante. Anche se un decisore illuminato fosse in grado di farlo, la mancanza di concorrenza probabilmente non creerebbe abbastanza incentivi per sviluppare completamente quella stessa soluzione promettente. In secondo luogo, il processo di selezione indurrà le imprese che sposano una soluzione perdente a fallire o a ridursi. Questo migliora i rendimenti potenziali per i vincitori, che guadagneranno quote di mercato. Allo stesso tempo, crea anche un importante effetto di ricaduta della conoscenza. Almeno una parte dei dipendenti licenziati da un'azienda perdente sarà probabilmente assunta da un'azienda vincente, portando con sé know-how ed esperienza (M. Cantamessa, F. Montagna [2016]).

Il valore dei brevetti come misura dell'innovazione è esaminato in dettaglio da Griliches [1990]: è possibile utilizzarne il valore come misura dell'innovazione poiché essi rappresentano un "quantum minimo di invenzione" che ha superato il controllo degli esaminatori di brevetti ed è di importanza economica dimostrata per l'attività industriale. Pakes e Griliches [1984] hanno formalizzato un'equazione di brevetto come una rappresentazione di una "funzione di produzione della conoscenza" che relaziona una conoscenza latente (rappresentata dai brevetti) a una serie di input. Le ICT potrebbero essere, a loro volta, considerate come qualche tipo di input di capitale avanzato, in aggiunta alla R&S, in una funzione di produzione della conoscenza. Essi misurano però solo un particolare segmento di tutte le innovazioni, vale a dire quelle associate a invenzioni originali e legalmente codificate. Nonostante tale limitazione, è indiscutibile l'incentivo creato da questo strumento giuridico. Se i brevetti non esistessero non ci sarebbe alcun motivo per gli individui e per le imprese di investire nello sviluppo tecnologico. Infatti, non sarebbe razionale perseguire qualsiasi sforzo di R&S sapendo che, dopo aver sostenuto i relativi costi e rischi, i concorrenti potrebbero semplicemente copiare l'invenzione e togliere all'azienda il vantaggio competitivo, senza nemmeno dover recuperare l'investimento. Un governo che concede un brevetto permette quindi all'inventore di diventare un monopolista e di "tassare i cittadini" come ricompensa per l'investimento fatto nello sviluppo dell'invenzione; il meccanismo dei brevetti dovrebbe poi autoregolarsi rispetto all'importanza dell'invenzione. Un'invenzione con poco appeal sul mercato sarà ricompensata con magri profitti di monopolio, mentre un'invenzione molto utile avrà un ampio successo commerciale e porterà a profitti abbondanti. Se i brevetti non esistessero, l'unica arma a disposizione degli investitori sarebbe la segretezza. Questo porterebbe a un processo inefficiente per la società di perseguire l'innovazione, dato che costringerebbe gli aspiranti imitatori a "partire da zero" invece di permettere loro di imparare dall'arte precedente e cercare di migliorarla ("stare sulle spalle di altri inventori"). Infine, il brevetto crea un sistema di diritti di proprietà legalmente definiti che permette il commercio della proprietà intellettuale, il che può portare a un uso sociale efficiente del risultato dell'attività inventiva. I governi possono fornire un ambiente economico e istituzionale che supporta o ostacola l'innovazione. Le scelte hanno un impatto su tutte le imprese che operano nell'economia, ma specialmente su quelle che investono in innovazione. Un ambiente favorevole può aumentare l'*'upside'* (cioè i potenziali ritorni sugli investimenti fatti, se hanno successo) e abbassare il *'downside'* (cioè la perdita di valore, se non hanno successo).

Un'economia dinamica e altamente competitiva, fornendo reti di telecomunicazioni efficienti e incentivi legali, sarà probabilmente in grado di aumentare l'offerta e la domanda per i prodotti e i servizi innovativi.

1.2.3 Guadagni generalizzati e nuove opportunità

La nostra società ha subito alcuni importanti cambiamenti negli ultimi decenni, soprattutto a causa dell'introduzione dell'*Information and Communication Technology* (ICT). Nate nel mondo militare, si sono diffuse in ogni attività economica e sono diventate il pilastro fondamentale dell'economia attirando un grande interesse perché in grado di superare molti degli svantaggi dei sistemi convenzionali (Takahashi et al. [2004]). La caratteristica principale è l'applicazione della conoscenza e allo stesso tempo la generazione progressiva di nuova e cumulata conoscenza. Queste tecnologie stanno guidando un'intera epoca di crescita economica e rientrano per questo nelle GPT⁵. Così, quando la banda larga ultraveloce si diffonde, genera guadagni di produttività generalizzati. Ciò implica un aumento dei rendimenti di scala e supporta il tasso di avanzamento tecnico a favore dell'innovazione.

L'avanzata può causare però altri effetti indesiderati poiché la concorrenza delle giovani imprese digitali può spingere i tradizionali operatori storici fuori dal mercato, aumentando così il tasso di uscita dal settore. Allo stesso tempo, si creano nuove opportunità per i potenziali entranti, permettendo alle startup di inserirsi con successo nel mercato, generando potenziali ricadute positive anche sull'attività innovativa degli incumbent (Aghion et al. [2009]). Gli effetti sulle imprese già presenti sono eterogenei: la maggior minaccia da parte di nuovi operatori tecnologicamente avanzati dovrebbe incoraggiare l'innovazione degli operatori storici, nei settori in cui gli incumbent sono inizialmente vicini alla frontiera tecnologica ("*escape-entry effect*"); mentre può scoraggiarne l'innovazione nei settori dove il divario tra le competenze possedute e quelle necessarie per rispondere efficacemente è evidente ("*discouragement effect*" of entry). Nel primo caso, sanno che possono ovviare all'ingresso innovando e pertanto maggiore minaccia di ingresso porterà ad attività innovative più intense. Questo fenomeno è molto simile all'effetto "nave a vela" (Geels [2005]): nel diciannovesimo secolo si verificò un significativo e improvviso aumento delle prestazioni nella tecnologia a vela a lungo raggio, solo quando la concorrenza dei battelli a vapore iniziò a sorgere. Nel secondo caso, le imprese storiche non hanno alcuna speranza di raggiungere le proposte innovative dei nuovi arrivati, quindi i loro investimenti non saranno incentivati. Tuttavia, anche a fronte di soluzioni tecnicamente migliori degli entranti, non è certo che quest'ultimi riescano a conquistare il mercato poiché il design dominante precedente potrebbe causare un effetto lock-in e contrastare il cambiamento tecnologico.

La motivazione per cui piccoli entranti colgono più spesso opportunità innovative può essere spiegata dalla differenza degli obiettivi rispetto agli incumbent (Swinney et al. [2011]). Gli operatori storici sono solitamente interessati alla redditività del business nel complesso, dunque, tendono a massimizzare il valore attuale netto di un portafoglio

⁵ General Purpose Technology: insieme di tecnologie in grado di influenzare e orientare un'intera economia. Le caratteristiche principali sono: la pervasività, il potenziale intrinseco di miglioramenti tecnici e la complementarità nelle applicazioni di prodotto e processo, ovvero la produttività nei settori a valle aumenterà come conseguenza dello sviluppo della tecnologia a monte (Bresnahan & Trajtenberg [1995])

aggregato di progetti. Al contrario, i nuovi arrivati si preoccupano della sopravvivenza, quindi cercheranno di minimizzare la probabilità di fallimento su un portafoglio limitato di progetti. Il momento in cui l'innovazione viene perseguita dipenderà dall'incertezza della domanda e dalla tendenza con cui l'investimento richiesto dovrebbe evolvere nel tempo. Se gli attori si aspettano che i costi di investimento ΔC diminuiscano, entrambi decideranno di aspettare. D'altro canto, se si aspettano che i costi rimangano costanti o aumentino, la loro risposta dipenderà dall'incertezza della domanda σ^2 (Figura 1-14). Se è bassa, entrambi entreranno presto. Tuttavia, nel contesto attuale l'incertezza della domanda per l'innovazione sarà generalmente alta, questo ridurrà allo stesso tempo la redditività prevista ma aumenterà le possibilità di sopravvivenza. Sarà quindi razionale per un entrante agire prima di un incumbent, creando così l'opportunità di un'innovazione dirompente e di affermare un paradigma emergente.

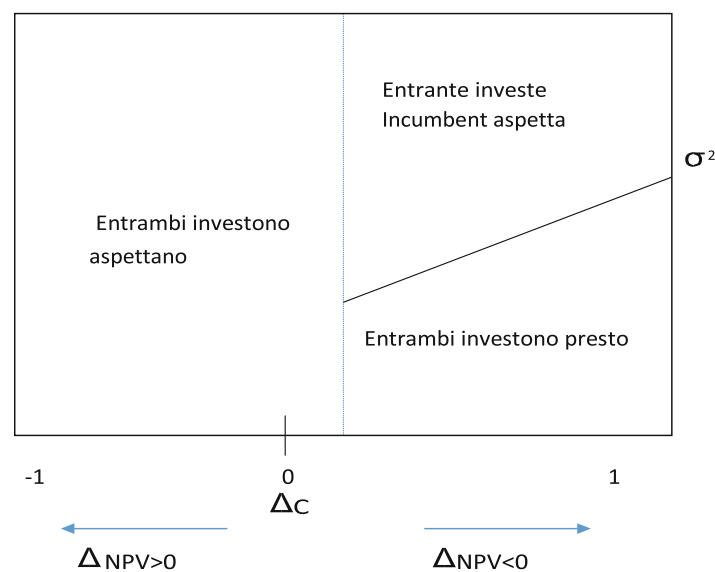


FIGURA 1-14 MODELLO DI SWINNEY

Un'altra spiegazione della reattività dei piccoli entranti potrebbe essere legata all'andamento delle vendite durante l'evoluzione di una nuova tecnologia. La curva di diffusione può essere approssimata a una curva di distribuzione normale (Figura 1-15), e poi divisa al punto medio e alle deviazioni standard -2, -1 e +1 (Rogers [1962]). I clienti tenderanno ad adottarla in momenti molto diversi lungo il ciclo di vita del prodotto. Il primo segmento rappresenta gli *'innovators'*, coloro che la adottano perché "guardano al futuro", e sono quindi pronti ad accettare una tecnologia relativamente immatura. Solitamente è una quota molto piccola del mercato raggiungibile (circa il 2%) e quindi non così attraente per le grandi imprese, dal punto di vista economico. Al contrario, le startup potrebbero trovarlo interessante e muoversi per prime poiché intravedono la possibilità di utilizzarlo come piccolo mercato di testa di ponte. Ciò permette ai nuovi entranti di reinvestire i margini

delle vendite in R&S, accelerare lo sviluppo tecnologico e anticipare il giorno in cui la nuova tecnologia diventerà competitiva.

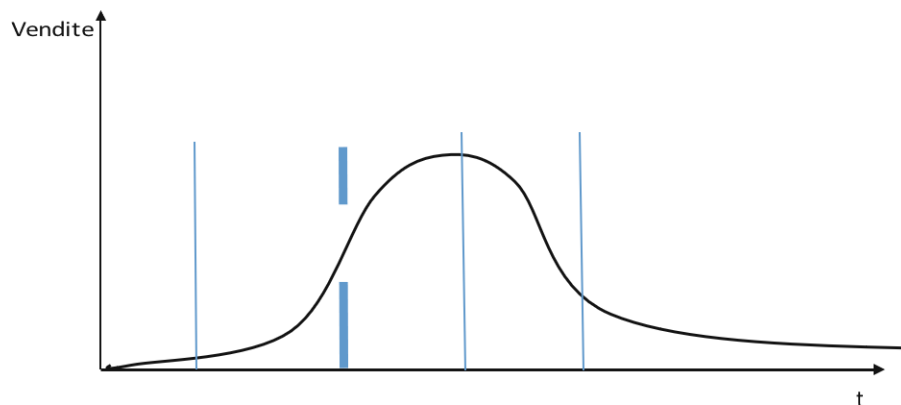


FIGURA 1-15 CURVA DI ROGERS

1.2.4 Eterogeneità negli effetti e asset complementari

Sebbene effetti indiretti rendano attraente il contesto per potenziali entranti, le grandi imprese incumbent potrebbero ottenere un ritorno maggiore dalla diffusione della banda larga ultraveloce, rendendo così le piccole imprese meno competitive. Questo accade perché l'implementazione dell'ICT avanzata aumenta sia la domanda di capitale umano che i costi irrecuperabili associati al cambiamento organizzativo (Brynjolfsson e Hitt [2000]). Una barriera all'investimento per le imprese di piccole dimensioni può essere infatti costituita dall'aumento dei costi e dall'assenza degli investimenti complementari (Bugamelli e Pagano [2004]).

La produttività marginale dell'ICT è infatti molto più alta del suo costo d'uso marginale se e solo se l'adozione è integrata con assets che beneficiano di tale tecnologia. A sostegno di tale affermazione, Polder et al. [2010] trovano che l'uso di internet a banda larga è particolarmente importante per le imprese dove la rete è positivamente correlata all'innovazione di prodotto, di processo o organizzativa. Cambini & Sabatino [2020] evidenziano che l'adozione, in assenza di investimenti complementari, non produce risultati significativamente migliori (e spesso persino peggiori) rispetto alla non adozione. Ciò implica che è la presenza di investimenti complementari che fa leva sui benefici dell'adozione dell'UBB per la performance aziendale e non viceversa. Sono d'accordo anche Colombo et al. [2013] che studiando un panel di 6 anni di PMI italiane, trovano che la produttività dell'impresa è influenzata positivamente dall'aumento dell'uso di applicazioni complementari alla banda larga se e solo se le applicazioni sono sufficientemente avanzate. Internet a banda larga è una condizione necessaria, ma non sufficiente per l'impatto positivo sull'innovazione e di conseguenza sul benessere economico.

Man mano che queste tecnologie diventano prevalenti nell'economia, le imprese che non sono in grado di implementare gli asset complementari nel loro business vengono dunque spinte fuori dal mercato (Aghion et al. [2009]). L'effetto benefico finale è che i costi

aggregati diminuiranno poiché in un mercato più concentrato sarà alleviata la cattiva allocazione delle risorse da parte di piccole imprese inefficienti.

Inoltre, la dimensione delle imprese è legata in modo cruciale anche al turnover (Cambini & Sabatino [2020]). Solo l'ingresso di grandi imprese è influenzato positivamente, in particolari settori industriali. Un risultato ancora più sorprendente emerge riguardo al tasso di uscita delle imprese. La diffusione dell'UBB aumenta significativamente il numero totale di piccole e medie imprese commerciali che chiudono, mentre non è significativo per le grandi imprese. Le infrastrutture avanzate a banda larga rimodellano dunque il paesaggio industriale verso un contesto più concentrato.

Tutte queste argomentazioni concordano che non tutte le attività commerciali beneficino allo stesso modo delle tecnologie a banda larga di ultima generazione. Inoltre, l'eterogeneità negli effetti si verifica anche nell'implementazione delle tecnologie digitali in paesi diversi. Le differenze di reddito, consumi e crescita possono essere in parte spiegate dalle ricadute positive delle ICT (Easterly [2001]) perché l'adozione efficiente non è sostenibile ovunque. Il cambiamento tecnologico, che si è verificato negli ultimi decenni, è strettamente legato alla trasparenza e favorisce i paesi con istituzioni più meritocratiche e sistemi di incentivi più obiettivi.

Pellegrino e Zingales [2017] hanno evidenziato che gli imprenditori perdono collettivamente da un ambiente meno incline (a causa di scarse competenze digitali) all'adozione delle ICT; tuttavia, essi perdono individualmente dall'adottare tecnologie basate sulla trasparenza quando i loro pari non lo fanno. Dato questo enigma e il fatto che le istituzioni di un paese sono intrinsecamente difficili da cambiare, in Italia le imprese non sono riuscite ad accedere ai pieni benefici dalla diffusione del digitale e la scena politica dovrebbe ripensare le istituzioni: non semplicemente appropriate in un frangente storico, ma in sintonia con il ritmo del progresso tecnologico. Sempre su questo filone della letteratura, Schivardi e Schmitz [2019] analizzano la relazione tra pratiche manageriali, adozione dell'IT, dimensioni delle imprese e produttività aggregata per i paesi del Sud Europa (Italia, Spagna e Portogallo) rispetto alla Germania, identificando le cattive pratiche di gestione come i principali driver di una minor crescita e adozione delle nuove tecnologie.

Le ICT influenzano anche la flessibilità posseduta dalle aziende di un paese di adattarsi alle contingenze delle economie, permettendo di modellare la loro offerta alle esigenze del mercato (Vilaseca et al. [2007]). Questo è diventato particolarmente rilevante negli ultimi anni poiché a causa della globalizzazione, le aziende stanno affrontando sfide significative originate dall'aumento della competitività, la comparsa di nuovi prodotti e le crescenti aspettative dei consumatori. Baquero Forero [2013] inoltre rileva che i paesi con maggior investimenti nelle infrastrutture di comunicazione mobile e nella banda larga sono tendenzialmente più vicini alla frontiera tecnologica globale. Le ICT possono avere un effetto determinante per colmare il divario tecnologico nei confronti delle nazioni leader e, rispetto all'accumulazione del capitale fisico, è ciò che conta di più quando si tratta di spiegare la stagnazione o la crescita di una certa economia.

1.2.5 Legame con altre forme di conoscenza

Il legame tra banda larga ultraveloce e altre forme di capitale di conoscenza può sostenere gli sforzi di R&S e innovazione delle aziende; tuttavia, non è ancora stato codificato il ruolo di questo contributo.

Bertschek et al. [2013], utilizzando dati a livello aziendale per la Germania tra il 2001 e il 2003, esaminano la relazione tra l'adozione della banda larga da parte delle imprese e i cambiamenti nella loro produttività del lavoro (misurata come vendite per dipendente) e nell'attività innovativa. L'adozione ha un effetto positivo significativo sull'innovazione con effetti di trattamento medi stimati nell'ordine del 41-45% per le innovazioni di prodotto e del 28-35% per le innovazioni di processo. Anche se la quantificazione esatta dell'impatto della banda larga sulle innovazioni di processo rimane difficile, gli effetti positivi e altamente significativi suggeriscono che, durante o dopo la fase di espansione del DSL, internet a banda larga ha aumentato la probabilità di rimodellare e riorganizzare processi aziendali. La grande dimensione del coefficiente potrebbe riflettere il fatto che il dispiegamento delle reti, come tecnologia di uso generale, beneficia di esternalità di rete e di conoscenza diffondendosi in tutta l'economia mondiale. Così, l'effetto stimolante di internet a banda larga potrebbe essere rafforzato da esternalità legate all'adozione da parte dei fornitori e dei clienti lungo la catena del valore. Tuttavia, questi effetti potrebbero essere sovrastimati a causa della correlazione negativa tra l'adozione della banda larga e fattori non osservati che aumentano l'attività innovativa. Ciò nonostante, essi sono altamente significativi e robusti rispetto alle diverse specificazioni. La banda larga nella sua fase di espansione ha permesso alle aziende di ottimizzare i processi aziendali e sviluppare o migliorare nuovi prodotti e servizi esistenti. L'impatto sulla produttività invece è eterogeneo tra le imprese e non statisticamente diverso da zero. Ciò suggerisce che la riorganizzazione è stata accompagnata da una fase di sperimentazione e apprendimento, che è tipica dell'introduzione di una general purpose technology. Gli effetti potrebbero sorgere nel lungo periodo (Bertschek et al. [2013]).

N. Bloom et al. [2010] valutano se il capitale ICT possa sostenere gli sforzi di R&S delle imprese. Si concentrano inizialmente sui brevetti come la principale misura di innovazione senza ottenere però un impatto significativo. Giustificano tali risultati sostenendo che il ruolo principale dell'ICT nell'innovazione risiede probabilmente nel suo utilizzo come strumento complementare alle innovazioni di prodotto e strutture di costo che non vengono misurate nel quadro dei formali diritti di proprietà intellettuale. La letteratura ha affrontato una sfida difficile nel tentativo di quantificare queste innovazioni. Una delle principali domande a livello aziendale è se diversi tipi di innovazione si precludono a vicenda, o se siano invece complementari.

Le linee di argomentazione che suggeriscono la sostituibilità sono:

- Le aziende sono tipicamente vincolate alla liquidità, ciò rende difficile per loro condurre allo stesso tempo due tipi di innovazione.
- Le capacità manageriali sono limitate e sarà molto complicato implementare i due processi in parallelo all'interno di un'organizzazione.

Gli argomenti a favore della complementarità sono:

- Sia le innovazioni di prodotto che di processo tendono ad aumentare la produzione di equilibrio e quindi ad incrementare la redditività dell'altra innovazione.

- Le capacità/competenze che permettono a un'azienda di affrontare un tipo di innovazione saranno probabilmente utili anche per implementare l'altro tipo.

La maggior parte degli studi esistenti ipotizzano che le innovazioni siano create e utilizzate nella stessa organizzazione, questo implica che non ci siano problemi di incompatibilità degli incentivi tra le organizzazioni. Il secondo presupposto è che la generazione e l'adozione dell'innovazione siano lo stesso processo. Ciò rimuove un ulteriore problema: una volta che un'innovazione c'è, viene adottata, e un'innovazione viene creata solo se è redditizia. Nel contesto europeo, le aziende sono però solitamente abbastanza capaci nel proporre innovazioni, ma molto più lente delle loro controparti d'oltreoceano nell'adottarle.

L'utilizzo delle ICT avviene con l'obiettivo di migliorare la redditività dell'azienda, aumentandone la capacità di generare ricavi (innovazione di prodotto) o abbassandone i costi operativi (innovazione di processo).

Per essere in grado di arrivare a conclusioni solide sulla relazione tra innovazioni di prodotto e di processo, N. Bloom et al. identificano due forme utili per le innovazioni di prodotto e di processo. Dato che la maggior parte dell'hardware è considerata tecnologia di uso generale (GPT), è utile concentrarsi sull'adozione di software che tipicamente cambiano una parte più strettamente definita delle routine aziendali. La scelta ricade sul software di gestione delle risorse umane come un'innovazione di processo, e il software di sviluppo delle applicazioni per quelle di prodotto. Il software di gestione delle risorse umane riguarda direttamente la gestione della forza lavoro piuttosto che un qualsiasi prodotto specifico venduto dall'azienda e i costi operativi degli adottanti probabilmente diminuiranno. Le APPS sono invece altamente specifiche dell'industria, e spesso supportano la personalizzazione di massa o specializzati componenti che entrano nel prodotto finale. Le applicazioni sono fonti di differenziazione tra le aziende e l'uso effettivo è idiosincratico, quindi difficile da esternalizzare.

L'identificazione dei diversi tipi di software come innovazioni di prodotto e di processo è cruciale per studiare le interazioni tra questi due tipi di innovazioni. La semplice ricerca di correlazioni nell'uso trascura l'eterogeneità non osservata che potrebbe essere il fattore chiave che guida l'adozione congiunta; ciò essenzialmente lascia il ricercatore a indovinare quale ruolo è giocato dalle complementarità tecnologiche e quale dall'eterogeneità non osservata.

Per affrontare questo, N. Bloom et al. [2010] hanno sviluppato un modello econometrico che tiene conto sia delle potenziali complementarità che l'eterogeneità non osservata. Come applicazione empirica, stimano l'effetto di un aumento esogeno della pressione competitiva, in quanto sono distinguibili diversi regimi concorrenziali e il conseguente impatto sui modelli di adozione di innovazioni di prodotto e di processo. Il contesto studiato è la scadenza del regolamento UE 1475/95, che in precedenza consentiva alle case automobilistiche francesi di limitare la concorrenza nel mercato attraverso rivenditori indipendenti. Un aumento della pressione competitiva identificata non ha avuto alcun effetto indipendente sull'adozione dell'innovazione di prodotto o di processo ma impatta sulla scala dei concessionari d'auto. In sostanza, i concessionari d'auto sono stati precedentemente limitati nella loro crescita non potendo aprire filiali al di fuori del territorio assegnato e non potendo vendere pezzi di ricambio ai rivenditori. L'eliminazione di questa restrizione ha permesso loro di crescere. In secondo luogo, registrano una correlazione positiva tra l'aumento di scala e l'adozione delle innovazioni di prodotto, che risultano essere sostitutive con quelle di processo. Questo suggerisce che ci possono essere limitazioni manageriali o finanziarie che impediscono l'adozione contemporanea di più

innovazioni. In sintesi, ci sono significative complementarità tra i diversi tipi di innovazioni e altre strategie aziendali. Tuttavia, il test del rapporto di verosimiglianza suggerisce però che sia le complementarità che l'eterogeneità non osservata contribuiscono a spiegare la variazione nel comportamento di adozione delle imprese, in particolare il fatto che le scelte di scala e le innovazioni di prodotto e/o di processo sono fatte simultaneamente.

1.2.6 Altre evidenze su ICT

Lo scopo di questa sezione è riportare altri contributi esistenti riguardanti il legame tra ICT e innovazione, non trattati precedentemente.

Tra i primi studi inerenti, Becchetti et al. [2003] registrano un impatto positivo degli investimenti in telecomunicazioni sulla creazione di nuovi prodotti e processi. Morikawa [2004] ha scoperto che le PMI giapponesi che usano i computer hanno maggiori probabilità di impegnarsi in attività innovative rispetto alle aziende senza applicazioni informatiche. Questo studio non ha però potuto rivelare se la forte relazione tra adozione e attività innovativa esistesse per tutti i tipi di applicazioni digitali. Koellinger [2008], analizzando un campione di aziende europee, ha identificato le tecnologie di e-business come importanti facilitatori dei processi innovativi. Hempell e Zwick [2008] adottano un approccio diverso: indagano in che misura l'uso delle ICT favorisca le attività innovative incentivando strutture organizzative più flessibili. Innanzitutto, distinguono tra flessibilità funzionale (la capacità dei lavoratori di cooperare e prendere decisioni decentralizzate) e flessibilità numerica (la riduzione dei costi fissi, principalmente a causa dell'esternalizzazione dei processi aziendali). L'uso delle tecnologie dell'informazione si è rivelato associato a un incremento di entrambi i tipi di flessibilità, ma le implicazioni per le attività di innovazione sono state diverse. Mentre la flessibilità funzionale è fortemente e positivamente associata alle innovazioni di prodotto, la flessibilità numerica tende a ridurre la capacità innovativa nel lungo periodo. Pertanto, è lecito aspettarsi che applicazioni ICT abbiano impatti diversi per innovazioni di processo e di prodotto. Quando un'azienda decide di disimpegnarsi dallo sviluppo in una determinata area tecnologica, disperde progressivamente le sue competenze e ciò potrebbe spiegare l'impatto negativo dell'aumento della flessibilità numerica sulle attività innovative interne. Infatti, Hempell [2005] aveva rilevato che l'esperienza innovativa, al contrario di quanto accade per gli investimenti convenzionali, gioca un ruolo importante per quanto riguarda l'uso di successo delle ICT poiché a differenza di altri beni sono fortemente legate allo sviluppo tecnologico e alla reingegnerizzazione dei processi.

Secondo D. Añón Higón [2011], le ICT operano principalmente come tecnologie di miglioramento dell'efficienza e riduzione dei costi, anche se specifiche applicazioni orientate al mercato (in particolare lo sviluppo di siti web), hanno il potenziale di generare un vantaggio competitivo attraverso l'innovazione di nuovi prodotti. Inoltre, ha sottolineato che l'impatto dipende dal tipo di applicazione e dalle diverse caratteristiche dell'azienda e del management, nonché dai fattori esterni. La diversa natura delle applicazioni ICT non solo ha un impatto particolare a seconda dell'attività di innovazione, ma può implicare requisiti diversi in termini di risorse e incentivi necessari per l'implementarle. Le politiche mirate ad aumentare il potenziale di innovazione delle aziende, attraverso il rafforzamento di nuove opportunità dal lato della domanda, dovrebbero concentrarsi su applicazioni mirate all'innovazione di prodotto. Invece, le politiche che mirano a migliorare l'efficienza delle aziende, attraverso l'introduzione di nuovi processi, dovrebbero dare la priorità allo sviluppo

di applicazioni ICT che integrano la produzione (ad esempio il CAD⁶) insieme alle applicazioni generali.

Le imprese ad alta intensità di conoscenza, che utilizzano software aziendali personalizzati in modo da soddisfare i loro requisiti specifici, hanno maggiori probabilità di innovare rispetto alle imprese che usufruiscono di software generici (Engelstätter et al. [2010]). Questo significa che l'incorporazione di esperienza e conoscenza a lungo termine nell'applicazione software aumenta il know-how IT e contribuisce all'attività di innovazione, poiché riflette tutti i processi di business necessari che portano alla loro ottimizzazione. Affinché ciò avvenga, il software aziendale deve essere sviluppato e applicato correttamente, l'azienda deve essere consapevole della propria struttura organizzativa, dei processi e degli obiettivi che vuole raggiungere.

Ollo López & Aramendia-Muneta [2012] registrano che l'uso di un sistema software specifico relativo al design (CAD) ha un effetto positivo sul lancio di un nuovo prodotto, e se utilizzato insieme con un ERP⁷, si genera complessivamente un impatto positivo anche sull'introduzione di un'innovazione di processo. Per quanto riguarda la competitività, registrano effetti diversi per le tecnologie di comunicazione. La LAN⁸ non è significativa, la LAN wireless⁹ impatta positivamente sulla concorrenza mentre le connessioni EDI¹⁰ hanno un effetto negativo.

Infine, Vilaseca et al. [2007] analizzano l'uso delle ICT nel marketing come un fattore in grado di abbattere le barriere all'innovazione e migliorare il processo di cooperazione. Attraverso le ICT è possibile stabilire e mantenere relazioni comunicative e di cooperazione sia all'interno che all'esterno dell'organizzazione. L'analisi dei dati ha dimostrato l'esistenza di una relazione diretta tra cooperazione e innovazione: l'88,4% delle organizzazioni, in cui ha avuto luogo innovazione di prodotto, coopera con istituzioni scientifiche. Inoltre, anche la cooperazione orizzontale e la cooperazione con gli agenti all'interno della catena del valore (fornitori e clienti) hanno ottenuto un punteggio elevato, rispettivamente 81,9% e 73,6%. I benefici dell'uso dell'ICT non si limitano al processo di cooperazione: registrano inoltre la creazione di team di lavoro flessibili e grazie allo stretto legame tra queste tecnologie e la possibilità di acquisire dati del mercato di appartenenza, il loro utilizzo (durante il processo innovativo) ha favorito lo sviluppo di nuovi beni e servizi differenziati con caratteristiche mirate a soddisfare le esigenze e i requisiti della domanda.

1.2.7 L'importanza delle startup

Secondo la nota definizione di Steve Blank [2010], una startup è “un'organizzazione temporanea in cerca di un business model replicabile e scalabile”.

⁶ Computer-Aided Drafting: il settore dell'informatica volto all'utilizzo della computer grafica per supportare l'attività di disegno tecnico

⁷ Enterprise Resource Planning: software di gestione che integra tutti i processi di business rilevanti di un'azienda con le funzioni aziendali

⁸ Local Area Network: rete informatica di collegamento tra più computer, estendibile anche a dispositivi periferici condivisi, che copre un'area limitata

⁹ Wireless Local Area Network: rete locale che sfrutta la tecnologia wireless, invece di una connessione cablata via cavo

¹⁰ Electronic Data Interchange: interscambio di dati tra sistemi informativi, attraverso un canale dedicato e in un formato definito in modo da non richiedere intervento umano

In un'economia moderna la nascita delle startup innovative è fondamentale. Queste aziende contribuiscono al dinamismo economico iniettando concorrenza nei mercati e stimolando l'innovazione: nuove idee sono portate sul mercato e trasformate in imprese economicamente sostenibili. Per un commento sul rischio, inerente al core business di questo particolare tipo di impresa, come stimolo per innovare, si rimanda al sottoparagrafo *1.2.2 Definire innovazione*.

Una panoramica sul mondo delle startup italiane è fornita dal rapporto del 20 settembre 2020 a opera del MISE e InfoCamere, con il supporto della Camera di Commercio. A quasi otto anni dall'introduzione della policy dedicata, le startup innovative italiane sono soprattutto microimprese e vantano un valore della produzione medio poco superiore a 201 mila euro. Ciò è anche dovuto al ricambio costante cui è soggetta questa popolazione: le imprese "best-performer", più consolidate per età e fatturato, tendono progressivamente a perdere lo status di startup innovativa. Per quanto riguarda la redditività, le startup mostrano un'incidenza di società in perdita più elevata della media (oltre il 52,6% contro il 30,8% complessivo). Tuttavia, le società in utile mostrano valori particolarmente positivi in termini di redditività (ROI, ROE) e valore aggiunto. Inoltre, presentano un tasso di immobilizzazioni, uno dei principali indicatori della propensione a investire delle aziende, di circa sette volte più elevato rispetto alle altre aziende comparabili.

Le analisi tradizionali sullo sviluppo di un'economia tendono a concentrarsi sulle grandi aziende e trascurano il vitale contributo delle piccole startup. Per le grandi società, la capacità di crescita è influenzata dalle condizioni generali del business, specifiche di ogni paese. Invece le startup influenzano la redditività attraverso innovazioni di processo e prodotto stimulate dal rischio e l'incertezza in cui si trovano a operare. L'incremento del numero porterà poi a un ulteriore aumento della concorrenza all'interno del mercato e un'influenza positiva sulla crescita economica del paese. Inoltre, un fattore base per la sopravvivenza delle startup è rappresentato dall'adozione delle ICT (Parker & Castelman [2007]) e ciò potrebbe incentivare le grandi imprese ad avvicinarsi al digitale e investire in queste tecnologie.

Quando si tratta di potere di creazione di posti di lavoro, non è la dimensione dell'azienda che conta quanto l'età (Kauffman Foundation, *"The Importance of Young Firms for Economic Growth"* [2014]). Le nuove e giovani aziende sono la fonte primaria di creazione di posti di lavoro: l'1% delle startup più dinamiche è responsabile della creazione del 40% dei nuovi posti di lavoro complessivamente generati in un anno: un fattore che dovrebbe spingere i governi a far nascere nuove startup piuttosto che focalizzarsi sul salvataggio dei business tradizionali. Più la popolazione è coinvolta nell'imprenditorialità, più aumentano i livelli di sviluppo economico (Z. Acs [2006]).

In conclusione, lo sviluppo economico di una nazione dipende fortemente dalla natura complementare dei meccanismi tra le piccole imprese e le società consolidate. Il valore benefico di questo meccanismo varia con il reddito nazionale, misurato dal PIL pro capite. A bassi livelli di reddito nazionale, le startup forniscono opportunità di lavoro e spazio per la creazione di mercati. Come il PIL pro-capite aumenta, l'emergere di nuove tecnologie ed economie di scala permette alle imprese più grandi di soddisfare la crescente domanda e di aumentare la propria quota all'interno del mercato. Infine, quando si verificano ulteriori aumenti di reddito, il ruolo giocato dalle startup aumenta a sua volta, poiché l'ambiente commerciale è maturo e permette lo sfruttamento di nuove opportunità. L'innovazione e la concorrenza saranno il motore del circolo virtuoso che si è generato.

Un noto detto cinese dice: "Dai un pesce a un uomo e lo nutrirai per un giorno. Insegnagli a pescare e lo nutrirai per tutta la vita". Facilitare la nascita di nuove startup (invece di

sussidiare le imprese tradizionali in crisi) dovrebbe essere il mantra di una nazione poiché attraverso l'innovazione e la creazione posti di lavoro, la società sarà più ricca.

1.2.8 Variabili significative per l'innovazione nelle imprese

Riprendendo le evidenze ottenute da Añón Higón [2011], la quale aveva sottolineato come i fattori esterni, quali caratteristiche dell'azienda e management, incidessero significativamente sull'impatto delle ICT sull'innovazione, si procede ad analizzare una serie di variabili che entrano in gioco e possono influenzare l'effetto della banda larga ultraveloce. Scopo di questa sezione è precisare che non è sufficiente e/o necessario lo status di startup per conseguire importanti risultati nelle attività innovative, ma molti fattori hanno il potenziale di condizionare tali processi.

Inizialmente, si discutono alcune variabili strutturali relative alle caratteristiche dell'azienda. Per quanto riguarda la dimensione, la relazione con l'innovazione è stata ampiamente studiata da Schumpeter [1942]. Le grandi aziende sono meglio posizionate per sviluppare e sfruttare nuove tecnologie grazie a una maggiore disponibilità di risorse, capacità di sfruttare economie di scala e superiore appropriabilità del valore creato. D'altro canto, le piccole imprese, caratterizzate da strutture più flessibili, possono essere più innovative nelle risposte ai cambiamenti dei bisogni dei clienti o delle condizioni ambientali (Rogers [2004]). Tuttavia, il dibattito sul ruolo delle dimensioni dell'impresa è ancora in corso in quanto gli studi empirici hanno raggiunto conclusioni contrastanti. Acs & Audretsch [1988] riportano che le piccole imprese hanno un vantaggio nelle industrie altamente innovative e nei mercati competitivi. Tale associazione risulta però molto più significativa con l'innovazione di processo, rispetto quella di prodotto. Le startup sembrano spesso soffrire a causa di una mancanza strutturale di risorse tangibili e intangibili (Wymer & Regan [2005]) e ciò ne ostacola l'adozione di nuovi processi innovativi. Le pratiche di *open Innovation* potrebbero risolvere questo problema (Bogers [2011]), permettendo alle aziende di attingere ad asset complementari e alla ricchezza (forse illimitata) di conoscenza che si può trovare nel mondo. Studiando l'adozione delle ICT, sono stati ottenuti risultati eterogenei a seconda dei metodi utilizzati per quantificare la dimensione aziendale. Sfruttando la metratura degli stabilimenti, Bayo-Moriones & Lera-López [2007] hanno ottenuto una relazione positiva; al contrario riferendosi al numero di dipendenti, sono state raggiunte evidenze opposte, ovvero che il grado di adozione aumenta al diminuire della dimensione.

Se ci focalizzassimo sugli anni di esperienza accumulata, si potrebbe attendere una relazione positiva tra l'età dell'azienda e l'innovazione. Tuttavia, molti studi hanno riscontrato un comportamento maggiormente proattivo, flessibile e aggressivo da parte delle aziende più giovani e dunque, la minore età dell'azienda è risultata spesso significativamente associata all'innovazione (Balasubramanian & Lee [2008]).

La proprietà straniera può poi implicare maggiori risorse finanziarie o l'accesso a nuovi tipi di conoscenza e tecnologia. Tuttavia, la R&S deve essere condotta vicino ai mercati di riferimento e dunque non è stata registrata una relazione univoca con la probabilità di introdurre nuovi prodotti e nuovi processi produttivi.

Le imprese appartenenti a un gruppo imprenditoriale sono generalmente più propense a innovare poiché il gruppo permette alle aziende di superare il problema della mancanza di risorse complementari necessarie, come la finanza, il capitale fisico o umano. La letteratura

sottolinea anche i potenziali benefici del trasferimento di tecnologia dalle società madri alle società figlie (Oakey et al. [1980]).

La crescita della concorrenza delle importazioni dalla Cina (e da altri paesi caratterizzati da basso costo del lavoro) ha inoltre costretto le imprese ad adottare le ICT e innovare per evitare la *mercificazione*. Così, il commercio ha il beneficio di indurre un cambiamento tecnico più veloce sia per le imprese che sopravvivono sia attraverso la riallocazione delle risorse delle altre. I risultati di N. Bloom et al. [2010] hanno evidenziato che il 15% dell'aumento dell'intensità ICT, nel settore manifatturiero tra il 2000-2007, è proprio spiegato dall'aumento del commercio. Nello specifico, circa il 4% di questo valore è derivato da un effetto di riallocazione, mentre la quota restante è legata all'aggiornamento intra-aziendale delle ICT.

C'è ancora un grande dibattito sulla direzione della causalità con l'esportazione: alcuni argomenti suggeriscono che il coinvolgimento espanderà la rete geografica degli imprenditori e li metterà in contatto con una fonte più ricca di conoscenza, competenza e tecnologia non disponibili nel mercato nazionale (Lachenmaie & Wößmann [2006]) e promuoverà inoltre l'adozione delle ICT (Bayo-Moriones & Lera-López [2007]).

Secondo la letteratura empirica, un'altra determinante importante del comportamento innovativo nelle piccole imprese è l'attitudine imprenditoriale. È riconosciuto il legame positivo tra le competenze, il livello di istruzione dell'imprenditore e il grado di performance raggiunto dall'azienda in attività innovative. L'istruzione favorisce anche la ricettività di nuove idee, che giocano un ruolo importante sia nel rilevare il bisogno che nel creare un ambiente favorevole per la sua implementazione (Damanpour & Schneider [2006]). I lavoratori con alti livelli di istruzione mostrano più abilità a sfruttare in modo efficiente il know-how dell'azienda e a tradurre la conoscenza tacita in innovazione (Vilaseca et al. [2007]). Le risorse di conoscenza sono strettamente legate alla flessibilità e alla capacità di aggiornare i prodotti o i processi di produzione, così come sviluppare nuove tecnologie radicali. L'assistenza o la consulenza del governo o altre società esterne può essere considerata come un input di conoscenza avanzata alla funzione di produzione dell'azienda e potrebbe quindi avere un impatto positivo sull'innovazione. A questo proposito, Freel [2000] ha dimostrato che le piccole imprese manifatturiere del Regno Unito riuscirono ad abbattere le barriere all'innovazione di prodotto grazie all'appoggio di società esterne.

Infine, una corrente della letteratura considera l'innovazione come il risultato di un processo interattivo e non lineare tra l'azienda e il suo ambiente (Dosi et al. [1988]). In questo approccio, i fattori ambientali giocano un ruolo importante nel determinare il risultato di tale processo. Un gran numero di studi ha testato l'ipotesi schumpeteriana di una relazione positiva tra potere monopolistico e sforzo innovativo. Scherer [1972], per esempio, ha trovato una relazione a forma di U invertita tra l'intensità dell'innovazione e la concentrazione del mercato. Dunque, non è chiara a priori la direzione dell'impatto sull'innovazione derivante dagli ostacoli alla concorrenza.

RIFERIMENTI BIBLIOGRAFICI

- [1] L. Abrardi e C. Cambini. *Ultra-fast broadband investment and adoption: A survey*. Telecommunications Policy, 43(3):183-198, 2019.
- [2] Z. Acs. *How Is Entrepreneurship Good for Economic Growth?* Innovations: Technology, Governance, Globalization, 1(1):97-107, 2006.
- [3] Z. Acs e D. Audretsch. *Innovation in large and small firms: An empirical analysis*. American Economic Review, 78(4):678-690, 1988.
- [4] P. Aghion, R. Blundell, R. Griffith, P. Howitt, e S. Prantl. *The effects of entry on incumbent innovation and productivity*. The Review of Economics and Statistics, 91(1):20-32, 2009.
- [5] D. Añón Higón. *The impact of ICT on innovation activities: Evidence for UK SMEs*. International Small Business Journal, 30(6):684-699, 2011.
- [6] N. Balasubramanian e J. Lee. *Firm age and innovation*. Industrial and Corporate Change, 17(5):1019-1047, 2008.
- [7] M. d. P. Baquero Forero. *Mobile communication networks and Internet technologies as drivers of technical efficiency improvement*. Information Economics and Policy, 25(3):126-141, 2013.
- [8] G. Battisti e P. Stoneman. *The intra-firm diffusion of new process technologies*. International Journal of Industrial Organization, 23(1):1-22, 2005.
- [9] A. Bayo-Moriones e F. Lera-López. *A firm-level analysis of determinants of ICT adoption in Spain*. Technovation, 27(6-7):352-366, 2007.
- [10] L. Becchetti, D. Bedoya, e L. Paganetto. *ICT Investment, Productivity and Efficiency: Evidence at Firm Level Using a Stochastic Frontier Approach*. Journal of Productivity Analysis, 20(2):143-167, 2003.
- [11] Bertschek, D. Cerquera, G. J. Klein. *More bits - more bucks? Measuring the impact of broadband internet on firm performance*. Information Economics and Policy, 25(3):190-203, 2013.
- [12] N. Bloom, M. Draca, T. Kretschmer, R. Sadun. *The Economic Impact of ICT*. Centre for Economic Performance, SMART N. 2007/0020, 2010.
- [13] M. Bogers. *The open innovation paradox: knowledge sharing and protection in R&D collaborations*. European Journal of Innovation Management, 14:(1)93-117, 2011.
- [14] Borusyak e J. Xavie. *Revisiting event study designs*. Available at SSRN 2826228, 2017.
- [15] J. Bouckaert, T. Van Dijk, e F. Verboven. *Access regulation, competition, and broadband penetration: An international study*. Telecommunications Policy, 34(11):661-671, 2010.
- [16] T. F. Bresnahan e M. Trajtenberg. *General purpose technologies 'Engines of growth'?* Journal of econometrics, 65(1):83-108, 1995.
- [17] W. Briglauer e K. P. Gugler. *Go for gigabit? First evidence on economic benefits of High-speed Broadband Technologies in Europe*. JCMS: Journal of Common Market Studies, 57: 1071-1090, 2019.
- [18] E. Brynjolfsson e L. Hitt. *Computing productivity: Firm-level evidence*. The Review of Economics and Statistics, 85(4):793-808, 2003.
- [19] M. Bugamelli e P. Pagano. *Barriers to investment in ICT*. Applied Economics, 36(20):2275-2286, 2004.
- [20] C. Cambini e L. Sabatino. *Digital Highways and Firm Turnover: Evidence from Italy*. Unpublished manuscript, 2020.

- [21] C. Cambini, M. R. Ward, e T. Kretschmer. *ICT and Innovation: Editorial*. Information Economics and Policy, 25(3):107-108, 2013.
- [22] M. Cantamessa e F. Montagna. *Management of Innovation and Product Development*. Springer London Ltd, 2016.
- [23] G. Canzian, S. Poy, e S. Schüller. *Broadband upgrade and firm performance in rural areas: Quasi-experimental evidence*. Regional Science and Urban Economics, 77:87-103, 2019.
- [24] G. Colombo, A. Croce, e L. Grilli. *ICT services and small businesses' productivity gains: An analysis of the adoption of broadband Internet technology*. Information Economics and Policy, 25(3):171-189, 2013.
- [25] Copenhagen Economics. *The economic impact of a European digital single market*. Marzo, 2010.
- [26] N. Czernich, O. Falck, T. Kretschmer, e L. Woessmann. *Broadband infrastructure and economic growth*. The Economic Journal, 121(552):505-532, 2011.
- [27] F. Damanpour e M. Schneider. *Characteristics of innovation and innovation adoption in public organizations: Assessing the role of management*. Journal of Public Administration Research and Theory, 19(3):495-522, 2009.
- [28] I. Diaz-Rainey. *Induced Diffusion: Definition, Review and Suggestions for Further Research*. Gennaio 2009.
- [29] M. Dodgson. *Generating technological innovation: E.B. Roberts (Ed.), (Oxford University Press, 1987)*. Research Policy, 18(5):305-306, 1989.
- [30] G. Dosi, C. Freeman, R. Nelson, G. Silverberg, e L. Soete. *Technical Change and Economic Theory*. Laboratory of Economics and Management (LEM), Sant'Anna School of Advanced Studies, Pisa, Italy, 1988.
- [31] B. Engelstätter e M. Sarbu. *Enterprise Software and Service Innovation: Standardization Versus Customization*. ZEW - Centre for European Economic Research, Discussion Paper No. 10-100, 2011.
- [32] S. Freel. *Barriers to product innovation in small manufacturing firms*. International Small Business Journal, 18(2):60-80, 2000.
- [33] J. J. Ganuza e M. F. Vicens. *Deployment of high-speed broadband infrastructures during the economic crisis. The case of Xarxa Oberta*. Telecommunications Policy, 35(9-10):857-870, 2011.
- [34] F. W. Geels. *Processes and patterns in transitions and system innovations: Refining the co-evolutionary multi-level perspective*. Technological Forecasting and Social Change, 72(6):681-696, 2005.
- [35] P. Giuri, S. Torrisi, e N. Zinovyeva. *Ict, skills, and organizational change: evidence from italian manufacturing firms*. Industrial and Corporate change, 17(1):29-64, 2008.
- [36] A. Goodman-Bacon, *Difference-in-differences with variation in treatment timing*. Journal of Econometrics, ISSN 0304-4076, 2021.
- [37] Z. Griliches. *Patent Statistics as Economic Indicators: A Survey*. National Bureau of Economic Research Working Paper Series, 28(4):1661-1707, 1990.
- [38] A. Grimes, C. Ren, e P. Stevens. *The need for speed: impacts of internet connectivity on firm productivity*. Journal of Productivity Analysis, 37(2):187-201, 2012.
- [39] L. Grzybowski, M. Hasbi, e J. Liang. *Transition from copper to fiber broadband: The role of connection speed and switching costs*. Information Economics and Policy, 42:1-10, 2018.
- [40] M. Hasbi. *Impact of very high-speed broadband on Local Economic growth: Empirical evidence*. International Telecommunications Society (ITS), 15(3):586-625, 2017.

- [41] T. Hempell. *What's spurious, what's real? Measuring the productivity impacts of ICT at the firm-level*. Empirical Economics, 30(2):427-464, 2005.
- [42] T. Hempell e T. Zwick. *New technology, work organization, and innovation*. Economics of Innovation and New Technology, 17(4): 331-354, 2008.
- [43] P. Koellinger. *The relationship between technology, innovation, and firm performance - empirical evidence from e-business in Europe*. Research Policy, 37:1317-1328, 2008.
- [44] T. Koonen. *Fiber to the Home/Fiber to the Premises: What, Where, and When?* Proceedings of the IEEE, 94(5):911-934, 2006.
- [45] S. Lachenmaier e L. Wosmann. *Does innovation cause exports? Evidence from exogenous innovation impulses and obstacles using German micro data*. Oxford Economic Papers, 58(2):317-350, 2006.
- [46] D. McCoy, S. Lyons, E. Morgenroth, D. Palcic, e L. Allen. *The impact of local infrastructure on new business establishments*. MPRA Paper 69074, University Library of Munich, Germany. 2016.
- [47] M. Morikawa. *Information technology and the performance of Japanese SMEs*. Small Business Economics, 23(3):171-177, 2004.
- [48] R. P. Oakey, A. T. Thwaites e P. A. Nash. *The regional distribution of innovative manufacturing establishment in Britain*. Regional Studies, 14(3):235-253, 1980.
- [49] A. Ollo López e M. E. Aramendia-Muneta. *ICT impact on competitiveness, innovation and environment*. Telematics and Informatics, 29(2):204-210, 2012.
- [50] A. Pakes e Z. Griliches. *Estimating Distributed Lags in Short Panels with an Application to the Specification of Depreciation Patterns and Capital Stock Constructs*. The Review of Economic Studies, 51(2):243-262, 1984.
- [51] C. Parker e T. Castleman. *New directions for research on SME-eBusiness: Insights from an analysis of journal articles from 2003 to 2006*. Journal of Information Systems and Small Business, 1(1):21-40, 2007.
- [52] B. Pellegrino e L. Zingales. *Diagnosing the Italian Disease*. Working Paper 23964, National Bureau of Economic Research, 2017.
- [53] M. Polder, G. Leeuwen, P. Mohnen, e W. Raymond. *Product, Process and Organizational Innovation: Drivers, Complementarity and Productivity Effects*. MERIT Working Paper (2010-035), 2010.
- [54] M. Rogers. *Networks, Firm Size and Innovation*. Small Business Economics, 22:141-153, 2004.
- [55] K. Rohman, e E. Bohlin. *Does broadband speed really matter as a driver of economic growth? Investigating OECD countries*. International Journal of Management and Network Economics, 2(4):336-356, 2012.
- [56] G. L. Rosston, S. Savage, e D. M. Waldman. *Household Demand for Broadband Internet Service*. Discussion Papers from Stanford Institute for Economic Policy Research, No 09-008, 2010.
- [57] F. M. Scherer. *Industrial Market Structure and Economic Performance*. American Journal of Agricultural Economics, 54(3):547-548, 1972.
- [58] F. Schivardi e T. Schmitz. *The IT Revolution and Southern Europe's Two Lost Decades*. Journal of the European Economic Association, 18(5):2441-2486, 2019.
- [59] R. Swinney, G. P. Cachon, e S. Netessine. *Capacity Investment Timing by Start-ups and Established Firms in New Markets*. Management Science, 57(4):763-777, 2011.
- [60] I. Takahashi, H. Tatemichi, T. Tanaka, S. Nishi, e T. Kunioka. *Environmental impact of information and communication technologies including rebound effects*. IEEE International Symposium on Electronics and the Environment, 13-16, 2004.

- [61] J. Vilaseca-Requena, J. Torrent-Sellens, e A. I. Jimenez-Zarco. *ICT Use in Marketing as Innovation Success Factor: Enhancing Cooperation in New Product Development Processes*. European Journal of Innovation Management, 10(2):268-288, 2007.
- [62] J. Wiens e C. Jackson. *The Importance of Young Firms for Economic Growth*. Kauffman Foundation, 2014.
- [63] S. Wymer e E. Regan. *Factors influencing e-commerce adoption and use by small and medium businesses*. Electronic Markets, 15(4):438-453, 2005.

Capitolo 2 - Analisi descrittiva

L'obiettivo di questo capitolo è analizzare i dati che saranno poi utilizzati nello studio econometrico finale. Durante lo sviluppo del modello, sono stati analizzati tre diversi dataset per l'arco temporale 2013-2019: il primo riguarda le startup innovative fornito dalla Camera di Commercio, il secondo contiene le registrazioni brevettuali a livello provinciale fornite dall'Ufficio Italiano Brevetti e Marchi (UIBM), infine il terzo rappresenta il cuore della studio e raccoglie informazioni granulari sulla diffusione della banda larga ultraveloce a livello comunale, fornito da Telecom Italia Mobile (TIM). Ogni comune presente è stato inoltre caratterizzato dall'eventuale presenza di università e grado di urbanizzazione (fonte ISTAT). I tre dataset sono stati poi aggregati in un unico file di lavoro, raggiungendo in questo modo 9886 osservazioni e 49 variabili.

2.1 Startup innovative

I dati forniti sulle startup innovative riguardano la ragione sociale, provincia, comune¹¹, codice ATECO, le date di iscrizione al registro delle imprese, alla sezione speciale e dell'effettivo inizio dell'attività nell'arco temporale dal 2013 al 2020. Le imprese sono inoltre caratterizzate dall'appartenenza a un settore specifico (agricoltura/pesca, industria/artigianato, commercio, servizi e turismo). Dopo alcune modifiche¹² al database, per ogni comune si è specificato il grado di urbanizzazione¹³, la posizione geografica¹⁴, le registrazioni brevettuali annue e eventuale presenza di polo universitario.

In un primo momento, la situazione è stata analizzata dall'alto suddividendo la nascita di nuove startup innovative nell'arco dei 7 anni. Gli andamenti considerando come variabili le date di "registro", "sezione" e "esercizio" non presentano particolari differenze, com'è possibile vedere dalla [Figura 2-1](#) in cui si riporta un quadro generale della situazione italiana. Al termine del 2019, 9916 startup complessive hanno ottenuto l'iscrizione. Le startup godono di particolari benefici nel caso soddisfino alcuni requisiti e riescano ad ottenere l'iscrizione alla sezione speciale. Tuttavia, tale processo non è immediato e può richiedere mesi o in casi particolari addirittura anni. Fino al termine del 2017 il numero di imprese che è riuscito ad ottenere tale agevolazione è minore del totale, per poi toccare un picco di 2968 unità nell'ultimo anno e attestarsi complessivamente a 9467 nell'intero

¹¹ Nel 2019 il territorio italiano è suddiviso in 107 province e 7914 comuni. I comuni in cui si è iscritta almeno una startup innovativa dal 2013 al 2019 sono 1660.

¹² Si è escluso dal database alcune imprese poiché aventi sede all'estero (*Uniwere Ug, Kaleao Limited, Amavido GmbH, Popsa holdings LTD, Inesse Corporation LTD, Wishraiser LTD, Deephound LTD, U-Earth biotech LTD, Geouniq sas, Europe trade LTD, Henshin group LTD*), rimosso il codice postale, il codice ateco2007 e NACE sono stati uniti, rimossa denominazione NACE poiché ridondante con la voce "attività", rimosse le date e aggiunto anno relativo ad ognuna, aggiunta posizione geografica, aggiunto il grado di urbanizzazione (fonte Eurostat), aggiunta eventuale presenza di università per ogni comune e numero di brevetti annui registrati per ogni provincia

¹³ Dal 2011 Eurostat classifica i comuni secondo tre gradi di urbanizzazione: alta, media e bassa basandosi sulla densità demografica e il numero di abitanti valutati entro griglie regolari con celle di un chilometro quadrato. In Italia risulta che il 67,9% dei comuni ricade nella classe di bassa urbanizzazione, area prevalentemente rurale, dove su una superficie del 72,5% si localizza una popolazione pari al 24,3%. Nei comuni ad alta urbanizzazione, che rappresentano solo il 3,3% del totale nazionale e con una superficie territoriale complessiva del 4,8%, è presente il 33,3% della popolazione italiana. Nel restante 28,7% dei comuni di grado medio di urbanizzazione, su un'estensione territoriale del 22,7%, si concentra il 42,4% della popolazione complessiva. Al termine del 2019, 187 comuni sono definiti ad alta urbanizzazione (1), 996 a media (2) e 475 a bassa (3).

¹⁴ L'Italia, in accordo con gli studi statistici dell'Istat, è stata divisa in tre zone: nord (Piemonte, Valle d'Aosta, Lombardia, Trentino-Alto Adige, Veneto, Friuli-Venezia Giulia, Liguria, Emilia-Romagna), centro (Toscana, Umbria, Marche, Lazio) e sud (Abruzzo, Molise, Campania, Puglia, Basilicata, Calabria, Sicilia, Sardegna).

periodo considerato. Concludendo, il 4,53% delle nuove startup innovative non ha ancora ottenuto l'iscrizione alla sezione speciale del RI al termine del 2019. L'effettivo avvio dell'attività economica invece può sia essere posticipata o anticipata rispetto l'iscrizione al Registro delle Imprese. Si evince infine che lo 0,12% delle startup innovative, che ha iniziato l'esercizio dell'attività economica, non abbia ancora ottenuto la certificazione dei loro dati costitutivi.

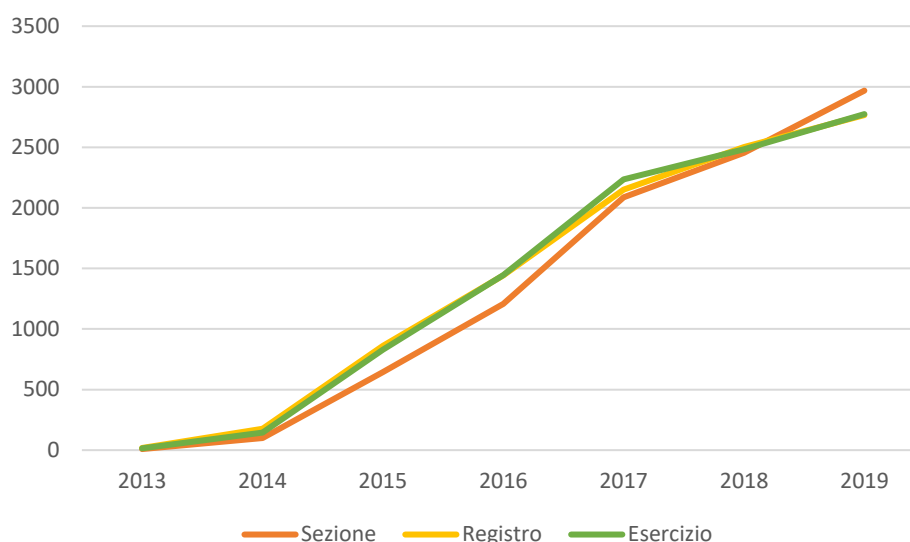


FIGURA 2-1 ANDAMENTO ISCRIZIONE AL REGISTRO DELLE IMPRESE, A SEZIONE SPECIALE E AVVIO ATTIVITÀ

Vista la stretta vicinanza tra le curve tracciate seguendo le tre date diverse, d'ora in avanti utilizzeremo l'iscrizione al registro delle imprese come variabile principale per analizzare la diffusione delle startup innovative e anche quando si parlerà di "nascita di startup" in modo generico faremo riferimento a questo valore.

Nel 2013 e 2014 un numero limitato di nuove startup sono presenti nella penisola italiana. Si sono verificate 195 iscrizioni al registro delle imprese e solo 160 hanno iniziato l'esercizio dell'attività economica. Dall'anno successivo la variazione marginale aumenta e si verifica una crescita pari al 485,87% rispetto al 2014, ovvero 860 nuove realtà innovative nascono nella penisola. Negli anni successivi la crescita marginale aumenta ancora fino a raggiungere l'iscrizione di 2766 startup innovative nel 2019.

2.1.1 Suddivisione geografica

Fino al 2014 il centro-sud era l'area geografica italiana col maggior numero di startup innovative ([Figura 2-2](#)); all'interno vi erano iscritte complessivamente 144 imprese di questo tipo contro le 51 nel nord. Successivamente si è verificata una forte inversione di tendenza che ritroviamo ancora oggi. Dal 2015 al 2019 le startup fondate annualmente nell'Italia settentrionale rappresentano infatti più della metà del totale, come evidenziato nella [Tabella 2-1](#).

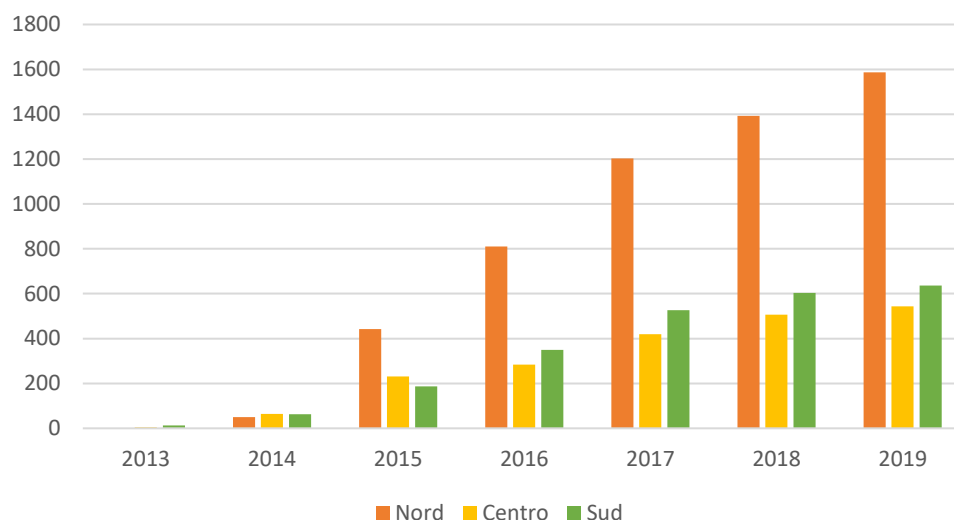


FIGURA 2-2 ISCRIZIONE AL REGISTRO DELLE IMPRESE

TABELLA 2-1 RAPPORTO TRA STARTUP INNOVATIVE ISCRITTE OGNI ANNO NEL NORD ITALIA E NELL'INTERA PENISOLA

2015	51.39%
2016	56.13%
2017	55.95%
2018	55.67%
2019	57.37%

Dopo una fase di stagnazione fino a inizio 2015, si manifesta infatti una presenza sempre più consistente nelle regioni settentrionali. Per quest'area geografica la funzione delle nuove imprese iscritte nel tempo è crescente ma la crescita marginale diminuisce in modo evidente nel 2017. La [Tabella 2-2](#) mette in luce la variazione marginale della crescita annua delle iscrizioni nel nord Italia. Anche il massimo incremento è nel 2015, il numero complessivo annuale continua, comunque, ad aumentare e al termine del 2019 sono state registrate complessivamente 5486 startup innovative.

TABELLA 2-2 CRESCITA MARGINALE ANNUA NORD ITALIA

2014	+49
2015	+392
2016	+368
2017	+393
2018	+190
2019	+194

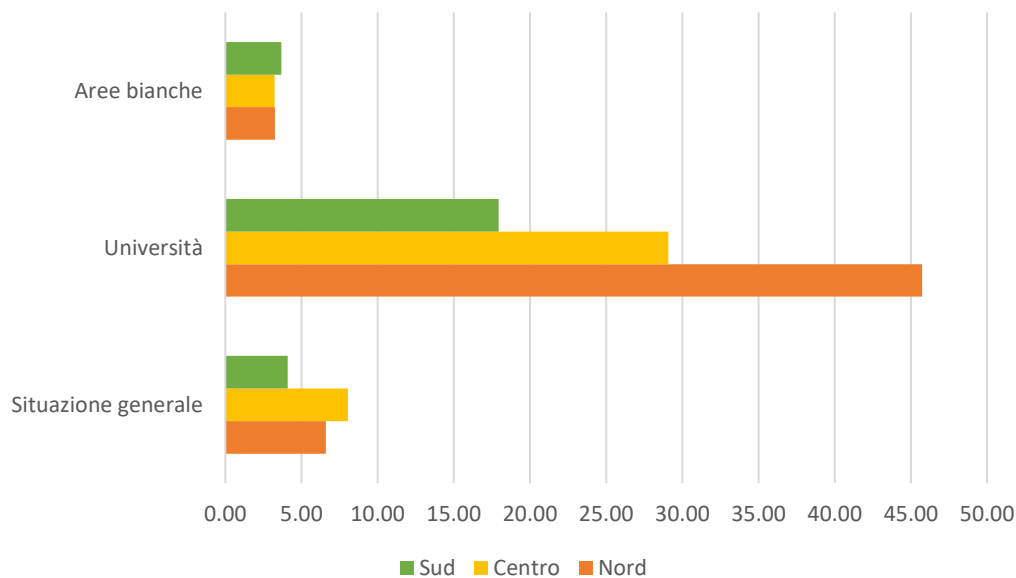


FIGURA 2-3 STARTUP MEDIE PER COMUNE

Una variabile interessante potrebbe essere le startup medie per comuni (Figura 2-3). È importante tener conto che in questo calcolo, al denominatore, si è incluso soltanto i comuni in cui è presente almeno una startup. Dunque, a causa di una diffusione capillare nel settentrione, le unità territoriale considerate di quest'area risultano molto maggiori rispetto al centro e sud Italia (rispettivamente 828, 254 e 578). Per questo motivo, si ricava che le



FIGURA 2-4 CARTINA ITALIA CON FOCUS SU STARTUP INNOVATIVE

startup medie per ogni comune siano superiori al centro, pari a 8,05, rispetto al nord (6,61) e sud (5,78). Se invece si considerassero solo le unità territoriali in cui è presente almeno un'università, i valori riscontrati sono molto superiori e il nord vanta 45,73 startup per ogni comune. Nelle aree bianche invece le startup presenti sono in residua quantità, da notare che il calo netto si verifica al centro-nord mentre il sud continua a presentare un numero di startup vicino alla situazione generale.

Nella **Errore. L'origine r** iferimento non è stata trovata. si riporta un

impatto visivo qualitativo sulla

diffusione delle startup nella penisola.

La dimensione delle sfere è proporzionale alle iscrizioni complessive al termine del 2019. Quattro regioni settentrionali (Piemonte, Lombardia, Veneto, Emilia-Romagna) hanno sul proprio territorio il 48,8% delle startup italiane.

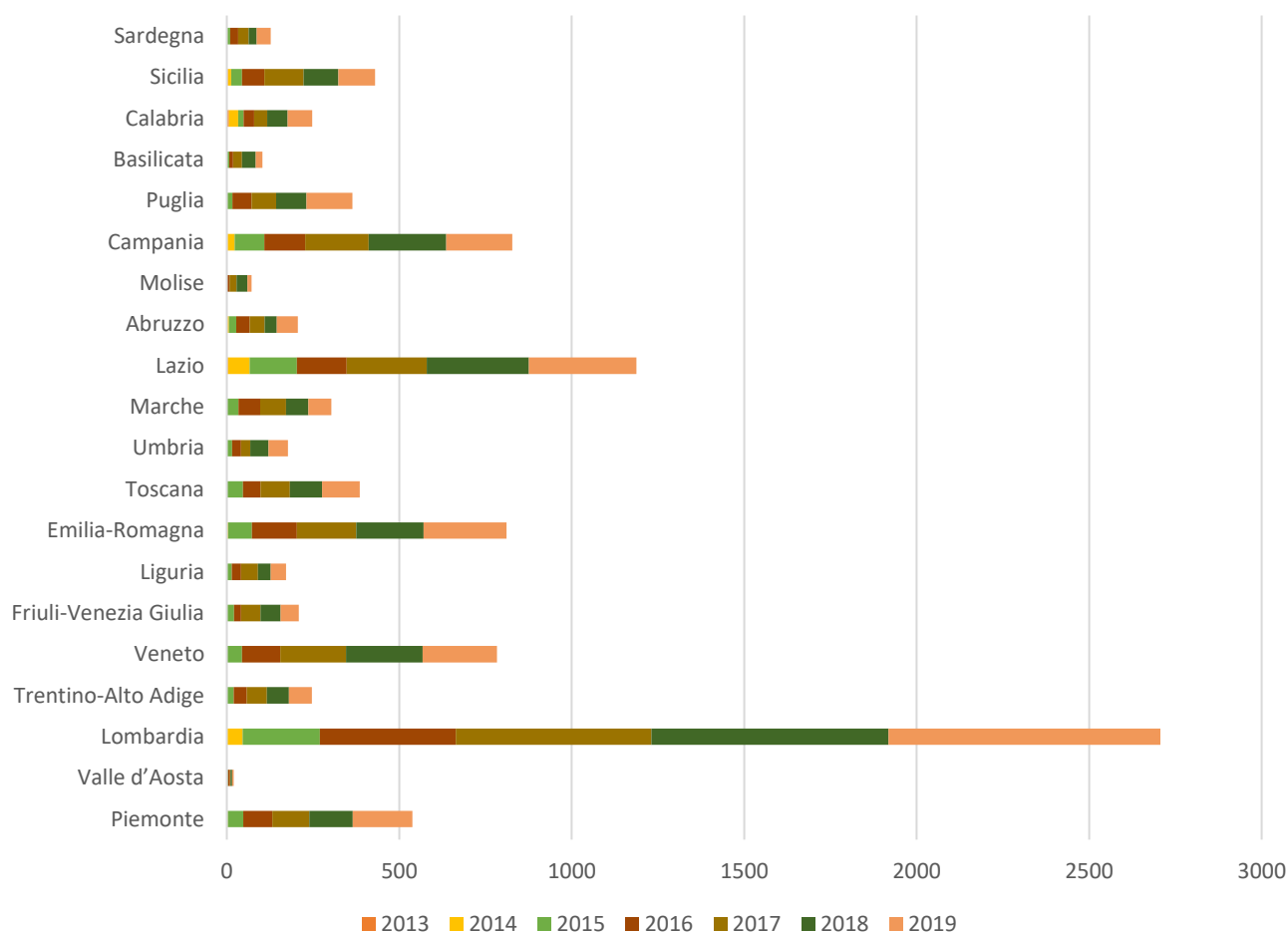


FIGURA 2-5 STARTUP INNOVATIVE ISCRITTE AL REGISTRO DELL'IMPRESE SUDDIVISE PER REGIONE DI APPARTENENZA

La [Figura 2-5](#) rappresenta le nuove startup iscritte suddivise tra le regioni di appartenenza ed evidenzia le crescite annue.

Le prime tre regioni per numerosità complessiva, al termine del 2019, risultano essere Lombardia, Lazio e Campania, con rispettivamente 2706, 1187, 828 startup. La Lombardia è nettamente la più prolifica e il totale delle iscrizioni presenti in questo territorio rappresenta il 27,29% del totale. Ognuna di queste tre regioni è rappresentante di un'area geografica diversa. A seguire troviamo Veneto (783) e Emilia-Romagna (811), ciò concorda con i risultati precedenti che mostravano la marcata diffusione concentrata nel nord Italia rispetto al resto della penisola.

In particolare, osservando le 3 regioni più "fertili" ([Figura 2-6](#)), si può notare come nei primi 2 anni la Lombardia non sia prima per numerosità. Dal 2015 in cui si verificano 224 iscrizioni, ovvero 1 unità in più rispetto alla somma di Veneto e Campania, raggiunge il

primato. Negli anni successivi il tasso di crescita è ancora più elevato e tocca il picco nel 2019 con l'iscrizione di 788 startup innovative. La variazione della crescita marginale è sempre positiva e negli ultimi 5 anni costantemente maggiore di 100 unità. Le altre due regioni non raggiungono mai una tale variazione e tra alti e bassi la Campania subisce una diminuzione della crescita marginale nel 2019.

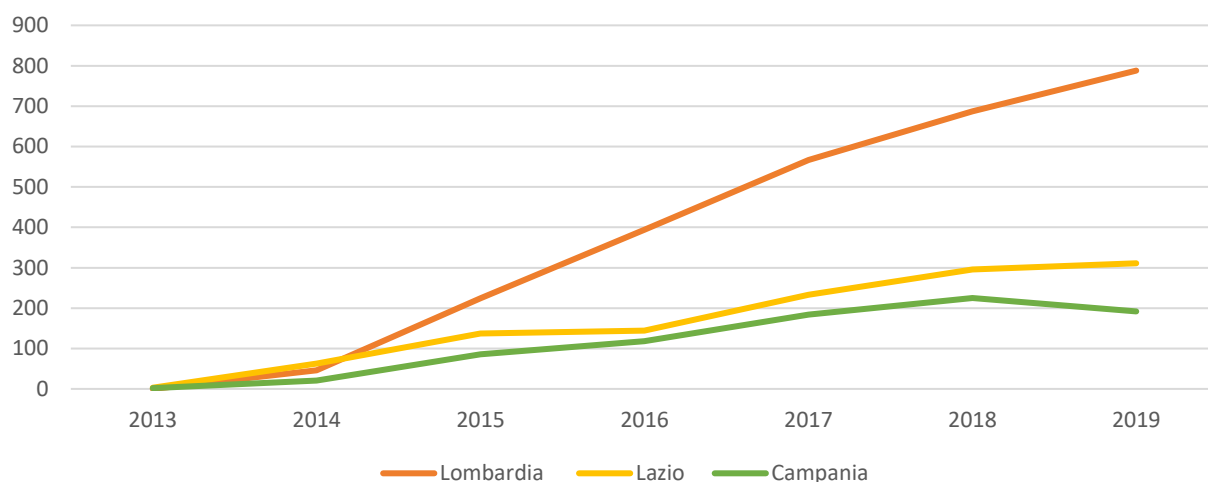


FIGURA 2-6 FOCUS SU LOMBARDIA, LAZIO E CAMPANIA

Focalizzandoci poi sulle 7 regioni con più di 400 startup innovative complessivamente iscritte (Figura 2-7), si è analizzato le variazioni percentuali delle iscrizioni annue dal 2016 al 2019.

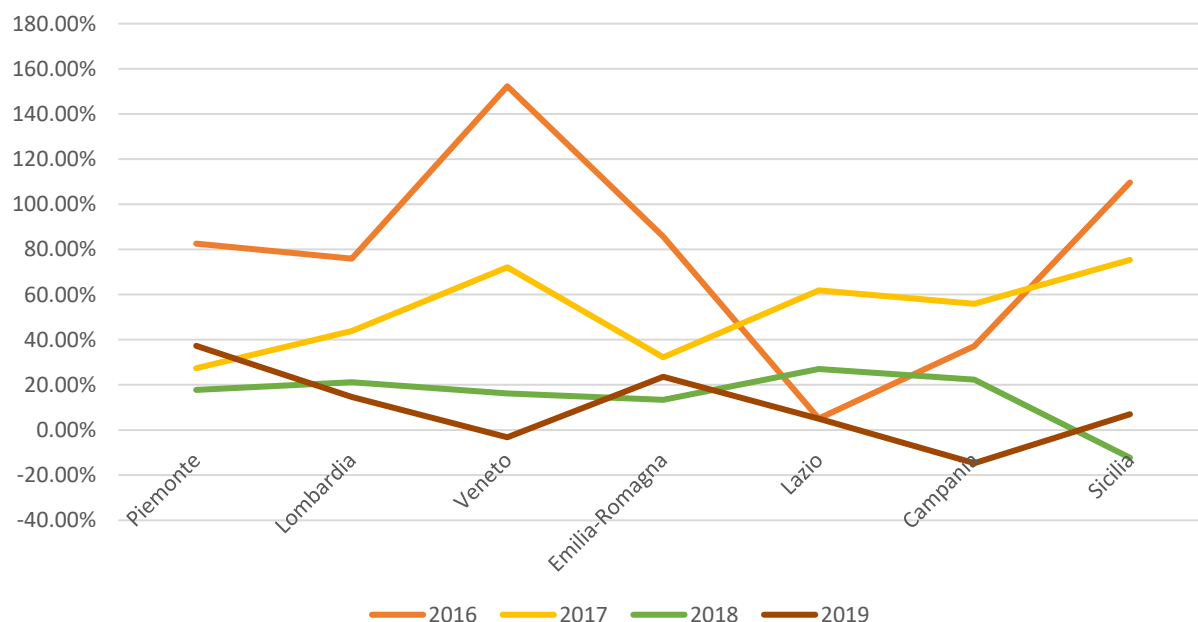


FIGURA 2-7 VARIAZIONE PERCENTUALE CRESCITA STARTUP INNOVATIVE

I primi anni sono chiaramente caratterizzati da una variazione percentuale maggiore poiché la situazione precedente presentava una numerosità bassa (1055 startup iscritte complessivamente in tutta Italia). Veneto, Campania e Sicilia osservano una variazione dell'incremento percentuale negativo in un anno, questo rallentamento non si verifica nelle altre regioni citate in cui ogni anno aumentano le nascite di startup. Il Veneto è la regione che presenta un intervallo percentuale maggiore passando dal +152,27% nel 2016 al -3,15% nel 2019.

La diffusione dell'innovazione è legata a particolari situazioni ambientali e ciò incentiva la creazione di società innovative intorno a pochi centri particolarmente attrattivi e tende a ostacolare la diffusione capillare tra le varie città della stessa regione. Ci focalizziamo sulle province in cui almeno 175 startup si sono registrate durante i sette anni e studiamo l'andamento nell'ultimo biennio (Figura 2-8). Milano è molto avanti rispetto alle altre città della penisola, contiene infatti al suo interno il 59,1% delle 2396 startup nate nelle 11 province analizzate. Roma è molto distante ma raggiunge 756 e si conferma il secondo polo innovativo italiano. Poi si trova Napoli, il cui trend è leggermente in calo. Anche Torino e Bologna superano 200 iscrizioni negli ultimi 3 anni ma solo nella prima citata si sta verificando una crescita in costante aumento. Brescia e Bari, inizialmente lontane per numerosità dalle altre province presenti nel grafico, hanno avuto una crescita esponenziale nel 2019, rispettivamente al 89% e 78% raggiungendo 146 e 130 iscrizioni.

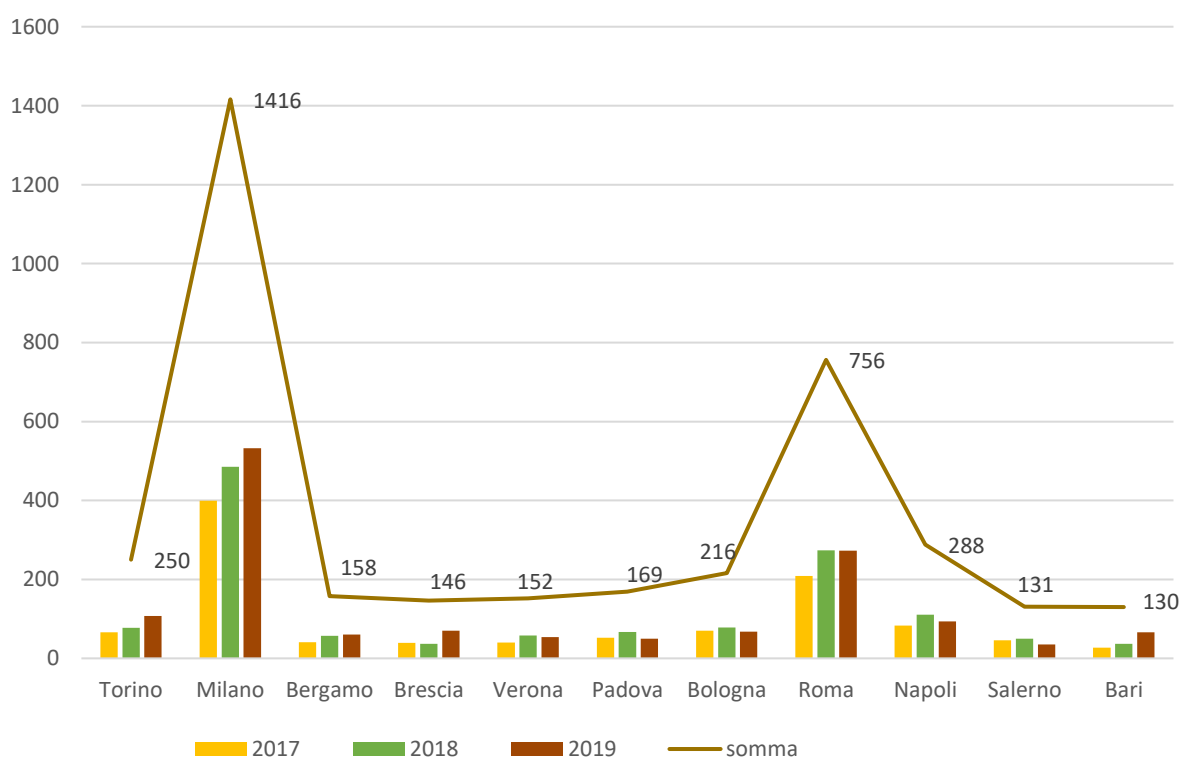


FIGURA 2-8 FOCUS SU PROVINCE PIÙ PROLIFICHE NEL BIENNIO 2017-2019

2.1.2 Suddivisione settoriali

Al momento dell'iscrizione al registro delle imprese, ad ogni startup viene attribuito un settore con la denominazione NACE. Per l'analisi si è individuato cinque macrosettori:

- Servizi
- Turismo
- Industria/Artigianato
- Agricoltura/Pesca
- Commercio

Sono presenti inoltre 17 imprese non classificate poiché in seguito all'iscrizione non è sempre possibile attribuire loro un determinato settore industriale. Possono essere considerate appartenenti a una classe di transizione: dopo che la startup avrà perfezionato la propria posizione, sarà poi assegnata a un settore specifico. Quindi le startup complessivamente assegnate risultano essere pari a 9899, suddivise come nella [Figura 2-9](#). Vista la maggior numerosità nell'industria/artigianato e soprattutto servizi, andiamo ad analizzare più nel profondo questi due settori. Si è poi studiato l'andamento delle iscrizioni al registro delle imprese nel corso del tempo per le tre aree geografiche suddividendo per settore di appartenenza.

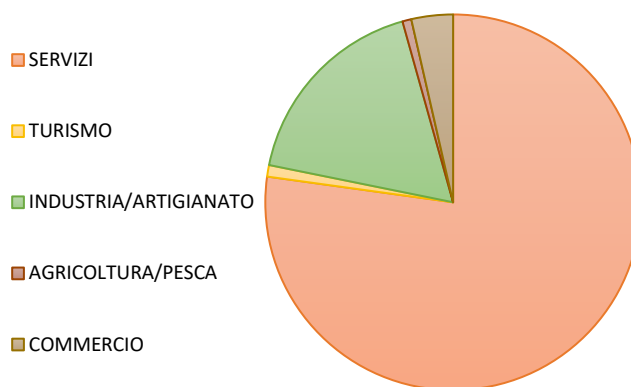


FIGURA 2-9 SUDDIVISIONE STARTUP PER SETTORE

L'andamento delle startup nel settore dei servizi non si discosta significativamente rispetto alla crescita complessiva. Dal 2013 al 2019 sono nate 7641 startup innovative in questo settore, pari al 77,2% del totale.

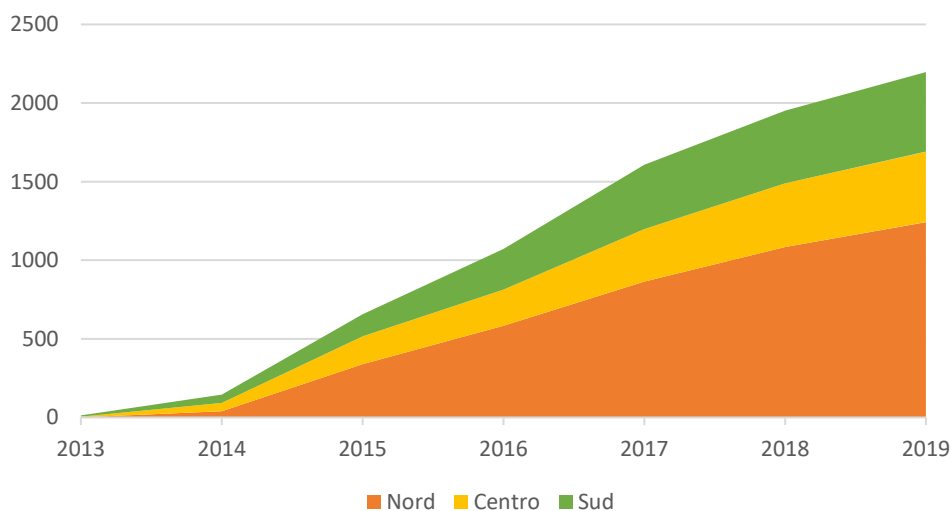


FIGURA 2-10 ANDAMENTO NASCITA STARTUP SETTORE DEI SERVIZI

Nell'industria/artigianato, le startup hanno avuto un tasso di crescita lineare e positivo, coerente con l'andamento generale, fino al 2017. Da questo anno, si è poi verificato un leggero rallentamento nell'Italia settentrionale. Questo lieve calo di tendenza nella zona trainante è stato tuttavia compensato dall'aumento della crescita marginale nel centro-sud. Inoltre, l'andamento complessivo non risente di questo rallentamento poiché nel settore dei servizi, il cui peso supera i $\frac{3}{4}$ del totale, si verifica un continuo incremento anche nell'Italia settentrionale.

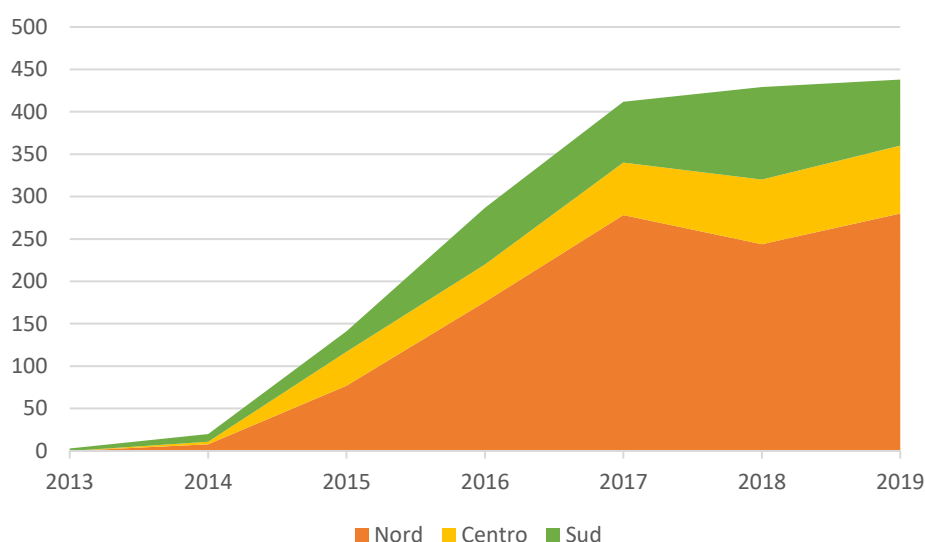


FIGURA 2-11 ANDAMENTO NASCITA STARTUP SETTORE INDUSTRIA/ARTIGIANATO

TABELLA 2-3 NUOVE STARTUP INNOVATIVE ISCRITTE PER SETTORE

	Servizi	Industria/Artigianato	Altri settori
2013	72.22%	16.67%	11.11%
2014	82.86%	11.43%	5.71%
2015	76.48%	16.41%	7.10%
2016	74.32%	19.92%	5.76%
2017	74.78%	19.17%	6.05%
2018	78.20%	17.19%	4.61%
2019	79.54%	15.86%	4.60%

Adesso analizziamo come sono variati i relativi pesi percentuali nel corso degli anni (Tabella 2-3). I tre settori meno impattanti (turismo, commercio e agricoltura/pesca) sono stati aggregati nella voce "Altri settori". Si verificano cambiamenti evidenti dal 2013 al 2019 per i diversi settori: i servizi raggiungono il 79,54% aumentando il proprio peso percentuale di circa 7,3 punti; l'industria artigianato dopo un leggero aumento a metà del periodo considerato, subisce un lieve calo e registra infine il 15,86%; il resto dei settori subisce invece una netta diminuzione di 6,5 punti percentuali, passando da 11,11% a 4,6% nei 7 anni presi in esame.

2.1.3 Suddivisione per grado di urbanizzazione

Il grado di urbanizzazione¹³ basandosi sulla densità di popolazione e sulla contiguità fra aree, classifica il territorio in tre tipologie: alta (1), media (2) e bassa urbanizzazione (3). I 187 comuni in cui è nata almeno una startup dal 2013 al 2019 sono definiti ad alta urbanizzazione, 996 a media e 475 a bassa.

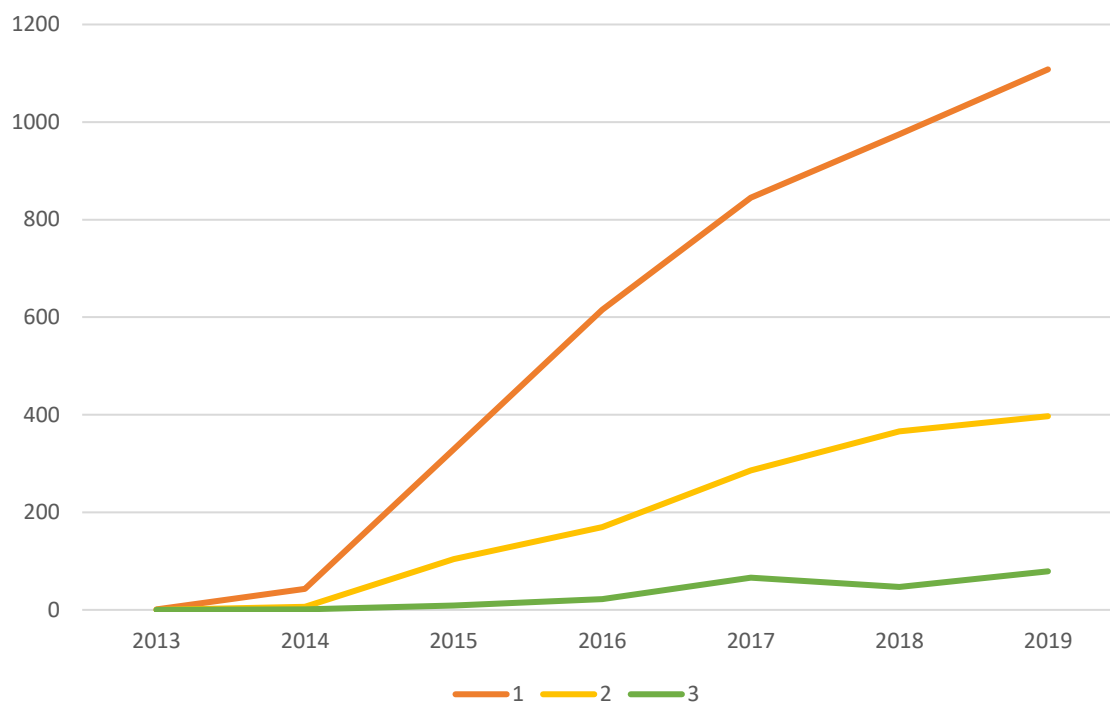


FIGURA 2-12 STARTUP SUDDIVISE PER GDU

Al termine del 2019, nelle zone ad alta urbanizzazione si riscontra il 67% delle iscrizioni complessive di nuove startup e il 26,77% in quelle a media. Le zone a bassa non raggiungono mai quota 200 iscrizioni annue. Concentrandosi sul Nord Italia, zona più ricca di startup innovative, il peso delle zone densamente popolate è ancora superiore: 3916 delle 5469 startup iscritte si trovano in comuni di questa categoria.

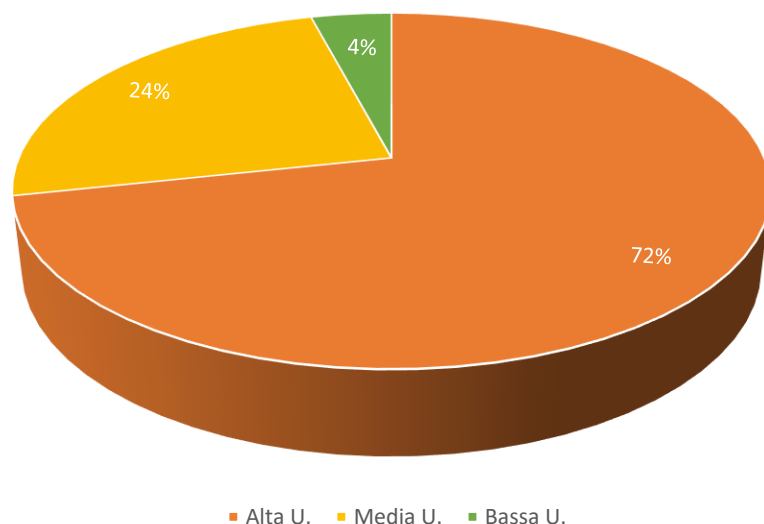


FIGURA 2-13 STARTUP NEL NORD ITALIA SUDDIVISE PER GDU

Al contrario al sud, il peso percentuale delle unità territoriali a medio-basso grado di urbanizzazione è maggiore contenendo complessivamente quasi il 43% delle startup innovative. Per quanto riguarda le aree prevalentemente rurali, nel sud si verificano 252 iscrizioni e nel nord 224; la spinta dell'Italia settentrionale è infatti legata principalmente alle zone a urbanizzazione elevata, in cui sono sorte complessivamente quasi il triplo di startup rispetto al meridione (3916 al nord e 1353 al sud).

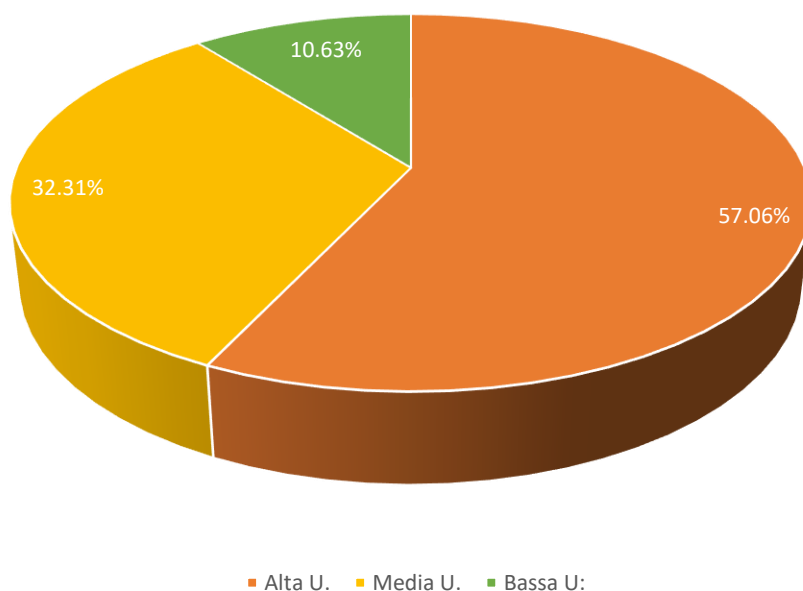


FIGURA 2-14 STARTUP NEL SUD ITALIA SUDDIVISE PER GDU

2.1.4 Analisi crescita startup nel 2020

Infine, per quanto riguarda lo studio descrittivo delle startup innovative si effettua una breve analisi di ciò che è successo nell'ultimo anno trattato nel dataset a disposizione, ovvero il 2020. Nella stima empirica non si sono inclusi tali variabile per due motivi principali: i dati riguardanti la diffusione della banda larga ultraveloce forniti da Telecom Italia riguardano l'arco temporale dal 2013 al 2019; inoltre, la pandemia Covid-19, nel 2020, ha generato una stagnazione della crescita della produttività su scala mondiale e tale esogeneità avrebbe presumibilmente impattato fortemente sui risultati empirici ottenuti.

Come per gli altri anni studiati, si analizzano le iscrizioni al Registro delle imprese startup effettuando una suddivisione geografica e settoriale.

Nella [Figura 2-15](#), è riportato il totale delle startup innovative nascenti al Nord, Centro e Sud Italia. L'area geografica col maggior incremento è il Settentrione in cui si registrano 1162 iscrizioni; mentre il Centro e Sud contano rispettivamente 405 e 520 nuove startup innovative. Dal 2015 il rapporto medio di iscrizioni nell'Italia settentrionale rispetto al totale era pari a 55,31% raggiungendo il picco nel 2019 con 57,38%. Nel 2020 tale rapporto risulta essere pari a 55,68%.

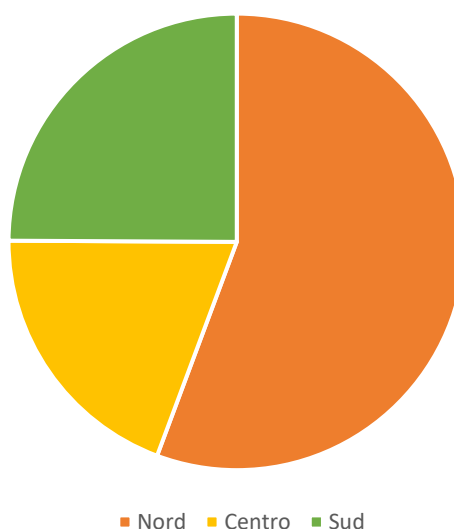


FIGURA 2-15 NUOVE STARTUP INNOVATIVE ISCRITTE AL REGISTRO DELLE IMPRESE NEL 2020

Nella suddivisione a livello regionale ([Figura 2-16](#)), le prime tre regioni per risultano Lombardia, Lazio e Veneto, la quale con 191 iscrizioni supera la Campania con 179, che risultava essere la terza più prolifica nell'intero arco temporale studiato ([Figura 2-5](#)). L'incremento di startup medio per regione è pari a 104.35, ciò evidenzia una diminuzione del 25% rispetto al 2019 in cui si erano registrano 138.3 iscrizioni medie. Oltre alla Lombardia che raggiunge quota 607, le altre che superano la media nazionale sono Veneto, Emilia-Romagna, Toscana, Lazio, Campania e Puglia.

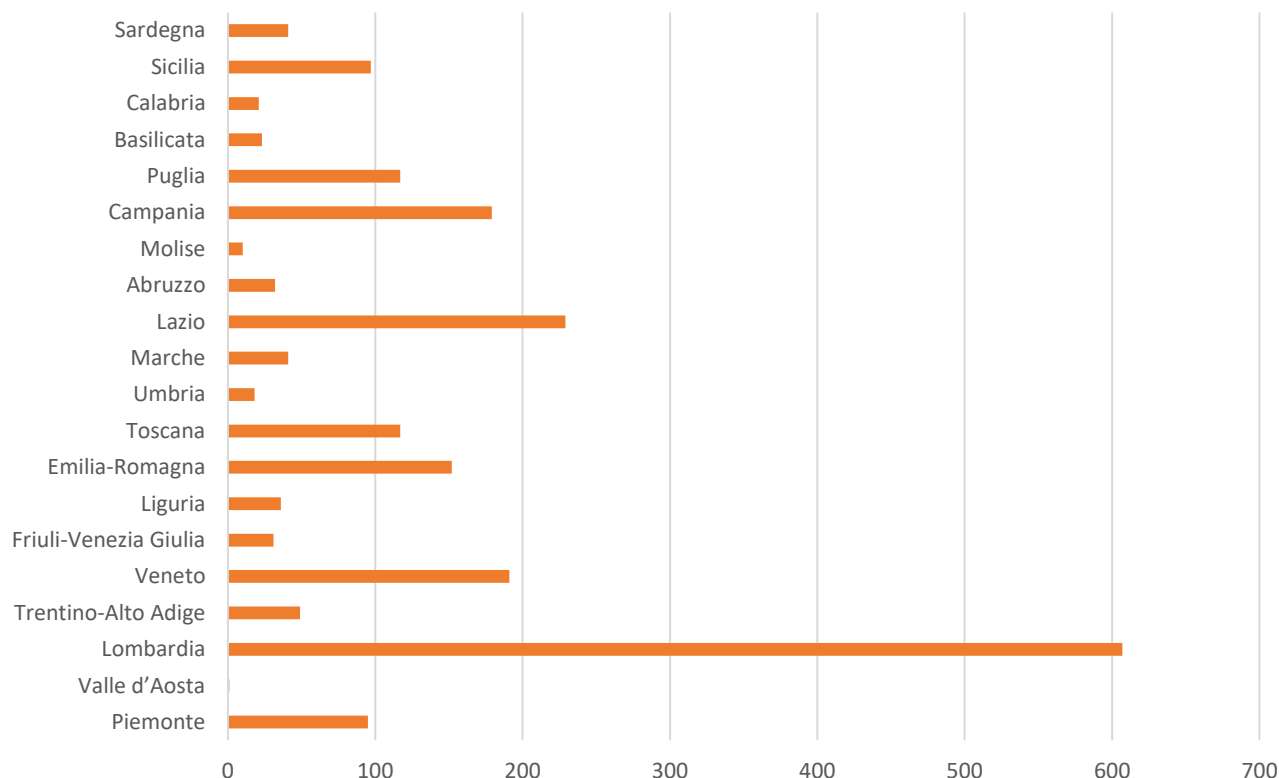


FIGURA 2-16 STARTUP INNOVATIVE ISCRITTE AL RI NEL 2020, SUDDIVISE PER REGIONE DI APPARTENENZA

Focalizzandoci adesso sulle province che hanno riscontrato la maggior crescita di startup innovative nel 2020 ([Figura 2-17](#)) è possibile fare un confronto con quanto riscontrato precedentemente per quanto riguarda il triennio 2017-2019. Salerno e Verona non fanno più parte degli undici enti locali in cui si è verificato il maggior incremento: al loro posto, si posizionano dunque Firenze e Palermo rispettivamente con 39 e 37 nuove startup innovative.

La media di iscrizioni nelle undici province più prolifiche risulta pari a 101.73, minore dell'11.5% rispetto alle 115 iscrizioni medie rispetto al triennio precedente. Le uniche città in cui la crescita è aumentata sono Padova (+0.66%) e Bari (+43.08%). Le maggiori variazioni negative si sono riscontrate a Bologna (-37.5%), Torino (-26.80%), Brescia (-23.97%) e Roma (-23.81%).

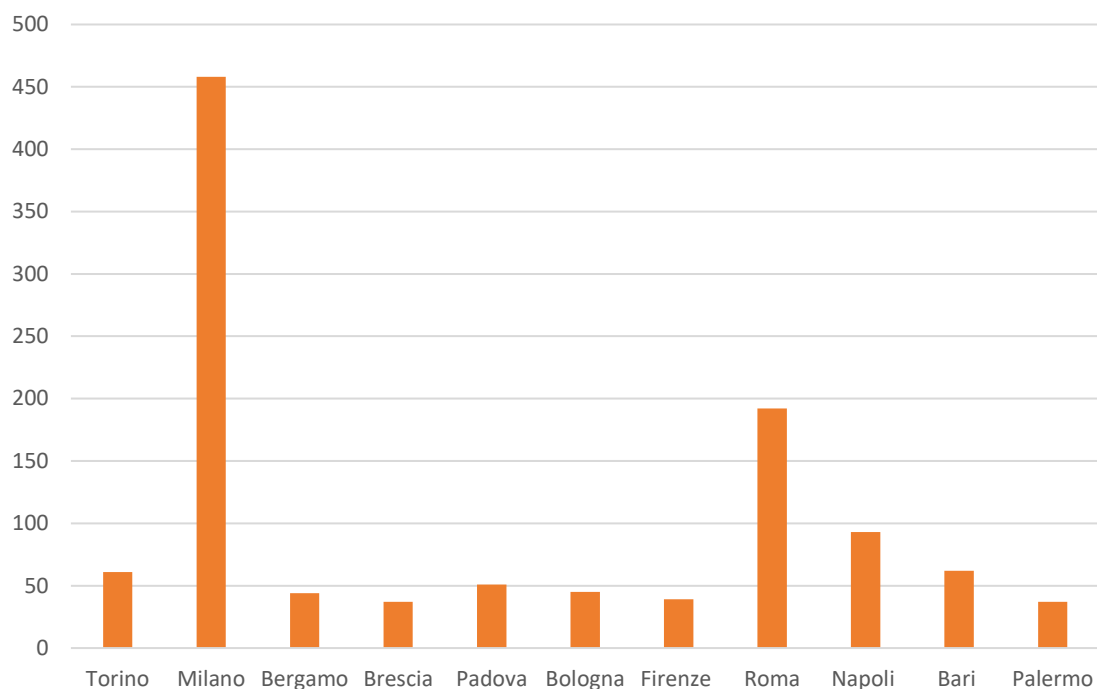


FIGURA 2-17 FOCUS SU PROVINCE PIÙ PROLIFICHE NEL 2020

In linea con quanto ottenuto studiando i periodi precedenti ([Tabella 2-3](#)), la maggior parte delle startup innovative mirano a fornire servizi. Infatti, il 79.33% fa parte di questo settore e solo il 20.67% ha ottenuto una classificazione diversa. Nella [Figura 2-18](#) sono riportati i valori in termini assoluti, includendo la categoria dell'industria e artigianato poiché è il secondo settore con maggiori registrazioni, raggiungendo il 16.42%.

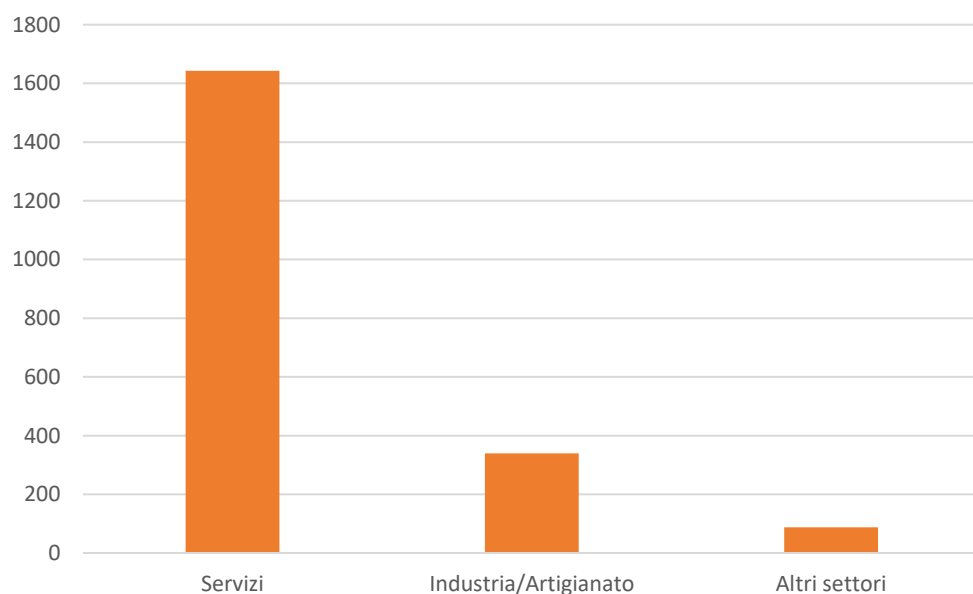


FIGURA 2-18 NUOVE STARTUP INNOVATIVE ISCRITTE PER SETTORE NEL 2020

Per quanto riguarda il grado di urbanizzazione, le nuove startup innovative del 2020 seguono gli andamenti degli anni precedenti riportati in [Figura 2-12](#), seppure subendo una diminuzione compresa tra il 22% e 27% in tutte le tre aree geografiche. Tuttavia, le proporzioni tra le varie zone non presentano variazioni significative, in quanto in quelle altamente popolate si registrano ancora il 67% e circa il 27% nelle aree a media urbanizzazione.

2.2 Brevetti

In Italia nel 2019 sono stati registrati mediamente 542,7 brevetti¹⁵ per ogni provincia. Le registrazioni di nuovi brevetti hanno avuto andamento eterogenei nel tempo nelle diverse aree geografiche. Al contrario di quanto si possa pensare, la nascita di startup innovative e la registrazione di nuove invenzioni sembrano correlate ma non in tutte le regioni (ad esempio la Campania terza per numero di startup innovative iscritte ma solo dodicesima per le registrazioni brevettuali). La [Figura 2-19](#) ha l'asse delle ordinate logaritmico; quindi, il divario con la Lombardia a può non sembrare significativo ma in realtà in questa regione sono stati registrati 106255 (il 42,6% del totale) mentre nel resto d'Italia 143206. La regione in cui si verifica una differenza maggiore rispetto alle iscrizioni di startup innovative è il Piemonte in cui sono stati registrati 48131 brevetti, mentre solo 538 startup sono nate nel periodo studiato.

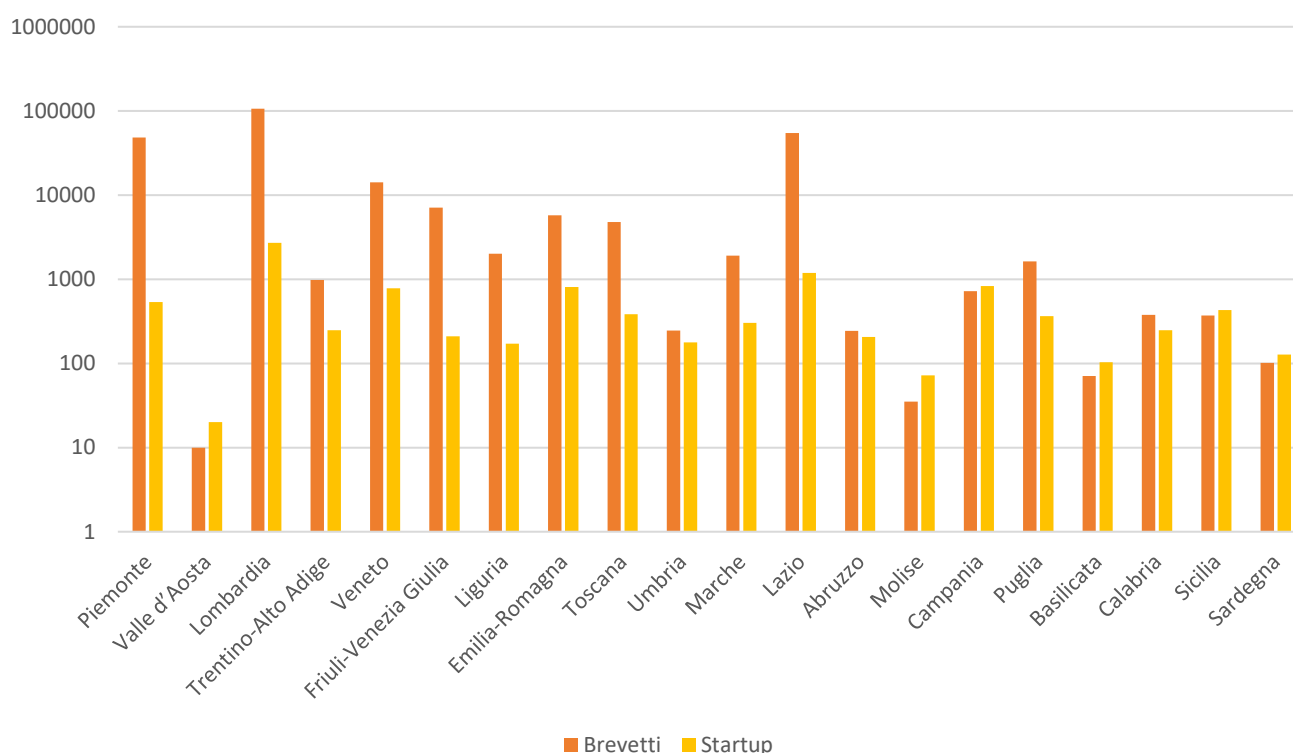


FIGURA 2-19 BREVETTI E STARTUP INNOVATIVE REGISTRATE DAL 2013 AL 2019

¹⁵ Il brevetto, in diritto, è un titolo giuridico in forza del quale al titolare viene conferito un diritto esclusivo di sfruttamento dell'invenzione, in un territorio e per un periodo ben determinato, e che consente di impedire ad altri di produrre, vendere o utilizzare l'invenzione senza autorizzazione.

Vista l'elevata quantità di registrazioni in Lombardia, si analizza l'andamento dei brevetti nell'arco temporale suddiviso tra le 12 province (Figura 2-20). Milano vanta il 94,3% dei brevetti lombardi nell'arco temporale 2013-2019. A seguire Brescia con 3235 brevetti e Monza e della Brianza con 977. Nelle restanti 9 province si sono verificate in totale 1579 registrazioni brevettuali (circa l'1,5%).

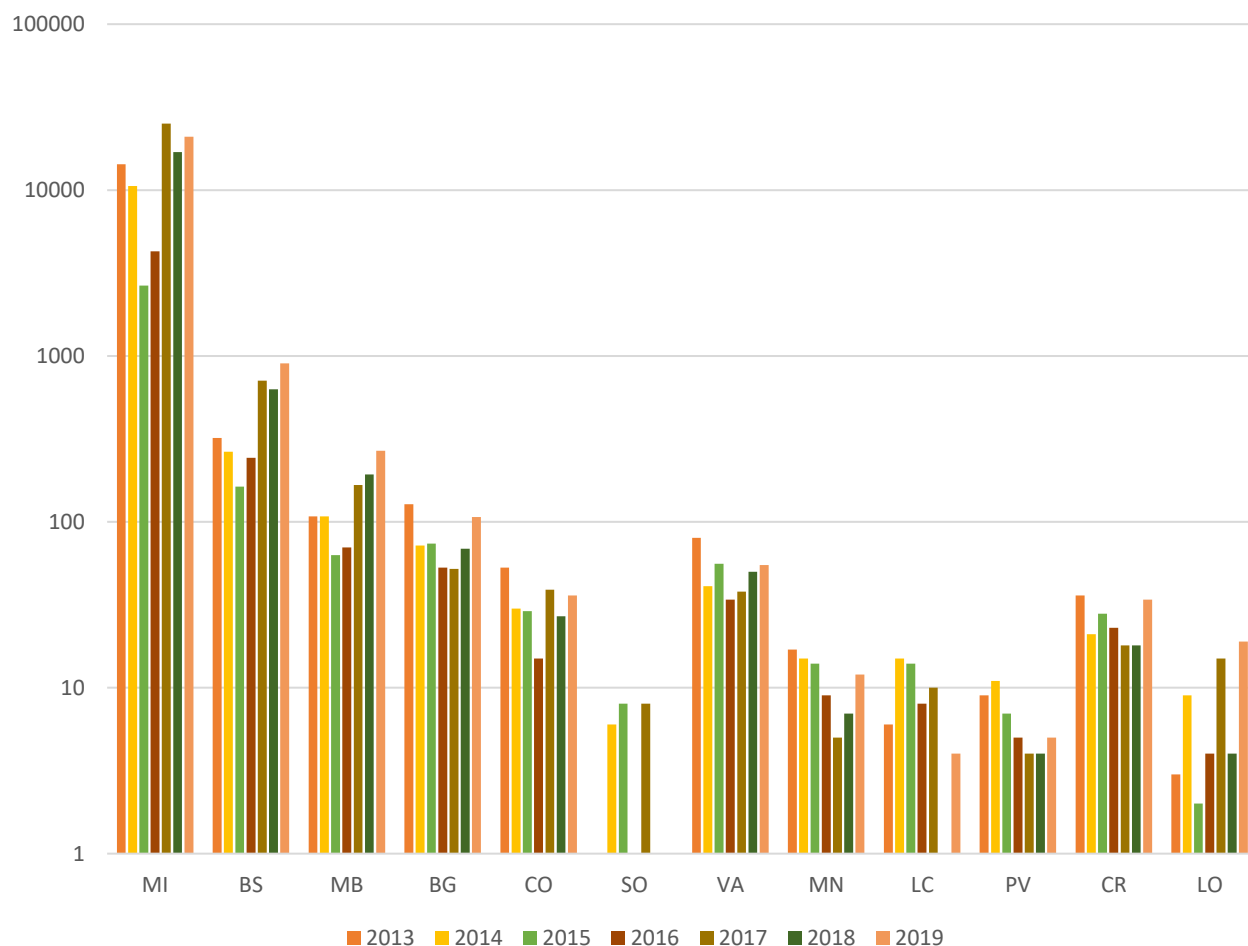


FIGURA 2-20 ANDAMENTO REGISTRAZIONI BREVETTUALI NELLE PROVINCE DELLA LOMBARDIA

Tornando a prendere in considerazione l'intera penisola, le registrazioni annue hanno andamenti molto simili nel nord e centro Italia (Figura 2-21). È interessante notare che anche nel sud la crescita marginale sia sempre concorde con quella delle altre due aree geografiche, ovvero ogni anno l'effetto di un aumento o diminuzioni delle registrazioni si riscontra con lo stesso segno in tutte le suddivisioni territoriali.

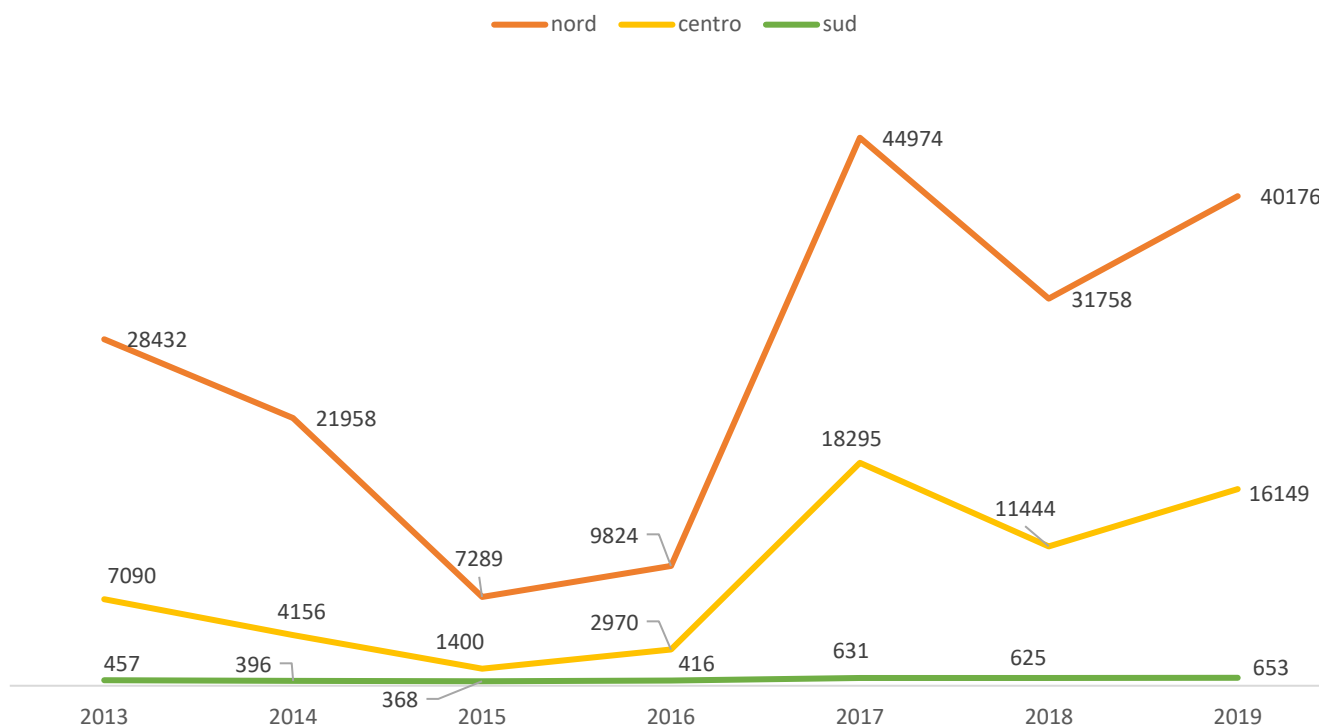


FIGURA 2-21 ANDAMENTO REGISTRAZIONI BREVETTUALI ANNUE

2.3 Diffusione UBB

La strategia empirica si avvale di un unico set di dati che raccoglie informazioni granulari sulla diffusione della banda larga ultraveloce (UBB) a livello comunale nel periodo 2013-2019. Queste informazioni contengono dati sulla diffusione della banda larga in Italia e specifiche descrizioni sull'infrastruttura di TIM, sia a livello locale che nazionale. Per avere un quadro completo, integriamo i dati di TIM sulla banda larga con informazioni aggiuntive sui comuni coperti dall'unico operatore alternativo di UBB, Open Fiber. Si analizza principalmente gli effetti dell'introduzione scaglionata dell'UBB a partire dal 2015, poiché la maggior parte della popolazione ha accesso a connessioni ADSL già all'inizio dell'arco temporale.

Il database contiene varie informazioni a livello comunale:

- Coverage ADSL 7mb: la copertura con ADSL ad almeno 7mb
- Coverage 20mb: la copertura con ADSL ad almeno 20 mb
- Coverage TIM UBB: copertura con banda larga ultraveloce da parte di TIM
- Sede OLT: osserviamo quali comuni hanno almeno un CO dotato di OLT.
- OLT distance: la distanza minima in chilometri tra ogni comune e la OLT più vicina (terminale di linea ottica).
- OPB distance: la distanza tra ogni comune e il suo nodo di backbone più vicino (nodi dell'infrastruttura Kaleidon)

- Coverage OF FTTH: copertura con banda larga ultraveloce da parte di Open Fiber
- Dummy UBB: copertura con banda larga ultraveloce da parte di almeno uno tra TIM e Open Fiber

2.3.1 Livello aggregato

Come panoramica iniziale, nella [Tabella 2-4](#) si riportano le statistiche riassuntive al termine del 2019.

TABELLA 2-4 VALORI STATISTICI DIFFUSIONE RETI

	Obs	Media	Std. Dev	Min	Max
<i>UBB</i>	7873	0.549	0.498	0.000	1.000
<i>TIM UBB</i>	7873	0.532	0.499	0.000	1.000
<i>OF UBB</i>	7873	0.036	0.173	0.000	1.000
<i>ADSL 7Mbit/s</i>	7873	0.963	0.177	0.000	1.000
<i>ADSL 20Mbit/s</i>	7873	0.841	0.341	0.000	1.000
<i>OLT distance</i>	7873	5.372	5.584	0.000	59.996
<i>OPB distance</i>	7873	43.615	25.344	0.000	297.726

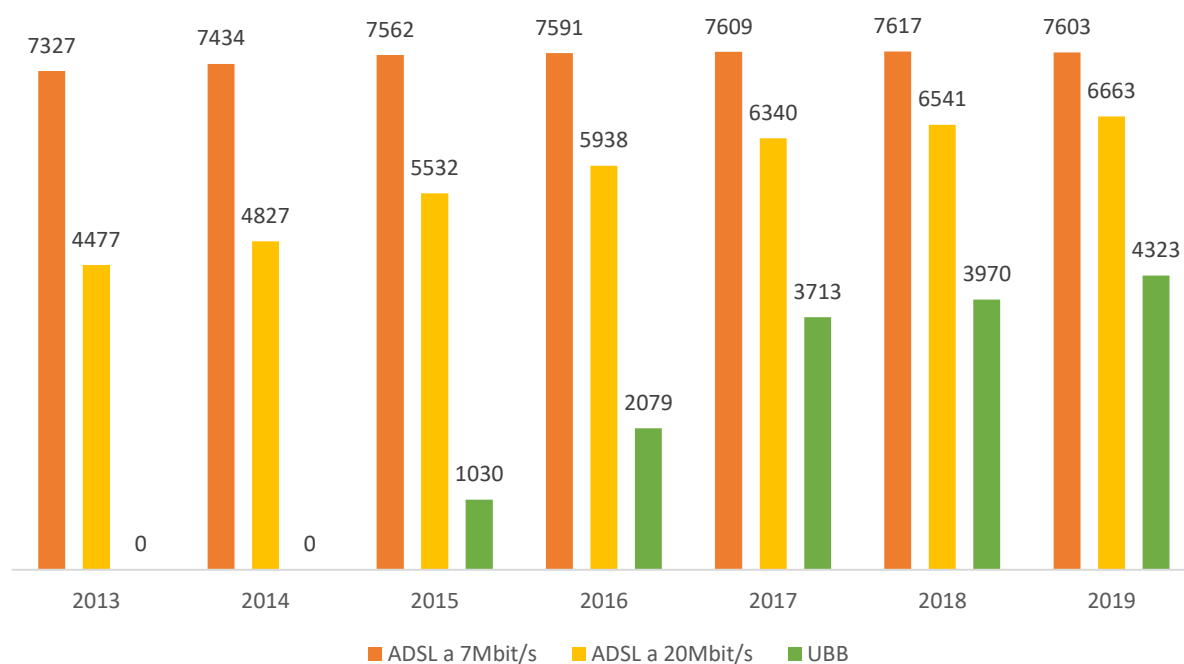
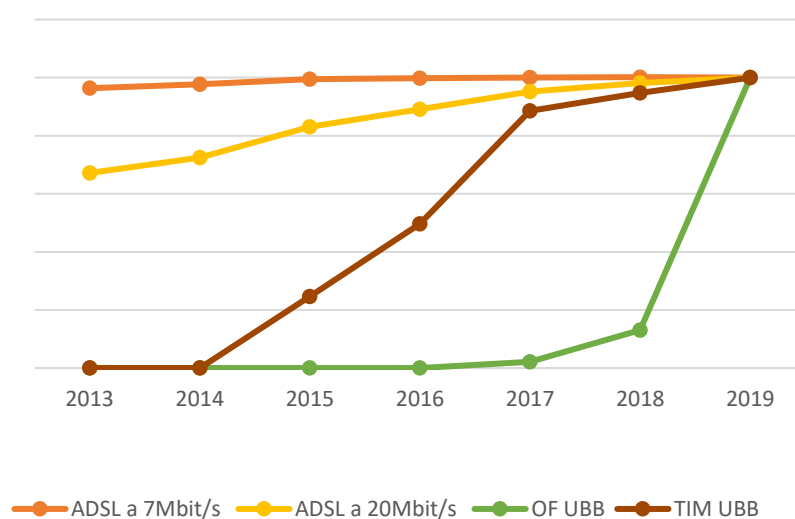


FIGURA 2-22 DIFFUSIONE DELLE TECNOLOGIE

Come si evince dalla [Figura 2-22](#), il 96,07% dei comuni italiani è coperto con ADSL a 7Mbit/s al termine del 2019, l'84,19% dispone anche di ADSL a 20Mbit/s e solo il 54,62% delle unità territoriali ha accesso alla banda larga ultraveloce. L'ADSL 7 Mbit/s ha avuto un andamento leggermente crescente (già nel 2013 copriva il 91,46% dei comuni) e il calo tra il 2018 e 2019 non è reale poiché legato alla diminuzione delle unità territoriali che sono passata da 7954 a 7914 in quell'anno. L'ADSL 20Mbit/s registra una crescita lineare aumentando di più di 28 punti percentuali nel periodo di tempo considerato. La fibra ottica ha fatto il suo ingresso sul territorio italiano nel 2015 coprendo 1030 comuni. L'evoluzione di questa tecnologia è stata rapida nei primi tre anni, raggiungendo il 46,52% nel 2017; mentre si è verificato un rallentamento della crescita nel 2018 e 2019, in cui è stata fornita solamente a 610 nuovi comuni.

FIGURA 2-23 GRADO DI COPERTURA

L'evoluzione delle tecnologie nel tempo (fino ai valori raggiunti al termine del 2019) non è lineare ma segue la ben nota curva a S (Geroski [2000]), la crescita tende a diminuire dopo aver raggiunto un sufficiente grado di maturità. La fase iniziale è caratterizzata da un inizio esitante (incubazione), seguito poi da un rapido incremento e diffusione delle prestazioni.



Adesso analizziamo l'andamento nelle tre aree geografiche ([Figura 2-24](#)).

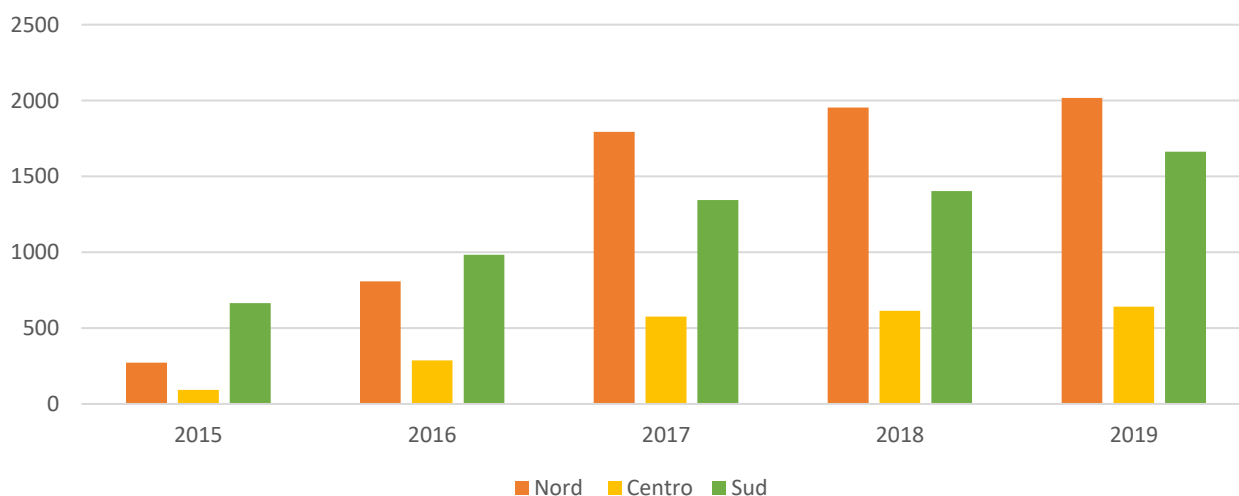


FIGURA 2-24 COPERTA UBB NORD-CENTRO-SUD

Nell'ultimo anno, il 46,04% del nord ha accesso alla banda larga ultraveloce, il 66,19% delle unità del centro e il 65,19% del sud. La copertura ancora distante dal 100% è legata al fatto che la fibra si sia diffusa sul territorio italiano dal 2015, molto più tardi rispetto all'ADSL, in uso dal 2000. Per i primi due anni, il sud era l'area geografica col maggior numero di comuni coperti. Il 2017 è stato l'anno in cui nord Italia ha beneficiato di un evidente incremento rispetto al resto della penisola: sono stati raggiunti da questa tecnologia complessivamente 985 nuovi comuni, mentre al centro-sud 649. Nonostante 2018 comuni settentrionali abbiano accesso a reti NGA, la copertura relativa è circa di 20 punti percentuali inferiore rispetto alle altre due aree geografiche.

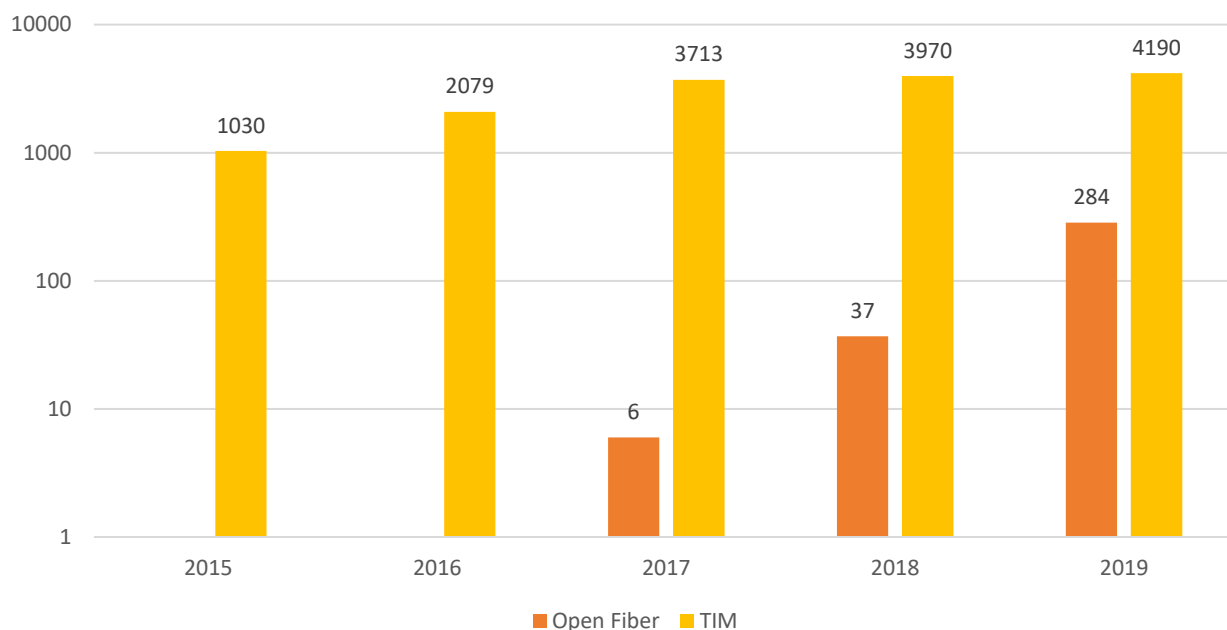


FIGURA 2-25 COPERTURA OF E TIM

Telecom Italia copre approssimativamente il 100% dei comuni con accesso alla tecnologia UBB fino al 2018, mentre nell'ultimo anno Open Fiber raggiungendo 284 unità territoriali è riuscita ad arrivare al 3,08% (Figura 2-25). OF si sta diffondendo principalmente nelle zone a fallimento di mercato, ha infatti installato principalmente le tecnologie FTTH in aree bianche (93 del nord, 33 del centro e 91 al sud).

Una domanda possibile è se la diffusione dell'UBB nelle aree bianche da parte di Open Fiber è correlata con la nascita di nuove startup innovative. Inizialmente le aree bianche totali risultano essere 18898. Open Fiber ha installato la tecnologia FTTH in 254 di queste. Al termine del 2019, 38 nuove startup iscritte al registro delle imprese nei comuni raggiunti, vale a dire 0,15 per ogni comune. Nelle restanti aree bianche (18644), sono nate 0,094 startup per ogni comune.

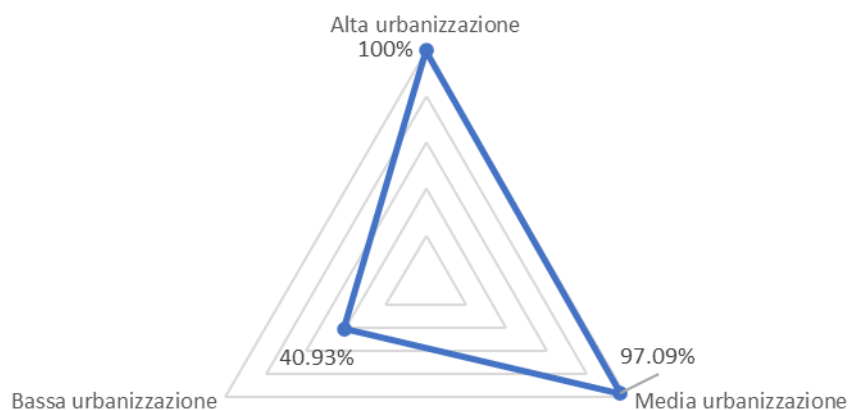


FIGURA 2-26 COPERTURA UBB SUDDIVISA PER GRADO DI URBANIZZAZIONE

Suddividendo poi i comuni per grado di urbanizzazione, TIM copre con banda larga ultraveloce 259 unità territoriali della prima categorizzazione, 1920 della seconda e 2001 della terza al termine del 2019 (Figura 2-25). È interessante notare che tutte le zone ad alta urbanizzazione sono raggiunte dall'incumbent delle telecomunicazioni. Open Fiber ha coperto 55 comuni della prima categoria (il 21,24% del totale) e poco più del 3,16% delle zone a medio-bassa GDU permettendo di raggiungere rispettivamente 1935 e 2119 comuni. Le aree a bassa urbanizzazione registrano il più basso tasso di copertura e 11884 dei comuni non ancora coperti sono aree bianche.

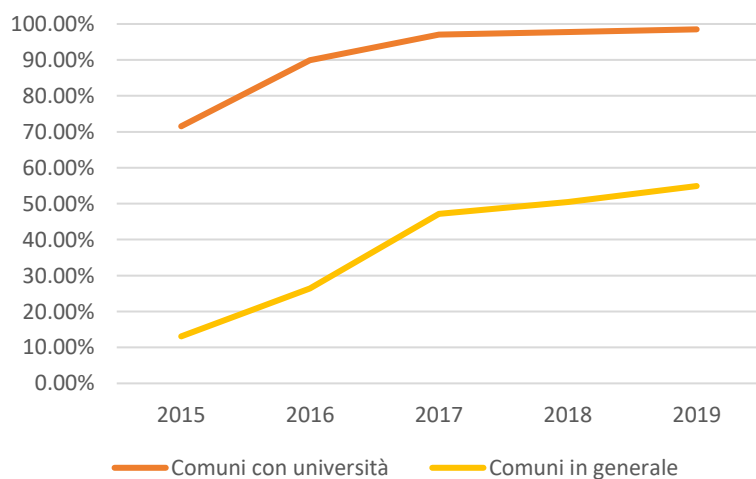


FIGURA 2-27 GRADO DI COPERTURA COMUNI CON UNIVERSITÀ

Per i comuni in cui è presente almeno un'università, si registra uno stadio di avanzamento della copertura quasi completo. Già nel 2015 c'è una differenza pari a 58,45 punti percentuali, questa aumenta (fino a un $\Delta=63,48\%$) nel 2016 in cui i comuni con università sfiorano il 90%.

Al termine del 2019 solo l'1,5% di questi comuni non è coperto con reti NGA.

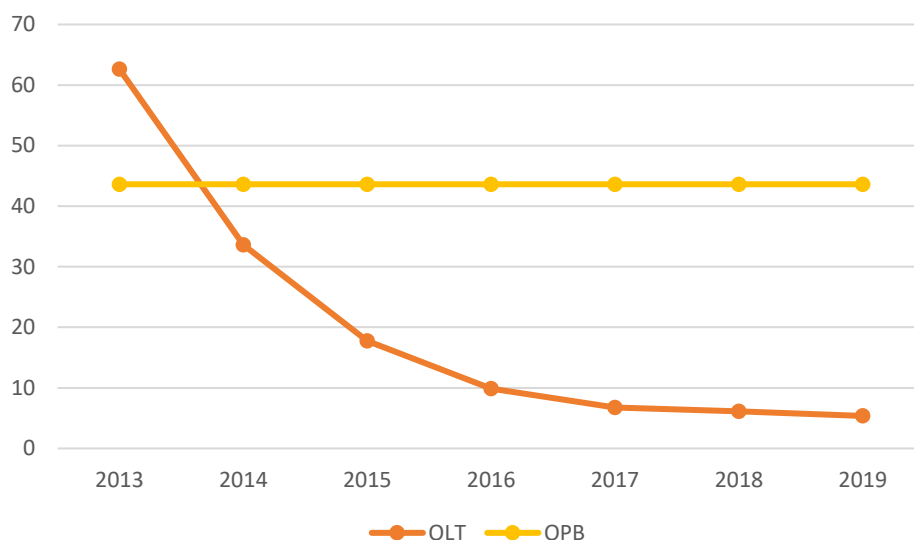


FIGURA 2-28 OLT E OPB DISTANCE

La [Figura 2-28](#) rappresenta l'andamento delle distanze medie (minime) dei comuni dal terminale di linea ottica e dal nodo di backbone.

La distanza OLT dal 2013 al 2016 è diminuita annualmente di un valore percentuale compreso tra 47 e 44 rispetto all'anno precedente. Nel 2017 si è ridotta ulteriormente del 31,7% mentre negli ultimi due anni il miglioramento è stato contenuto e la distanza media è passata da 6,75 km a 5,37 km. La distanza OPB è sempre rimasta costante e pari a circa 43 km. La [Figura 2-29](#) descrive la relazione tra la distanza OLT e il tasso di copertura. Solo le aree in cui la distanza minima è minore di 2 km si avvicinano al 100%. Meno del 5% dei comuni distanti dal terminale di linea ottica più vicino di un valore superiore a 8 km, ha accesso alle reti UBB.

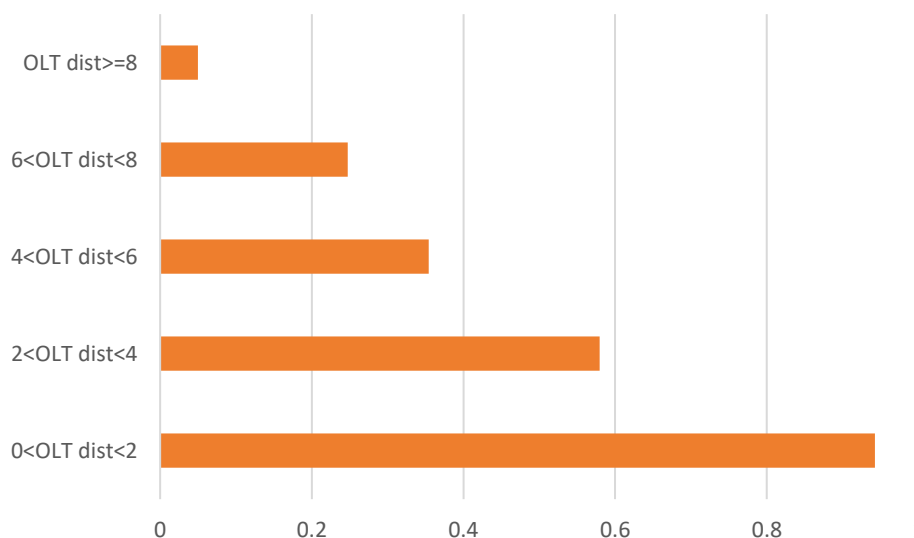


FIGURA 2-29 RELAZIONE TRA COPERTURA UBB E OLT DISTANCE

2.3.2 Livello regionale

Adesso spostiamo l'attenzione sull'infrastruttura a livello regionale, mettendo in evidenza quanti comuni hanno almeno un CO dotato di OLT, qual è la distanza dal terminale più vicino e infine dal nodo dell'infrastruttura Kaleidon (OPB).

TABELLA 2-5 DISTANZE DA OLT E OPB MEDIE SUDDIVISE PER REGIONE

	Comuni con sede OLT		OLT distance	OPB distance
PIEMONTE	123	10.41%	7.73	42.43
VALLE D'AOSTA	3	4.05%	9.97	77.24
LOMBARDIA	392	26.03%	3.86	26.50
TRENTINO-ALTO ADIGE	29	10.28%	8.62	27.50
VENETO	197	34.99%	3.19	35.55
FRIULI-VENEZIA GIULIA	47	21.86%	5.93	30.76
LIGURIA	47	20.09%	6.35	41.32
EMILIA-ROMAGNA	144	43.90%	5.10	31.98
TOSCANA	143	52.38%	3.23	44.42
UMBRIA	24	26.09%	7.26	41.37
MARCHE	55	24.23%	5.47	46.09
LAZIO	118	31.22%	4.54	58.69
ABRUZZO	53	17.38%	7.96	48.55
MOLISE	9	6.62%	11.75	78.26
CAMPANIA	175	31.82%	3.25	44.39
PUGLIA	148	57.59%	2.58	69.59
BASILICATA	43	32.82%	5.72	93.18
CALABRIA	105	25.99%	4.10	53.68
SICILIA	172	43.99%	3.58	56.82
SARDEGNA	73	19.36%	8.33	54.84

Le distanze sono valori medi, calcolati dividendo la somma totale delle distanze in km per ogni comune con il totale dei comuni nella regione di appartenenza (Tabella 2-5). I comuni lombardi sono mediamente i più vicini ai nodi OPB (26,50km). Trentino, Veneto, Friuli-Venezia Giulia e Emilia-Romagna hanno distanze medie inferiori a 36km. Tra i 41 e 69 km medi troviamo la maggior parte delle regioni, mentre le peggior posizionate sono la Valle d'Aosta, Molise, Puglia e Basilicata, i cui comuni distano rispettivamente 77,24 km, 78,26 km, 69,59 km e 93,18 km dal nodo di backbone più vicino. La distanza OLT è influenzata dal numero di comuni aventi almeno una sede e dalla conformazione geografica della

regione. Le regioni aventi almeno il 40% di comuni provvisti di OLT, sono caratterizzate da valori medi della distanza compresi tra 2,58 e 5,1 km. Nonostante la proporzione inferiori di sedi OLT, Lombardia, Veneto; Campania e Sicilia usufruiscono di distanze medie molto basse, minori di 4 km. I comuni con distanze medie maggiori di 7 km sono: Valle d'Aosta, Trentino-Alto Adige, Piemonte, Umbria, Abruzzo, Molise, e Sardegna. Le 7 regioni peggiori hanno il 12,83% dei comuni con CO dotato di OL e distanze medie dei pari 8,23 km, sinonimo di una rete meno performante.

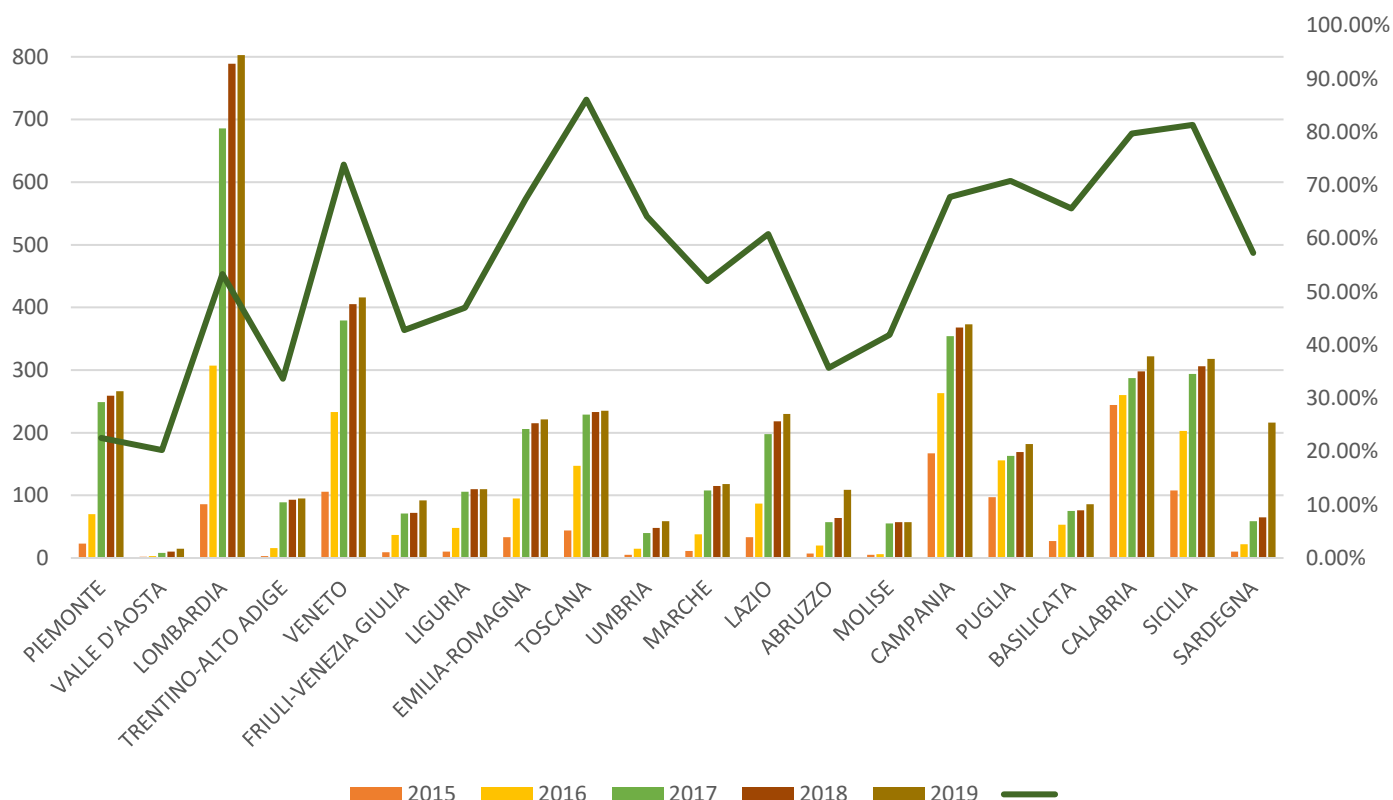


FIGURA 2-30 ANDAMENTO COMUNI COPERTI CON UBB E PERCENTUALE FINALE

La Toscana con l'86,08% e la Sicilia col 81,33% sono le regioni con la percentuale maggiore di comuni raggiunti dalla banda larga ultraveloce (Figura 2-30). In valori assoluti, la Lombardia presenta invece il più alto numero di comuni coperti (803) e un grado di copertura pari al 53,32%. A seguire il Veneto con 416 comuni (73,89%) e la Campania 373 (67,82%). I fanalini di coda sono Valle d'Aosta e Piemonte, in cui meno del 23% dei comuni hanno accesso alle reti NGA. Anche Trentino-Alto Adige, Liguria, Abruzzo e Molise non raggiungono il 50% dei comuni. L'anno in cui complessivamente è stato raggiunto il maggior numero di comuni è stato il 2017 in cui si è aggiunto 1634 unità. Anche i due anni precedenti hanno superato quota 1000 comuni, mentre nel 2018 e 2019 si è verificato un rallentamento e solamente 257 e 353 nuovi comuni hanno avuto accesso alla rete UBB.

Nella Tabella 2-7 seguente si riportano i sette maggior incrementi annuali suddivisi per regione. Mediamente ogni anno, le reti UBB sono diffuse in 43,23 nuove unità territoriali di ogni regione. In Lombardia si registra il valore più alto pari a 160,6, circa il doppio

rispetto tutte le altre regioni (Tabella 2-6). A seguire si registrano Veneto, Campania, Calabria, Sicilia i cui incrementi medi annui sono superiori a 63,5 comuni.

TABELLA 2-6 COMUNI RAGGIUNTI ANNUALMENTE CON RETI UBB

			<i>PIEMONTE</i>	53.2
			<i>VALLE D'AOSTA</i>	3
			<i>LOMBARDIA</i>	160.6
			<i>TRENTINO-ALTO ADIGE</i>	19
			<i>VENETO</i>	83.2
			<i>FRIULI-VENEZIA GIULIA</i>	18.4
			<i>LIGURIA</i>	22
			<i>EMILIA-ROMAGNA</i>	44.2
			<i>TOSCANA</i>	47
			<i>UMBRIA</i>	11.8
			<i>MARCHE</i>	23.6
			<i>LAZIO</i>	46
			<i>ABRUZZO</i>	21.8
			<i>MOLISE</i>	11.4
			<i>CAMPANIA</i>	74.6
			<i>PUGLIA</i>	36.4
			<i>BASILICATA</i>	17.2
			<i>CALABRIA</i>	64.4
			<i>SICILIA</i>	63.6
			<i>SARDEGNA</i>	43.2

<i>Lombardia 2016</i>	+379	TABELLA 2-7 MAGGIORI INCREMENTI ISCRIZIONI ANNUALI SUDDIVISI PER REGIONE
<i>Calabria 2015</i>	+244	
<i>Lombardia 2015</i>	+221	
<i>Piemonte 2016</i>	+179	
<i>Campania 2015</i>	+167	
<i>Sardegna 2019</i>	+151	
<i>Trentino-Alto Adige 2016</i>	+146	

Infine, spostiamo l'attenzione sulle province che hanno registrato almeno 1000 brevetti nel 2019: Torino, Milano, Bologna, Vicenza, Udine e Roma (Figura 2-31). Al termine dello studio, il 53,13 % di questi comuni è coperto da UBB. La provincia di Milano vanta 97 comuni coperti su 107, pari al 90,65%. Bologna e Vicenza hanno rispettivamente l'82,05% e il 65,71% dei comuni con accesso a tale rete, Roma poco meno del 64%. A Udine complessivamente 40 unità territoriali su 121 sono state raggiunte. Percentuale simile per Torino, in cui, nonostante sia la prima per numerosità assoluta (109 comuni), ancora 194 devono essere coperti da reti NGA.

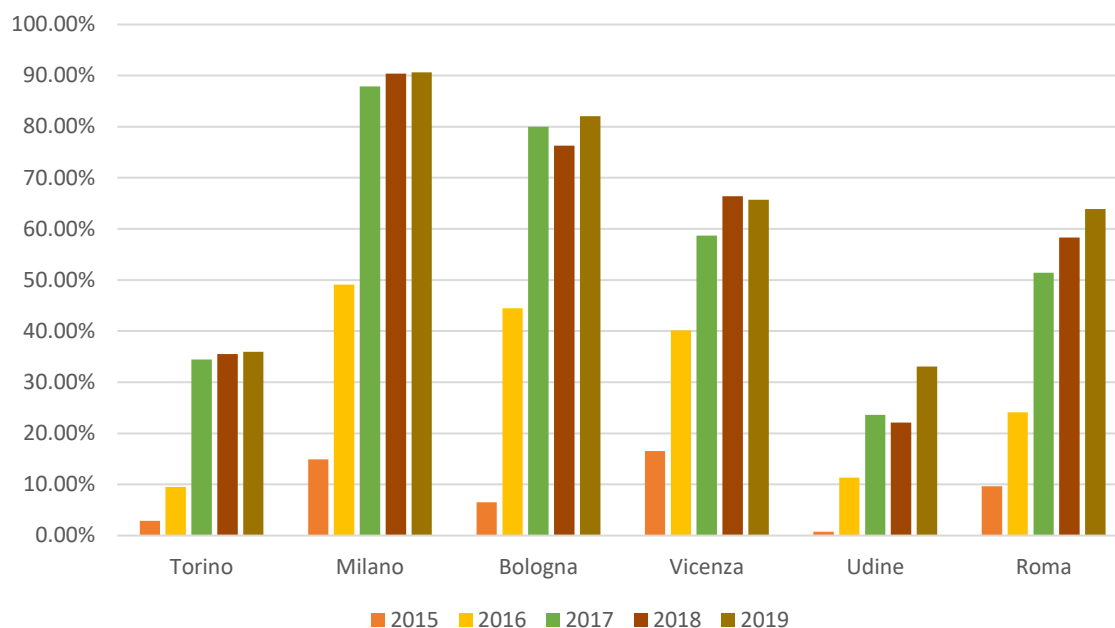


FIGURA 2-31 COPERTURA UBB NELLE PROVINCE CON MAGGIOR NUMERO DI BREVETTI REGISTRATI NEL 2019

Se si considerassero solo i comuni in cui è stata costituita almeno una startup dal 2013 al 2019, si restringerebbe l'analisi descrittiva della diffusione della banda larga ultraveloce a 1660 enti locali (828 nel Nord, 254 centro e 578 sud Italia).

TIM copre il 99,5% delle unità territoriali in cui sono nate startup innovative. Ciò avviene perché OF ha sviluppato principalmente la sua rete nelle zone a fallimento di mercato che tendenzialmente sono le stesse in cui i processi innovativi e di sviluppo dell'economia sono più carenti e per questo con meno probabilità si sarà iscritta almeno una startup nell'arco temporale studiato. OF ha installato tecnologie FTTH rispettivamente in 19 di questi comuni del nord, 4 del centro, 5 del sud e tra queste in 15 aree bianche.

Capitolo 3 - Modello empirico

Nel terzo e ultimo capitolo è analizzato l'effetto causale¹⁶ di una variazione nel numero di comuni coperti con banda ultraveloce sulla nascita di nuove startup innovative, sfruttando due modelli econometrici diversi: Ordinary Least Squares e Instrumental Variables.

Il software statistico con cui si è eseguito le regressioni presenti nell'elaborato di tesi è *Stata 12*.

3.1 Ordinary Least Squares

3.1.1 Modello

Il modello su cui si basa l'analisi empirica è stato costruito attraverso un processo incrementale. Inizialmente si è identificata un'equazione lineare a un'unica variabile dipendente, per poi arrivare a una maggior caratterizzazione, aggiungendo controlli, effetti fissi temporali e inerenti al comune o alla regione di appartenenza e infine trend temporali condizionati alle variabili di riferimento.

La variabile dipendente studiata è sempre il numero di iscrizioni annuali al Registro delle Imprese di startup innovative ($Y_{m,t}$), riferito a un determinato comune e anno.

La variabile principale d'interesse è l'accesso alla connessione ultraveloce ($UBB_{m,t}$) del comune m nell'anno t . È una dummy e assume valore 1 nel caso in cui il comune abbia accesso alle reti NGA nell'anno preso in esame.

Attraverso la regressione lineare, è possibile stimare l'effetto atteso sulla variabile dipendente di una variazione unitaria della copertura, studiando la pendenza della retta di regressione.

In prima battuta, è pertanto utile fare riferimento alla seguente equazione¹⁷.

$$Y_{m,t} = \beta_0 + \beta_1 UBB_{m,t} + u_{m,t}$$

Per stimare i parametri β_0 e β_1 , si utilizza il metodo dei minimi quadrati ordinari ("OLS").

¹⁶ Un effetto causale è definito come un effetto misurato in un esperimento controllato casualizzato ideale. Si immagini un esperimento controllato casualizzato ideale per misurare l'effetto:

- In tale esperimento la banda larga ultraveloce sarebbe assegnata casualmente ai comuni, che avrebbero caratteristiche diverse (per quanto riguarda GDU, ADSL, università...)
- Poiché la copertura UBB è assegnata casualmente, tutte le caratteristiche (e quindi gli errori) sarebbero distribuiti in modo indipendente dalla variabile indipendente.

¹⁷ Dove le variabili identificano:

- $Y_{m,t}$ il numero di startup innovative iscritte nel comune m nell'anno t
- $UBB_{m,t}$ l'accesso alla connessione ultraveloce nel comune m nell'anno t
- β_1 il coefficiente angolare
- β_0 l'intercetta
- $u_{m,t}$ il residuo di regressione, costituito da fattori omessi

$$\min_{b_0, b_1} \sum_{i=1}^n [Y_{m,t} - (b_0 + b_1 UBB_{m,t})]^2$$

Lo stimatore OLS minimizza la somma dei quadrati delle differenze tra i valori reali (osservati) di $Y_{m,t}$ e i valori predetti in base alla retta di regressione stimata, ovvero $b_0 + b_1 UBB_{m,t}$.

Gli stimatori¹⁸ OLS della pendenza e dell'intercetta possono essere calcolati nel modo seguente¹⁹:

$$\widehat{\beta}_1 = \frac{\sum_{m=1}^n (X_m - \bar{X})(Y_m - \bar{Y})}{\sum_{m=1}^n (X_m - \bar{X})^2} = \frac{s_{xy}}{s_x^2}$$

$$\widehat{\beta}_0 = \bar{Y} - \widehat{\beta}_1 \bar{X}$$

Infine, si ricavano valori predetti $\widehat{Y}_i$ e i residui di regressioni $\widehat{u}_i$:

$$\widehat{Y}_i = \widehat{\beta}_0 + \widehat{\beta}_1 X_i, i = 1, \dots, n$$

$$\widehat{u}_i = Y_i - \widehat{Y}_i, i = 1, \dots, n$$

Gli errori $\widehat{u}_i$ si differenziano in due tipologie. Se la $var(u|X_i = x)$ è costante, ossia se la varianza della distribuzione di u condizionata a X non dipende da X , allora l'errore è detto omoschedastico. In caso contrario, u è eteroschedastico.

È necessario prestare particolare attenzione poiché esistono due formule per calcolare gli errori standard di $\widehat{\beta}_1$ che portano però a risultati diversi: in quanto una di queste è applicabile solo in caso di omoschedasticità pura, mentre l'altra è definita "robusta all'eteroschedasticità".

Nel caso in cui gli errori siano eteroschedastici e si utilizzasse la formula dell'omoschedasticità pura, gli errori standard saranno sbagliati: lo stimatore della varianza di $\widehat{\beta}_1$ incoerente e l'inferenza errata. Inoltre, gli SE saranno sottostimati e pertanto risulterà elevato il rischio di rifiutare erroneamente l'ipotesi nulla. Per maggior accuratezza, nelle regressioni successive, si utilizzeranno sempre errori standard robusti all'eteroschedasticità.

Per approcciare l'analisi empirica, si effettua adesso una regressione lineare di base (Tabella 3-1), ricavando approssimativamente la pendenza della retta di regressione: si riscontra che a un incremento unitario dell'accesso alla connessione ultraveloce, corrisponda un aumento del numero di startup pari a 0,588. Chiaramente tale stima non è precisa ma è comunque possibile riferirsi a tale valore come punto di partenza prima di proseguire con lo sviluppo del modello.

¹⁸ L'intercetta stimata $\widehat{\beta}_0$, la pendenza stimata $\widehat{\beta}_1$ e il residuo $\widehat{u}_i$ sono calcolati per un campione di n osservazione di X_i e $\widehat{Y}_i$ con $i = 1, \dots, n$. Queste sono stime dell'intercetta β_0 , della pendenza β_1 e dell'errore u_i nella popolazione.

¹⁹ Nelle formule sotto riportate la variabile indipendente UBB è indicata con X e nel pedice i si identifica sia il comune che l'anno

TABELLA 3-1 APPROCCIO ALL'ANALISI EMPIRICA

	startup
UBB	0.588*** (0.064)
_cons	0.0180*** (0.002)
<i>N</i>	55111
<i>R</i> ²	0.004

Standard errors in parentheses

Source: auto.dta

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

L'errore, sempre presente all'interno delle regressioni, si verifica a causa di fattori, o variabili, che influenzano $Y_{m,t}$ ma non sono inclusi nella funzione. Talvolta l'omissione di queste variabili può portare a una distorsione dello stimatore OLS.

Ciò accade quando si verificano due condizioni per quanto concerne la variabile omessa Z : essere una determinante di $Y_{m,t}$ (cioè Z è parte dell'errore) e correlata col regressore UBB (cioè $\text{corr}(Z, UBB) \neq 0$).

In queste circostanze, potrebbe essere interessante cercare di comprendere la direzione della distorsione da variabile omessa, ovvero tentare di prevedere la correlazione tra l'incremento di startup innovative e l'errore (indicato ρ_{UBBu} nell'equazione seguente).

$$\hat{\beta}_1 = \beta_1 + \left(\frac{\sigma_u}{\sigma_{UBB}} \right) \rho_{UBBu}$$

A prima vista, si potrebbe semplicemente immaginare che gli shock non osservabili della domanda, che influenzano l'introduzione della fibra, abbiano un impatto sulla nascita delle startup innovative nella stessa direzione. Per esempio, il verificarsi di un evento esogeno che aumenti il potere d'acquisto dei consumatori, implicherebbe presumibilmente un effetto positivo sull'attrattività di quell'area dal punto di vista degli operatori di telecomunicazioni. Pertanto, essi saranno così incentivati a investire in connessioni ultraveloci. Meno evidenti le conseguenze derivanti dagli shock dal lato dell'offerta, i quali potrebbero implicare una relazione opposta tra la nascita di startup e la diffusione della banda larga ultraveloce. In mercati maggiormente competitivi le aziende avranno incentivi ad adottare ICT avanzate, che a loro volta aumenterebbero la domanda di connessioni ultraveloci; allo stesso tempo però diminuirà l'entrata di nuove imprese, poiché i margini si ridurranno e si creeranno ulteriori e solide barriere d'ingresso.

Un altro esempio plausibile potrebbe essere la presenza di almeno un polo universitario all'interno di un comune. Tale variabile soddisfa le due condizioni sopra citate, in quanto:

- Influenza verosimilmente il carattere tecnico e innovativo di una certa area, e di conseguenza il numero di startup nascenti. Z è una determinante di Y .
- I comuni meno popolati (caratterizzati probabilmente da "scarsa probabilità di avere un'università") tendono ad avere una domanda di banda inferiore e di conseguenza, gli

operatori TLC avranno incentivi ridotti a raggiungerli con le reti NGA: Z è correlata con UBB , ovvero $\rho_{Xu} \neq 0$ e lo stimatore OLS $\hat{\beta}_1$ risulterà distorto e inconsistente.

In questo caso, se non si considerasse l'effetto della variabile di controllo, si sovrastimerebbe l'effetto della banda larga ultraveloce.

Per superare la distorsione, si sfrutta quindi la regressione sotto riportata, in cui la variabile omessa non è più omessa.

$$Y_{m,t} = \beta_0 + \beta_1 UBB_{m,t} + \beta_2 Z_{m,t} + u_{m,t}$$

Z identifica la variabile di controllo, β_2 è l'effetto su $Y_{m,t}$ di una variazione in Z mantenendo $UBB_{m,t}$ costante; mentre β_1 misura l'effetto sulla variabile dipendente di una variazione in $UBB_{m,t}$ mantenendo costante Z . Come in precedenza, è possibile sfruttare lo stimatore OLS e incrementando il numero di regressori (fino a un valore pari a k), il problema di minimizzazione evolverà nell'equazione seguente.

$$\min_{b_0, \dots, b_k} \sum_{i=1}^n [Y_{m,t} - (b_0 + b_1 UBB_{m,t} + b_2 Z_{m,t} + b_3 X_{3m,t} \dots + b_k X_{km,t})]^2$$

Lo stimatore OLS è considerato efficiente soltanto se quattro particolari assunzioni risultano rispettate:

$$Y_{m,t} = \beta_0 + \beta_1 UBB_{m,t} + \beta_2 Z_{m,t} + \beta_3 X_{3m,t} \dots + \beta_k X_{km,t} + u_{m,t}$$

1. La distribuzione di u condizionata alle X ha media nulla, cioè $E(u_i | UBB_{m,t} = x_1, \dots, X_{km,t} = x_k) = 0$.
2. $(UBB_{m,t}, \dots, X_{km,t}, Y_{m,t})$ sono *i.i.d.*, ovvero prelevate dalla stessa popolazione (identicamente distribuite) e scelte a caso (perciò i valori di (X, Y) per unità diverse sono indipendentemente distribuiti).
3. Gli outlier sono improbabili, ovvero $UBB_{m,t}, Z_{m,t}, \dots, X_{km,t}, e Y_{m,t}$ hanno momenti quarti finiti: $E(UBB_{m,t}^4) < \infty, \dots, E(X_{km,t}^4) < \infty, E(Y_{m,t}^4) < \infty$.
4. Non vi è collinearità perfetta, ovvero quando uno dei regressori è funzione lineare esatta degli altri. Anche in caso di collinearità imperfetta, si generano grandi errori standard per uno o più coefficienti OLS. Se lo esigesse il modello o si dovesse includere queste variabili per evitare distorsione da variabile omessa, è necessario accettare che i loro coefficienti siano stimati con alta varianza campionaria.

La regressione multipla ha pertanto il potenziale di controllare l'effetto sulla variabile dipendente generato da diverse variabili. Tuttavia, ciò potrebbe compromettere la prima assunzione dei minimi quadrati oppure in caso di indisponibilità di dati, non riuscire comunque nell'intento di colmare la distorsione da variabile omessa. In quest'ultima circostanza si può trovare una soluzione sfruttando i dati panel, poiché consentono di controllare l'effetto di alcune variabili omesse senza la necessità di osservarle direttamente.

Il dataset su cui si basa questo lavoro di tesi è dunque un panel bilanciato. Non sono infatti presenti unità con osservazioni mancanti, ovvero le variabili sono osservate per tutte le unità (pari a M comuni) e in tutti gli anni dell'arco temporale (T , dal 2013 al 2019).

Adesso, è necessario definire gli effetti fissi a livello comunale e successivamente gli effetti temporali. Per quanto riguarda i primi, si può controllare per fattori (Z_m) che variano tra i comuni, ma restano costanti nel tempo, causando la distorsione da variabili omesse nel caso

in cui fossero omessi o inosservati e non misurati. L'idea chiave intorno cui ruota la logica è la seguente: se una variabile omessa non variasse nel tempo, allora qualsiasi variazione in Y nel tempo non potrebbe essere causata dalla variabile omessa.

$$Y_{m,t} = \beta_0 + \beta_1 UBB_{m,t} + \beta_2 Z_m + u_{m,t}$$

A titolo di esempio, si effettua una regressione per il comune di Torino ($m = Torino$):

$$\begin{aligned} Y_{Torino,t} &= \beta_0 + \beta_1 UBB_{Torino,t} + \beta_2 Z_{Torino} + u_{Torino,t} \\ &= (\beta_0 + \beta_2 Z_{Torino}) + \beta_1 UBB_{Torino,t} + u_{Torino,t} \\ &= \alpha_{Torino} + \beta_1 UBB_{Torino,t} + u_{Torino,t} \end{aligned}$$

$\alpha_{Torino} = \beta_0 + \beta_2 Z_{Torino}$ è l'intercetta per Torino e costante nel tempo, mentre β_1 è la pendenza. L'intercetta è specifica, mentre la pendenza è la stessa in tutti i comuni: gli effetti fissi a livello comunale sono dunque rappresentabili graficamente con rette parallele.

D'altro canto, ci potrebbero essere anche variabili omesse variabili nel tempo e costanti tra i comuni. Il modello di regressione può essere ricomposto con una nuova intercetta che varia da un anno al successivo in modo da includere gli effetti temporali.

$$Y_{i,t} = \beta_0 + \beta_1 UBB_{i,t} + \beta_3 S_t + u_{i,t}$$

Se si considerasse ad esempio l'anno 2013:

$$\begin{aligned} Y_{m,2013} &= \beta_0 + \beta_1 UBB_{m,2013} + \beta_3 S_{2013} + u_{m,2013} \\ &= \beta_0 + \beta_3 S_{2013} + \beta_1 UBB_{m,2013} + u_{m,2013} \\ &= \lambda_{2013} + \beta_1 UBB_{m,2013} + u_{m,2013} \end{aligned}$$

Dove λ_{2013} rappresenta l'intercetta per l'anno 2013, mentre β_1 la pendenza.

In conclusione, si mette insieme quanto appena detto e si esegue la stima considerando sia gli effetti fissi tra i comuni che gli effetti temporali.

$$Y_{m,t} = \beta_1 UBB_{m,t} + \alpha_m + \lambda_t + u_{m,t}$$

Sotto le assunzioni dei minimi quadrati nella versione per dati panel, lo stimatore OLS con effetti fissi di β_1 ha distribuzione normale.

Le prime due ipotesi per dati panel presentano particolari estensioni rispetto al caso precedente, mentre la terza e la quarta sono identiche.

La prima assunzione è $E(u_{m,t} | UBB_m, \dots, UBB_{mT}, \alpha_m) = 0$.

Essa impone che $u_{m,t}$ abbia media zero, dato l'effetto fisso e l'intera storia della variabile dipendente per l'unità corrispondente. Ciò significa che non vi siano effetti passati omessi.

La seconda assunzione è che $(UBB_{m1}, \dots, UBB_{mT}; u_{m1}, \dots, u_{mT})$ siano identicamente indipendentemente distribuite dalla distribuzione congiunta. Non richiede che le osservazioni siano *i.i.d.* nel tempo per la stessa unità poiché sarebbe irrealistico. Il fatto che in un comune si sia registrato un elevato aumento di startup innovative quest'anno è un buon predittore di un'elevata crescita anche nell'anno seguente. Similmente, il termine

d'errore per un'unità in un anno è plausibilmente correlato con il suo valore l'anno dopo, ovvero la $corr(u_{m,t}, u_{m,t+1})$ è plausibilmente diversa da zero. Quest'ultima è definita autocorrelazione. Ciò si verifica quando una variabile Z , osservata in diverse date t , è correlata con sé stessa. Ovvero se $corr(Z_t, Z_{t+j}) \neq 0$ per date $j \neq 0$.

La $cov(Z_t, Z_{t+j})$ è detta j -esima autocovarianza di Z_t . Se le unità sono ottenute per campionamento casuale semplice, allora $u_{i1}, \dots, u_{iT}$ è indipendente da $u_{j1}, \dots, u_{jT}$ per unità diverse con $i \neq j$. Però, se i fattori omessi che comprendono u_{mT} sono autocorrelati, allora u risulta essere serialmente correlato. Di conseguenza, i consueti errori standard OLS (sia di omoschedasticità pura sia robusti all'eteroschedasticità) saranno in generale sbagliati perché si basano sull'assunzione che l'errore non sia serialmente correlato.

Gli errori standard OLS rischiano pertanto di sottostimare l'incertezza del campionamento reale poiché in caso di autocorrelazione, non si hanno molte informazioni riguardanti l'effetto casuale, così come si avrebbero se $u_{m,t}$ fosse incorrelata. Il problema può essere risolto usando errori standard "clustered", i quali permettono di stimare la varianza di β_1 quando le variabili sono *i.i.d.* tra le unità ma sono potenzialmente autocorrelate in un'unità.

L'equazione, per la media stimata con dati panel, è la seguente:

$$Y_{mt} = \mu + u_{mt}$$

Lo stimatore della media μ è $\bar{Y} = \frac{1}{nT} \sum_{m=1}^n \sum_{t=1}^T Y_{mt}$.

È utile scrivere $\bar{Y}$ come media tra le unità del valore medio per ciascuna unità.

$$\bar{Y} = \frac{1}{nT} \sum_{m=1}^n \sum_{t=1}^T Y_{mt} = \frac{1}{n} \sum_{m=1}^n \left(\frac{1}{T} \sum_{t=1}^T Y_{mt} \right) = \frac{1}{n} \sum_{m=1}^n (\bar{Y}_m)$$

dove $\bar{Y}_m = \frac{1}{T} \sum_{t=1}^T (Y_{mt})$ è la media campionaria per l'unità i .

Poiché le osservazioni sono *i.i.d.* tra le entità, $(\bar{Y}_m, \dots, \bar{Y}_n)$ sono *i.i.d.* Quindi, se n è grande, vale la seguente approssimazione.

$$\bar{Y} = \frac{1}{n} \sum_{m=1}^n \bar{Y}_m \xrightarrow{d} N(0, \frac{\sigma_{\bar{Y}_m}^2}{n}), \text{ dove } \sigma_{\bar{Y}_m}^2 = var(\bar{Y}_m)$$

L'errore standard di $\bar{Y}$ è la radice quadrata di uno stimatore di $\sigma_{\bar{Y}_m}^2/n$.

Ciò fornisce la formula dell'errore standard per dati raggruppati per $\bar{Y}$ calcolata usando dati panel:

$$\bar{Y} = \sqrt{\frac{s_{\bar{Y}_m}^2}{n}}, \text{ dove } s_{\bar{Y}_m}^2 = \frac{1}{n-1} \sum_{m=1}^n (\bar{Y}_m - \bar{Y})^2$$

Rispetto all'errore standard della media campionaria, in questo caso i dati sono le medie di unità *i.i.d.* $(\bar{Y}_m, \dots, \bar{Y}_n)$ anziché una singola osservazione *i.i.d.* per ciascuna unità.

La caratteristica principale che ne giustifica l'importanza è che nella derivazione dell'errore standard per dati raggruppati non si è mai assunto che le osservazioni siano *i.i.d.* in un'unità, acconsentendone implicitamente alla correlazione seriale.

La correlazione seriale determina $\sigma_{\bar{Y}_i}^2$ modificando Y_{it} .

$$\sigma_{\bar{Y}_m}^2 = var(\bar{Y}_m)$$

$$\begin{aligned}
&= \text{var} \left(\frac{1}{T} \sum_{i=1}^T Y_{it} \right) = \frac{1}{T^2} \text{var}(Y_{m1} + \dots + Y_{mT}) \\
&= \frac{1}{T^2} \{ \text{var}(Y_{m1}) + \dots + \text{var}(Y_{mT}) + 2\text{cov}(Y_{m1}, Y_{m2}) + \dots + 2\text{cov}(Y_{m1}, Y_{mT}) \}
\end{aligned}$$

Se Y_{mt} fosse serialmente correlata, la formula consueta risulterebbe errata e sottostimerebbe la varianza di $\bar{Y}_m$ nel caso in cui le autocovarianze siano positive. Il vantaggio degli errori standard per dati raggruppati è che, operando al livello delle unità e delle loro medie $\bar{Y}_m$, non occorre preoccuparsi di stimare le autocovarianze sottostanti, le quali sono automaticamente stimate dalla formula dell'errore standard.

Di seguito, si riportano i calcoli:

$$\begin{aligned}
\bar{Y} &= \sqrt{\frac{S_{\bar{Y}_m}^2}{n}} \\
s_{\bar{Y}_m}^2 &= \frac{1}{n-1} \sum_{m=1}^n (\bar{Y}_m - \bar{Y})^2 = \frac{1}{n-1} \left(\sum_{m=1}^n \frac{1}{T} \sum_{t=1}^T (Y_{mt} - \bar{Y})^2 \right) \\
&= \frac{1}{n-1} \sum_{m=1}^n \left(\frac{1}{T} \sum_{t=1}^T (Y_{mt} - \bar{Y}) \right) \left(\frac{1}{T} \sum_{s=1}^T (Y_{ms} - \bar{Y}) \right) \\
&= \frac{1}{n-1} \sum_{m=1}^n \frac{1}{T^2} \sum_{t=1}^T \sum_{s=1}^T (Y_{ms} - \bar{Y})(Y_{mt} - \bar{Y}) \\
&= \frac{1}{T^2} \sum_{t=1}^T \sum_{s=1}^T \left[\frac{1}{n-1} \sum_{m=1}^n (Y_{ms} - \bar{Y})(Y_{mt} - \bar{Y}) \right]
\end{aligned}$$

Il termine finale tra quadre, $\frac{1}{n-1} \sum_{m=1}^n (Y_{ms} - \bar{Y})(Y_{mt} - \bar{Y})$, stima l'autocovarianza tra Y_{ms} e Y_{mt} .

Infine, può essere utile citare due valori con cui è possibile misurare la bontà della regressione, ovvero il SER e l' R^2 .

Il SER misura la dispersione delle Y attorno alla retta di regressione:

$$SER = \sqrt{\frac{1}{n-k-1} \sum_{m=1}^n \widehat{u}_{mt}^2}$$

L' R^2 rappresenta invece la frazione della varianza spiegata dalla regressione, è priva di unità e può variare tra 0 (nessun adattamento) e 1 (perfetto adattamento).

$$R^2 = \frac{ESS}{TSS} = 1 - \frac{SSR}{TSS}$$

$$ESS = \sum_{m=1}^n (\widehat{Y}_m - \bar{Y})^2, \quad SSR = \sum_{m=1}^n (\widehat{u}_m)^2, \quad TSS = \sum_{m=1}^n (Y_m - \bar{Y})^2$$

Aggiungendo un ulteriore regressore, l' R^2 aumenta sempre: il modello sembra risultare migliore ma in realtà “nasconde” lo scostamento esistente tra la retta di regressione e i dati osservati.

Riferendosi a un modello a più regressori (k), è possibile ovviare a questo problema utilizzando l' R^2 corretto:

$$R^2 = 1 - \left(\frac{n-1}{n-k-1} \right) \frac{SSR}{TSS}$$

3.1.2 Stima empirica

3.1.2.1 Generale

Durante la costruzione del modello empirico, si è incrementalmente incluso un maggior numero di variabili di controllo con lo scopo di individuare quali siano le specificazioni più esaustive e osservare la direzione delle variazioni del coefficiente relativo alla banda larga ultraveloce.

Alcuni comuni potrebbero essere particolarmente attrattivi per le nuove startup a causa di ragioni geografiche o la vicinanza ad attività economiche complementari. Esistono infatti molte variabili non osservabili, presumibilmente incluse nel termine di errore del modello, che potrebbero essere correlate all'espansione dell'UBB, rendendo così le stime incoerenti. Inoltre, ogni anno, mantenendo costanti le condizioni al contorno, si attende un aumento complessivo delle startup innovative e quindi anche l'effetto temporale non può essere trascurato.

Includendo nelle regressioni gli effetti predeterminati del comune, insieme ai controlli osservabili e variabili nel corso degli anni, non sarà necessario che i tempi di introduzione delle connessioni ultraveloci non siano correlati alle caratteristiche del comune.

Il modello di partenza è il seguente²⁰.

$$Y_{m,t} = \beta_0 + \beta_1 UBB_{m,t} + \beta_2 X_{r,t} + \alpha_t + \gamma_m + u_{m,t}$$

L'obiettivo dell'analisi empirica è verificare un'ipotesi (H_0), quale $\beta_1 = 0$, determinando sperimentalmente se sia corretta oppure no. L'ipotesi alternativa sarà pertanto $\beta_1 \neq 0$.

A questo scopo si utilizza la t-student: $t = \frac{\text{stimatore} - \text{valore ipotizzato}}{\text{errore standard dello stimatore}} = \frac{\hat{\beta}_1 - \beta_{1,0}}{SE(\hat{\beta}_1)}$

²⁰ Dove le variabili identificano:

- $Y_{m,t}$ il numero di nuove startup innovative in un certo comune m e anno t .
- $UBB_{m,t}$ la variabile principale di interesse che indica l'accesso alla connessione ultraveloce nel comune m nell'anno t . È una dummy e assume valore 1 se il comune ha accesso alle connessioni ultraveloci
- β_1 la pendenza della retta di regressione, ovvero l'effetto atteso su $Y_{m,t}$ di una variazione unitaria in $UBB_{m,t}$
- $X_{r,t}$ un vettore contenente gli effetti fissi sia a livello di regione che di anno. Si è generato per identificare in modo univoco la combinazione tra le due variabili che lo compongono.
- α_t l'effetto fisso a livello di anno
- γ_m l'effetto fisso a livello di comune
- $u_{m,t}$ il residuo di regressione costituito da fattori omessi

dove $SE(\hat{\beta}_1)$ è la radice quadrata di uno stimatore della varianza della distribuzione campionaria di $\hat{\beta}_1$. Al fine di consentire la correlazione in un determinato cluster, si è incluso gli errori standard clustered per comune.

Si rifiuta H_0 a livello di significatività del 5% se $|t| > 1,96$, mentre dell'1% se $|t| > 2,58$.

Un'affermazione equivalente può essere sostenuta utilizzando il p-value. Tale valore è la probabilità di osservare un altro valore di $\hat{\beta}_1$ diverso da 0 almeno quanto la stima realmente ottenuta con i nostri dati, supponendo che l'ipotesi nulla sia corretta. Si rifiuta H_0 al livello di significatività dell' $x\%$ se il valore p è minore di quel valore. Queste procedure si affidano all'approssimazione che $\hat{\beta}_1$ sia distribuito normalmente; ciò si verifica in campioni ad elevata numerosità e nel caso trattato, n è pari a 55111, ovvero grande abbastanza per avere un'approssimazione eccellente.

TABELLA 3-2 MODELLO LINEARE E LOGARITMICO, CON UN SINGOLO REGRESSORE E FIXED EFFECTS

	(1)	(2)	(3)	(4)	(5)	(6)
	startup	startup	startup	startup	log(startup)	log(startup)
UBB	0.588*** (0.0643)	0.589*** (0.0632)	0.266*** (0.0470)	0.259*** (0.0559)	0.0809*** (0.00336)	0.0778*** (0.00371)
cons	0.0180*** (0.0019)	0.00229*** (0.00057)	0.000583 (0.042)	0.0580 (0.053)	0.00129 (0.0022)	0.0225*** (0.0027)
Year FE		X	X		X	
Municipality FE			X	X	X	X
Year-Region FE				X		X
N	55111	55111	55111	55111	55111	55111
R^2	0.004	0.004	0.673	0.673	0.551	0.556

Standard errors in parentheses

Source: auto.dta

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Come è facilmente osservabile dalla [Tabella 3-2](#), i coefficienti relativi alla copertura UBB sono sempre positivi ed è possibile rifiutare l'ipotesi nulla a livello di significatività del 5% in tutte le quattro le regressioni riportate.

Nella prima colonna [1] non si cattura alcun effetto fisso. La presenza della tecnologia UBB corrisponde a un aumento pari al 0,588 delle iscrizioni annuali al Registro delle Imprese da parte di startup innovative.

L'introduzione degli effetti temporali [2] incide leggermente sul valore dei coefficienti relativi alla copertura con fibra. Maggiormente impattante si rivela essere l'effetto fisso a livello di comune [3], il quale rende evidente come l'effetto precedente fosse notevolmente sovrastimato: il valore del coefficiente diminuisce complessivamente del 54,84%. Inoltre, nelle regressioni [1] e [2], anche il valore dell' R^2 (pari a 0,04) evidenziava scarsa bontà del modello.

Includendo una nuova variabile che misura l'effetto fisso relativo all'anno e alla regione [4], il coefficiente subisce un'ulteriore riduzione fino a 0,259. Nelle regressioni [5] e [6], il modello utilizzato non è più lineare essendo la variabile dipendente sfruttata logaritmica²¹. L'effetto, sul numero di nuove startup innovative, di una variazione in UBB dipende dal valore di UBB, ovvero l'effetto marginale di X non è più costante. Essendo un modello log-lineare, una variazione unitaria nella copertura con le reti UBB ($\Delta UBB = 1$) produce una variazione di $100\beta_1\%$ nella variabile dipendente. La variabile UBB continua a essere statisticamente significativa, anche se ulteriormente attenuata; mentre il relativo errore standard è diminuito ma meno che proporzionalmente.

Includendo gli effetti temporali annuali [5], risulta che la presenza della banda larga ultraveloce corrisponda a un aumento dell'8,09% delle nuove startup innovative. Considerando invece il vettore che identifica anche gli effetti relativi alla regione di appartenenza [6], l'impatto si riduce di 0,31 punti percentuali.

TABELLA 3-3 MODELLO LOGARITMICO CON REGRESSORI MULTIPLI E FIXED EFFECTS

	(1)	(2)	(3)	(4)
	log(startup)	log(startup)	log(startup)	log(startup)
UBB	0.0798*** (0.00334)	0.0802*** (0.00334)	0.0746*** (0.00366)	0.0431*** (0.00399)
OLT distance	0.00025*** (0.000028)		0.00089*** (0.00009)	0.00103*** (0.000087)
ADSL 20Mbit/s		-0.0507*** (0.00246)	-0.0472*** (0.00262)	-0.0362*** (0.00261)
sede OLT				0.0915*** (0.00599)
_cons	-0.0144*** (0.00314)	0.0302*** (0.00211)	0.0611*** (0.00384)	0.0462*** (0.00358)
Year FE	X	X		
Municipality FE	X	X	X	X
Year-Region FE			X	X
N	55111	55111	55111	55111
R ²	0.552	0.553	0.559	0.563

Standard errors in parentheses

Source: auto.dta

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Terminata la prima stima decisamente grossolana, si procede ad aggiungere variabili di controllo, come riportato in [Tabella 3-3](#). Si è pertanto incluso la distanza OLT, ADSL 20Mbit/s e la sede OLT. Queste variabili risultano fortemente significative in tutte le quattro regressioni e studiare cosa accade al coefficiente di UBB può essere un interessante spunto

²¹ La variabile create è $\log(startup)$ pari a $\log(startup + 1)$

di riflessione.

Nella colonna [1], includendo la distanza OLT, l'effetto della banda larga ultraveloce diminuisce passando dal valore precedente di 0,0809 (riportato nella regressione [5] della [Tabella 3-2](#)), a 0,0798. Ciò accade a causa della correlazione negativa esistente (pari a -0,2881) tra la copertura UBB e la distanza OLT, trattata precedentemente nell'analisi descrittiva e riportata in [Figura 2-29](#). La lunghezza dell'ultimo miglio approssima il costo di implementazione necessario per raggiungere i potenziali clienti con la linea ottica. In questa prospettiva, la distanza tra ogni comune e la sua OLT più vicina costituisce una naturale variazione di costo per il dispiegamento della rete UBB. Tale chilometraggio è sempre significativo ma sorprendentemente ha un impatto positivo, seppur lieve, sulla nascita di startup innovative: un aumento di 1km della distanza minima tra un comune e la sede OLT più vicina, ha un effetto sulla variabile dipendente pari a 0,025%.

L'introduzione della copertura ADSL nella regressione [2] comporta un lieve aumento del coefficiente UBB. La correlazione tra le due variabili è pari a 0,291. L'ADSL impatta negativamente sulla nascita di nuove startup innovative e per questo se non si includesse nella specificazione, l'effetto della variabile d'interesse risulterebbe sottostimato.

Nelle regressioni successive ([3], [4]), si cattura l'effetto fisso a livello anno-regione e sono incluse contemporaneamente più variabili di controllo.

Tra ADSL 20Mbit/s e la distanza OLT esiste una correlazione negativa pari a -0,232 e l'effetto risultante (riportato nella colonna [3]), è una variazione in senso opposto dei due coefficienti, rispetto alle regressioni precedentemente trattate in cui erano non erano incluse congiuntamente. La presenza della rete UBB comporta un incremento del 7,46% della variabile dipendente.

Infine, nella regressione [4] è inclusa la sede OLT. L'effetto della banda larga ultraveloce diminuisce del 42,1% (rispetto alla precedente specificazione) e ciò è probabilmente dovuto alla forte correlazione esistente con la variabile aggiunta, pari a 0,608. La sede OLT ha un impatto fortemente significativo e positivo sulla variabile dipendente: la presenza di almeno un CO dotato di OLT all'interno di un comune, genera un aumento delle iscrizioni del 9,15%.

L'impatto della distanza OLT risulta altamente significativo e positivo anche nelle ultime due regressioni, sebbene il valore del coefficiente sia molto basso: a un incremento unitario della distanza minima, corrisponde un aumento percentuale delle iscrizioni di startup innovative compreso tra lo 0,0889% e lo 0,103%.

Prima di procedere, si cerca di valutare qualitativamente quali variabili non osservabili abbiano il potenziale di influenzare sia la nascita delle startup sia la diffusione della fibra. Sotto questo punto di vista, oltre all'UBB, è probabile che anche l'ADSL sia una variabile endogena e ciò potrebbe anche giustificarne il relativo coefficiente negativo. A inizio del 2013, la maggior parte dei comuni sono già coperti da tale tecnologia e includendo gli effetti fissi a livello di comune, il coefficiente catturerà variazioni positive soltanto nei comuni in cui precedentemente non era presente. Pertanto, gli incrementi inclusi nella stima econometrica riguarderanno soltanto quei piccoli comuni, caratterizzati al tempo stesso da scarsa numerosità di startup innovative. Al contrario, i comuni in cui sono già presenti molteplici startup, presumibilmente vantano già la tecnologia ADSL e non saranno catturati nella stima. In conclusione, si è deciso di focalizzare la nostra strategia empirica soltanto sulla fibra per cui è possibile catturare la variazione nel tempo, a partire dall'ingresso avvenuto nel 2015.

Un'ulteriore variabile potenzialmente endogena potrebbe essere l'FTTH, riferita alla banda

larga ultraveloce di qualità superiore. Tuttavia, si ritiene comunque interessante valutare l'effetto della fibra "più potente" e le variazioni conseguenti della variabile d'interesse. Questa architettura permette di raggiungere la massima velocità di trasmissione disponibile e potrebbe a sua volta generare un impatto aggiuntivo rispetto all'effetto della variabile principale d'interesse. Si assume pertanto che quest'ultima tecnologia si diffonda per caratteristiche idiosincratice del comune e che tale correlazione sia time invariant, in modo da catturare la potenziale endogeneità includendo gli effetti fissi. Se questa ipotesi non fosse rispettata, anche l'FTTH risulterebbe endogena.

La sede OLT e la distanza OLT risultano altamente e negativamente correlate, poiché nei comuni in cui la dummy è vera (ovvero dove è presente almeno un CO dotato di OLT), la distanza minima in chilometri tra ogni comune e il terminale di linea ottica più vicino è pari a zero; mentre ha valore positivo negli altri. Per evitare che i loro coefficienti siano stimati in modo impreciso e con alta varianza campionaria, si è quindi deciso di includere solo la variabile relativa alla distanza.

TABELLA 3-4 MODELLO LOGARITMICO CON ULTERIORI VARIABILI DI CONTROLLO

	(1) log(startup)	(2) log(startup)	(3) log(startup)	(4) log(startup)	(5) log(startup)
UBB	0.0680*** (0.00325)	0.0654*** (0.00399)	0.0797*** (0.00449)	0.0761*** (0.00488)	0.0666*** (0.00448)
OLT distance	0.000135*** (0.000024)	0.000312*** (0.000075)	0.000254*** (0.000036)	0.000936*** (0.00012)	0.000756*** (0.00011)
gdu	-0.0674*** (0.00409)	-0.0749*** (0.00475)	0 (.)	0 (.)	0 (.)
università	0.536*** (0.0396)	0.530*** (0.0393)	0 (.)	0 (.)	0 (.)
aree bianche	0.00456 (0.00293)	-0.00271 (0.00330)	0 (.)	0 (.)	0 (.)
patents	8.38e-08 (0.000001)	0.0000006 (0.000001)	0.0000003 (0.0000008)	0.000001 (0.0000009)	0.0000003 (0.0000008)
FTTH					0.326*** (0.0347)
_cons	0.149*** (0.00930)	0.188*** (0.0126)	-0.0147*** (0.00390)	0.0297*** (0.00342)	0.0260*** (0.00350)
Year FE	X		X		
Municipality FE			X	X	X
Year-Region FE		X		X	X
N	55111	55111	55111	55111	55111
R ²	0.240	0.250	0.552	0.558	0.571

Standard errors in parentheses

Source: auto.dta

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Come evidenziato in [Tabella 3-4](#), la presenza della banda larga ultraveloce è sempre altamente significativa ($p\text{-value} < 1\%$) e genera un effetto compreso tra 6,54% e 7,97%. Nella colonna [1] si includono gli effetti temporali annuali, mentre nella [2] gli effetti fissi sono considerati su base anno-regione; dalla terza in poi si inseriscono gli effetti comunali e nell'ultima si aggiunge un ulteriore variabile di controllo, FTTH. Essa identifica l'accesso alla tecnologia Fiber to the Home nel comune m nell'anno t regressione.

La distanza OLT è fortemente significativa in tutte le regressioni e il relativo coefficiente aumenta nelle specificazioni in cui si includono gli effetti temporali legati ad anno-regione ([2], [4] e [5]). Nelle ultime due regressioni la variazione unitaria della lunghezza genera un incremento che va dallo 0,94% allo 0,75% del numero di nuove startup innovative. Tali valori risultano essere inferiore di almeno un ordine di grandezza rispetto ai coefficienti delle altre variabili incluse.

Nelle prime due regressioni, in cui sono esclusi gli effetti fissi legati al comune, sono presenti covariate di controllo che conseguentemente risultano omesse poiché altamente collineari. Esse (GDU, università e aree bianche) sono variabili relative al comune e costanti nel tempo. Pertanto, le descrizioni che seguono riguardano esclusivamente le colonne [1] e [2].

Il GDU, il cui dominio è compreso tra 1 (alta urbanizzazione) e 3 (bassa urbanizzazione), è altamente significativo e come si attendeva, di segno negativo poiché nei comuni scarsamente popolati il numero di startup innovative tenderà presumibilmente ad essere inferiore. Una variazione unitaria di categorizzazione, ad esempio da bassa a media, corrisponde a un incremento di startup compreso tra il 6,74% e il 7,49%.

L'effetto complessivamente maggiore si registra per la variabile dummy inerente alla presenza di università. Tale variabile assume valore 1 se nell'anno t nel comune m , esiste un'università avente almeno un iscritto. L'effetto è dunque altamente significativo e la relativa presenza implica un aumento superiore al 53% del numero di startup nascenti. Vista l'elevata incidenza della variabile, in [Tabella 3-8](#) si analizzeranno i dati empirici suddividendo i comuni in base ai valori della dummy corrispondente.

Contrariamente a quanto ci si potesse aspettare, il coefficiente relativo alle zone a fallimento di mercato, aree bianche, non è significativamente diverso da zero a livello del 5%. Pertanto, non è possibile affermare che nel caso in cui un'area bianca sia raggiunta dalla fibra, le startup innovative nascenti tenderanno a crescere.

La variabile patents rappresenta le registrazioni brevettuali in una certa provincia e anno t . Non è altamente correlata agli effetti fissi comunali e quindi è inclusa anche nelle regressioni [3], [4] e [5]. Il relativo coefficiente non è mai significativamente diverso da zero ($p\text{-value} > 5\%$).

L'effetto della banda larga ultraveloce risulta sottostimato nelle regressioni [1] e [2], in cui risultava rispettivamente pari a 6,8% e 6,54%. Tale affermazione è giustificata dai risultati ottenuti considerando gli effetti fissi a livello di comune, dato che si registra un incremento di quasi un punto percentuale dell'effetto della variabile d'interesse (regressioni [3] e [4]). Infine, aggiungendo l'FTTH ([5]), l'impatto dell'UBB diminuisce fino al 6,66% poiché catturato da quest'ultima, la quale ingloba la parte della fibra di qualità superiore.

Non si è inclusa la distanza OPB, nonostante possa avere un impatto fortemente significativo. Omettendo la distanza OPB a favore degli effetti fissi a livello di comune, si sta però implicitamente assumendo che qualunque sia il legame esistente tra la localizzazione dei nodi OPB e la nascita di nuove imprese, tale correlazione sia costante nel tempo.

Concludendo, si ritiene che le regressioni migliori includano l'effetto temporale a livello di anno-regione e municipality FE (colonne [4] e [5]); perciò da qui in avanti si sfrutterà tale specificazione scomponendo il dataset e i risultati in base a particolari suddivisioni.

Il modello che verrà preso a riferimento è dunque il seguente.

$$Y_{m,t} = \beta_0 + \beta_1 UBB_{m,t} + \beta_2 X_{r,t} + \beta_3 X_{3m,t} + \beta_4 X_{4p,t} + \beta_5 X_{5m,t} + \gamma_m + u_{m,t}$$

Per riepilogare, la variabile $X_{3m,t}$ rappresenta la distanza OLT, $X_{4p,t}$ il numero di brevetti a livello provinciale, $X_{5m,t}$ la tecnologia FTTH.

L'effetto fisso inerente al comune (γ_m) sarà dunque perennemente incluso nelle specificazioni successive e per quanto riguarda l'effetto temporale si considererà a livello di anno-regione ($X_{r,t}$).

Concludiamo questa prima analisi generale quindi stimando che l'effetto dell'adozione della banda larga sui comuni è positivo, statisticamente significativo e si attesta intorno al 6,66% quando consideriamo gli effetti fissi.

3.1.2.2 Settori

Come descritto nel capitolo 2, le startup appartengono a 5 macrosettori.

Vista l'alta rilevanza del settore dei servizi, il quale include il 77,19% delle registrazioni, si è analizzato separatamente confrontandolo con un'unica variabile comprendente le startup innovative appartenenti all'industria/artigianato, commercio, agricoltura/pesca e turismo.

Per quanto riguarda la variabile principale d'interesse, i risultati ottenuti nel settore dei servizi, riportati nella [Tabella 3-5](#), sono prossimi a quanto ottenuto nel caso generale descritto nella [Tabella 3-4](#). La presenza della fibra, in un certo comune, genera un incremento di iscrizioni di startup operanti nei servizi del 6,04% e 5,2% nel caso si includa anche la variabile relativa alla tecnologia FTTH. Per gli altri settori, la variabile UBB genera ancora un effetto positivo anche se con impatto minore: al dispiegamento delle reti corrisponde un aumento compreso tra il 2,62% e 2,17% delle iscrizioni annuali. Confrontando le analoghe specificazioni, l'effetto risulta più che dimezzato, ovvero inferiore rispettivamente del 56,6% e 58% passando dalla prima alla seconda suddivisione.

La distanza OLT risulta sempre altamente significativa ($p\text{-value} < 1\%$) e i relativi coefficienti variano tra le specificazioni in modo analogo alla variabile d'interesse. L'effetto maggiore si rileva nella specificazione [1], in cui una variazione unitaria è associata a un incremento delle nuove startup nel settore dei servizi dello 0,12%. Per gli altri settori si registra ancora un impatto positivo di valore compreso tra lo 0,06% e 0,05%.

La tecnologia FTTH presenta, ancora una volta, il coefficiente maggiore di un ordine di grandezza rispetto alla variabile UBB. Per le startup che offrono servizi, la presenza della connessione di alta qualità in un certo comune, genera un aumento del 28,2%, mentre negli altri settori non raggiunge il 15%. L'inclusione di questa variabile cattura una parte dell'effetto della fibra, il cui relativo coefficiente diminuisce del 13,9% nella prima suddivisione e del 17,2% nella seconda. Le ulteriori considerazioni riguardanti le differenze intra-settoriali sono in linea con quanto evidenziato per la variabile principale di interesse.

Le variabili relative a gdu, università, aree bianche e patents perdono di significatività rispetto al caso generale ($p\text{-value} > 5\%$). Per quanto riguarda il grado di urbanizzazione e la presenza di università, si è verificata una netta diminuzione dei coefficienti, non seguita

però da una conseguente riduzione della deviazione standard, la quale è rimasta dello stesso ordine di grandezza riscontrato nel caso generale.

TABELLA 3-5 SUDDIVISIONE SETTORIALE

	(1) log(servizi)	(2) log(altri)	(3) log(servizi)	(4) log(altri)
UBB	0.0604*** (0.00454)	0.0262*** (0.00278)	0.0520*** (0.00415)	0.0217*** (0.00256)
OLT distance	0.00116*** (0.000161)	0.000618*** (0.000100)	0.000947*** (0.000137)	0.000504*** (0.0000838)
gdu	-0.00205 (0.00572)	-0.00236 (0.00362)	-0.00220 (0.00547)	-0.00244 (0.00349)
università	-0.00738 (0.0185)	-0.00665 (0.0103)	-0.00421 (0.0173)	-0.00498 (0.00969)
aree bianche	0 (.)	0 (.)	0 (.)	0 (.)
patents	0.0000012 (0.000001)	0.00000013 (0.00000059)	0.00000034 (0.00000073)	-0.00000032 (0.00000058)
FTTH			0.282*** (0.0331)	0.149*** (0.0228)
_cons	0.0178 (0.0152)	0.0121 (0.00974)	0.0136 (0.0145)	0.00995 (0.00941)
Year FE				
Municipality FE	X	X	X	X
Year-Region FE	X	X	X	X
N	48385	48385	48385	48385
R ²	0.565	0.465	0.578	0.475

Standard errors in parentheses

Source: auto.dta

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

3.1.2.3 GDU

Si è poi suddiviso il database in base al grado di urbanizzazione di ciascun comune.

A differenza dei risultati commentanti precedentemente, la copertura UBB non è perennemente significativa nelle varie regressioni.

Nelle colonne pari della [Tabella 3-6](#) si è inclusa anche la variabile relativa alla tecnologia FTTH, mentre gli effetti temporali sono sempre a livello anno-regione.

Per quanto riguarda i comuni ad alta urbanizzazione, nella regressione [1] l'impatto della variabile UBB non risulta fortemente significativa. Precedentemente, nel caso generale, era superiore al 6,6% come registrato nella prima regressione di [Tabella 3-4](#); adesso è significativamente diversa da zero solamente nei comuni mediamente e scarsamente popolati con effetti decisamente attenuati.

Al contrario per la distanza OLT, le considerazioni sono opposte rispetto quanto appena detto riguardo la variabile principale di interesse. Si riscontrano valori di un ordine di grandezza superiore rispetto al caso generale ma tali coefficienti risultano significativi soltanto nei comuni ad alta urbanizzazione.

Nelle regressioni [1] e [2], si constata che un incremento di 1km della distanza minima tra il comune e la sede OLT più vicina, provoca un aumento compreso tra lo 0,805% e lo 0,709% delle startup innovative nascenti. Tale diminuzione, riportata nella colonna [2], di poco meno di un decimo di punto percentuale, è dovuta all'inclusione della variabile relativa all'FTTH. La relazione attesa tra la distanza OLT e la variabile dipendente era negativa, ovvero si riteneva presumibilmente che la crescita del numero di startup fosse generata da una diminuzione (e non da un aumento) della distanza minima. Tuttavia, si constata tale relazione positiva anche nelle regressioni precedenti.

La presenza della tecnologia FTTH, all'interno di un certo comune densamente popolato, genera un incremento del 32,3% delle iscrizioni.

Come per il caso generale, l'impatto delle registrazioni brevettuali a livello provinciale non risulta mai significativamente diverso da zero in nessuna delle regressioni descritte in tabella.

Nei comuni a media urbanizzazione si verifica il maggior impatto della banda larga ultraveloce. Il dispiegamento delle reti in un'unità territoriale comporta un effetto di circa lo 0,35%. Includendo l'FTTH (regressione [4]) si verifica un leggero aumento della variabile di interesse. La fibra "più potente" risulta fortemente significativa e genera un aumento delle nuove startup innovative di poco meno del 20%. Sia in questa suddivisione che nella successiva, il p-value relativo alla distanza minima tra comune e sede OLT risulta sempre superiore al 5%.

Le unità territoriali a basso grado di urbanizzazione (le quali risultano il 73,2% del totale) presentano effetti opposti per quanto riguarda la copertura UBB e FTTH, rispetto ai comuni densamente popolati. L'effetto dell'UBB è infatti significativo e pari all'1,6%. Tale impatto è diminuito di circa 5 punti percentuali rispetto al caso generale ([Tabella 3-4](#)), ed è l'unica variabile che risulta statisticamente diversa da zero a livello del 5% nelle regressioni [5] e [6].

TABELLA 3-6 SUDDIVISIONE PER GRADO DI URBANIZZAZIONE

	(1)	(2)	(3)	(4)	(5)	(6)
	log(startup)	log(startup)	log(startup)	log(startup)	log(startup)	log(startup)
	Alta Urbanizzazione		Media Urbanizzazione		Bassa Urbanizzazione	
UBB	0.00337 (0.0483)	0.0438 (0.0489)	0.035*** (0.0065)	0.036*** (0.0065)	0.016*** (0.0036)	0.017*** (0.0037)
OLT distance	0.0081** (0.0029)	0.0071** (0.0027)	0.00031 (0.0002)	0.00023 (0.0002)	0.00007 (0.0004)	0.00006 (0.0001)
gdu	0 (.)	0 (.)	0 (.)	0 (.)	0 (.)	0 (.)
università	0 (.)	0 (.)	0 (.)	0 (.)	0 (.)	0 (.)
aree bianche	0 (.)	0 (.)	0 (.)	0 (.)	0 (.)	0 (.)
patents	-0.0001 (0.0001)	-0.0007 (0.0004)	-0.0001 (0.0001)	-0.0001 (0.0001)	-8.25e-9 (0.0001)	-1.11e-8 (0.0001)
FTTH		0.323*** (0.0829)		0.198*** (0.0363)		-0.0129 (0.0130)
_cons	0.529*** (0.0559)	0.358*** (0.0512)	0.083*** (0.0065)	0.076*** (0.0064)	0.012*** (0.0018)	0.014*** (0.0018)
Year FE						
Municipality FE	X	X	X	X	X	X
Year-Region FE	X	X	X	X	X	X
N	1771	1771	17907	17907	35433	35433
R ²	0.819	0.824	0.406	0.411	0.258	0.258

Standard errors in parentheses

Source: auto.dta

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

3.1.2.4 Aree bianche, università e ricchezza locale

Si focalizza adesso l'analisi sulle aree bianche, particolarmente interessanti poiché sono le zone che hanno maggiormente attratto gli investimenti statali poiché i governi mirano a perseguire una diffusione capillare della banda larga ultraveloce.

Anche se la maggior parte delle aree bianche sono a bassa urbanizzazione, sarebbe un errore ipotizzare che tutti questi comuni appartengano a questa categoria e perciò è utile approfondirle in separata sede. In effetti le due caratterizzazioni non sono così strettamente legate come si potrebbe pensare, poiché esistono anche comuni a media o alta urbanizzazione che in passato non hanno beneficiato del dispiegamento della fibra.

TABELLA 3-7 FOCUS SU AREE BIANCHE

	(1) log(startup) Aree bianche	(2) log(startup)	(3) log(startup) Aree grigie e nere	(4) log(startup)
UBB	0.0684*** (0.00592)	0.0573*** (0.00597)	0.0778*** (0.00477)	0.0782*** (0.00462)
OLT distance	0.000663*** (0.000162)	0.000612*** (0.000158)	0.00100*** (0.000107)	0.000694*** (0.0000918)
gdu	0 (.)	0 (.)	0 (.)	0 (.)
università	0 (.)	0 (.)	0 (.)	0 (.)
aree bianche	0 (.)	0 (.)	0 (.)	0 (.)
patents	-0.000001 (0.000001)	-0.000001 (0.000001)	0.000002* (0.000001)	9.54e-08 (0.000001)
FTTH		0.133*** (0.0248)		0.537*** (0.0381)
_cons	0.0309*** (0.00505)	0.0218*** (0.00517)	0.0177*** (0.00358)	0.01000** (0.00363)
Year FE				
Municipality FE	X	X	X	X
Year-Region FE	X	X	X	X
N	18898	18898	36213	36213
R ²	0.498	0.503	0.579	0.603

Standard errors in parentheses

Source: auto.dta

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Nella [Tabella 3-7](#) si riportano le stime empiriche ottenute. Nelle aree bianche, la variabile d'interesse risulta statisticamente significativa a livello dell'1% ed un incremento unitario genera un effetto compreso tra il 6,84% e il 5,73%, minore rispetto al caso generale in cui si registravano valori compresi tra il 7,61% e il 6,66% ([Tabella 3-4](#)). La distanza OLT, oltre a essere fortemente significativa, è quasi esattamente costante (sia per quanto riguarda il valore del coefficiente che la deviazione standard) tra le specificazioni [1] e [2], nonostante la variabile FTTH sia inclusa solo nella seconda. La presenza di tale tecnologia genera un incremento pari al 13,3% delle iscrizioni di startup innovative.

Confrontando tali valori con i risultati registrati in [Tabella 3-6](#), si evidenzia che sebbene l'effetto della fibra ad alta qualità sia diminuito rispetto alle zone densamente popolate, è ancora presente una forte significatività, diversamente da quanto accadeva nelle zone a bassa urbanizzazione. Anche l'impatto della variabile principale d'interesse risulta essere aumentato di circa 4,6 punti percentuali.

Per quanto riguarda i restanti comuni, il dispiegamento dell'UBB genera un effetto di circa il 7,8% sulle startup innovative nascenti. Anche la distanza OLT è sempre fortemente significativa e di portata maggiore rispetto a quanto rilevato nelle aree bianche, tuttavia i relativi effetti sono ancora inferiori allo 0,1%. La variabile che è cresciuta maggiormente passando da una suddivisione all'altra è FTTH. La presenza di tale tecnologia in un'area grigia o nera genera infatti un effetto aggiuntivo del 53,7%. Da segnalare che nella regressione [3], le registrazioni brevettuali risultano, per la prima volta, significative a livello del 5% e con un effetto positivo sulla variabile dipendente.

TABELLA 3-8 SUDDIVISIONE IN BASE A PRESENZA DI UNIVERSITÀ

	(1) log(startup) Con università	(2) log(startup)	(3) log(startup) Senza università	(4) log(startup)
UBB	0.213*** (0.0558)	0.272*** (0.0544)	0.0429*** (0.00313)	0.0406*** (0.00317)
OLT distance	0.00906*** (0.00136)	0.00750*** (0.00122)	0.000391*** (0.0000490)	0.000365*** (0.0000476)
gdu	0 (.)	0 (.)	0 (.)	0 (.)
aree bianche	0 (.)	0 (.)	0 (.)	0 (.)
patents	-0.000008 (0.000008)	-0.000015 (0.000008)	0.000001 (0.0000008)	0.000001 (0.0000008)
FTTH		0.582*** (0.0624)		0.0763*** (0.0191)
_cons	0.455*** (0.0592)	0.297*** (0.0673)	0.0182*** (0.00227)	0.0155*** (0.00224)
Year FE				
Municipality FE	X	X	X	X
Year-Region FE	X	X	X	X
N	1871	1871	53240	53240
R ²	0.745	0.762	0.318	0.320

Standard errors in parentheses

Source: auto.dta

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Nella [Tabella 3-8](#) si analizza la suddivisione dei comuni in base alla presenza o meno di un'università. Per quanto riguarda la copertura con la fibra ottica, si rileva alta significatività sia in presenza (regressioni [1] e [2]) che in assenza di un ateneo ([3] e [4]). L'effetto è però notevolmente diverso nei due casi: nei comuni, in cui è probabilmente presente una maggior conoscenza tecnica e scientifica, si verifica un impatto compreso tra il 21,3% e 27,2%, ben maggiore del 4,29% e 4,06% registratosi negli altri. Proprio come per la banda larga ultraveloce, anche i coefficienti delle variabili di controllo differiscono di almeno un

ordine di grandezza, passando dalla prima alla seconda caratterizzazione. Tuttavia, è importante sottolineare che esclusi i brevetti, tutte le variabili presenti risultano altamente significative ($p\text{-value} < 1\%$), anche nei comuni in cui non è presente un'università. Nel caso si considerasse l'effetto della tecnologia FTTH, l'aumento unitario della distanza minima dall'OLT più vicino diminuirebbe le iscrizioni di quel comune dello 0,75% in presenza di università e dello 0,0365% in assenza. Nella specificazione [2], la variabile FTTH presenta il più alto coefficiente riscontrato tra i risultati finora ottenuti. La presenza di tale tecnologia genera un incremento del 58,2%. Focalizzandoci invece sui comuni dove non è presente un ateneo, tale variabile è ancora significativa ma si constata che il relativo coefficiente sia più che dimezzato.

Altri risultati interessanti sono ricavabili suddividendo i comuni in base alla ricchezza locale (Tabella 3-9). A questo scopo si sfruttano le statistiche fiscali delle dichiarazioni Irpef inerenti all'anno 2013, pubblicate dal Dipartimento delle Finanze. La variabile discriminante individuata è il reddito imponibile²² complessivamente generato in ogni comune. Le unità territoriali sono dunque suddivise in quattro macro-scaglioni: 3325 comuni hanno un reddito imponibile minore di 20 M€, 2420 compreso tra 20 M€ e 60M€, 1399 tra 60 M€ e 150 M€ e infine, 928 riescono a raggiungere e oltrepassare quota 150 M€ (Figura 3-1).

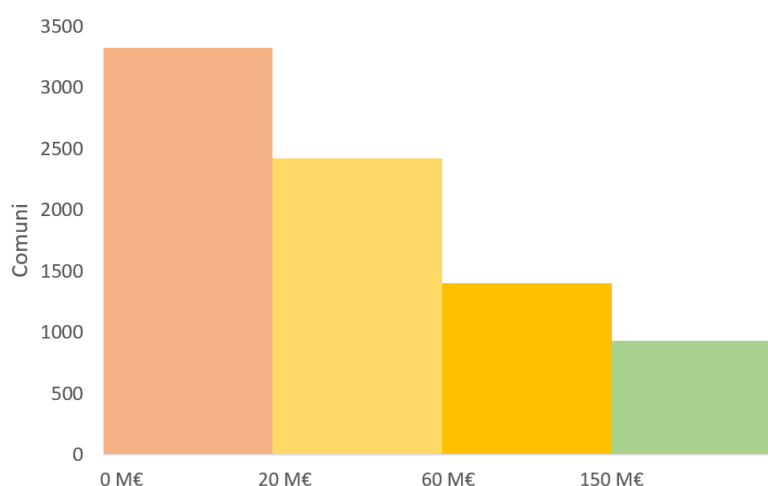


FIGURA 3-1 REDDITO IMPONIBILE

Dai risultati (Tabella 3-9) si evince che la banda larga ultraveloce non presenti un impatto significativo nei comuni appartenenti alle fasce medie di ricchezza. Presumibilmente l'effetto si consolida nei comuni ad alto reddito poiché oltre ad attività complementari, sono presenti anche le fonti finanziarie adeguate a garantire un flusso costante di investimenti a più startup innovative, riducendo il rischio per mezzo della diversificazione. Tuttavia, anche nei comuni in cui è presente minor ricchezza, la variabile principale d'interesse torna a essere significativa (e positiva) ma l'impatto è diminuito dell'80%. Il raggiungimento con

²² Reddito sul quale è effettivamente calcolata l'imposta. Si ottiene come somma algebrica del reddito complessivo (comprensivo del credito d'imposta per fondi comuni) al netto delle perdite compensabili (nei limiti del credito d'imposta per fondi comuni), delle deduzioni per abitazione principale e degli oneri deducibili.

le reti UBB di un comune avente un reddito imponibile complessivamente minore di 20 M€, comporta dunque un incremento delle startup innovative dell'1,82%.

Per la fibra ad alta qualità si riscontra un significativo e marcato impatto soltanto nelle aree caratterizzate da un elevato reddito imponibile. L'installazione dell'FTTH genera un incremento delle startup innovative, aggiuntivo rispetto la fibra standard, di oltre il 50% e quindi si riscontra che per raggiungere la quota maggioritaria dei benefici indiretti in queste aree, sia necessario l'ulteriore aggiornamento della tecnologia. Nei restanti comuni invece l'effetto inerente alla Fiber To The Home non risulta significativamente diverso da zero.

TABELLA 3-9 SUDDIVISIONE PER FASCE DI REDDITO

	(1) log(startup) <i>Bassa</i>	(2) log(startup) <i>Medio-Bassa</i>	(3) log(startup) <i>Medio-Alta</i>	(4) log(startup) <i>Alta</i>
UBB	0.0182*** (0.0038)	0.0027 (0.0037)	-0.00029 (0.0082)	0.0934*** (0.02415)
OLT distance	0.00009* (0.000045)	-0.00002 (0.00005)	-0.00001 (0.00013)	0.00279*** (0.00065)
gdu	0 (.)	0 (.)	0 (.)	0 (.)
aree bianche	0 (.)	0 (.)	0 (.)	0 (.)
patents	-0.000006 (0.000008)	-0.000014 (0.000008)	0.000001 (0.00008)	-0.00004 (0.000035)
FTTH	0.01644 (0.0197)	-0.0229 (0.0132)	0.0403 (0.0360)	0.5257*** (0.0454)
_cons	0.0055*** (0.0018)	0.0130*** (0.0031)	0.0506*** (0.0062)	0.1267*** (0.0259)
Year FE				
Municipality FE	X	X	X	X
Year-Region FE	X	X	X	X
N	23735	15757	9342	6277
R ²	0.3373	0.2199	0.2630	0.6614

Standard errors in parentheses

Source: auto.dta

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

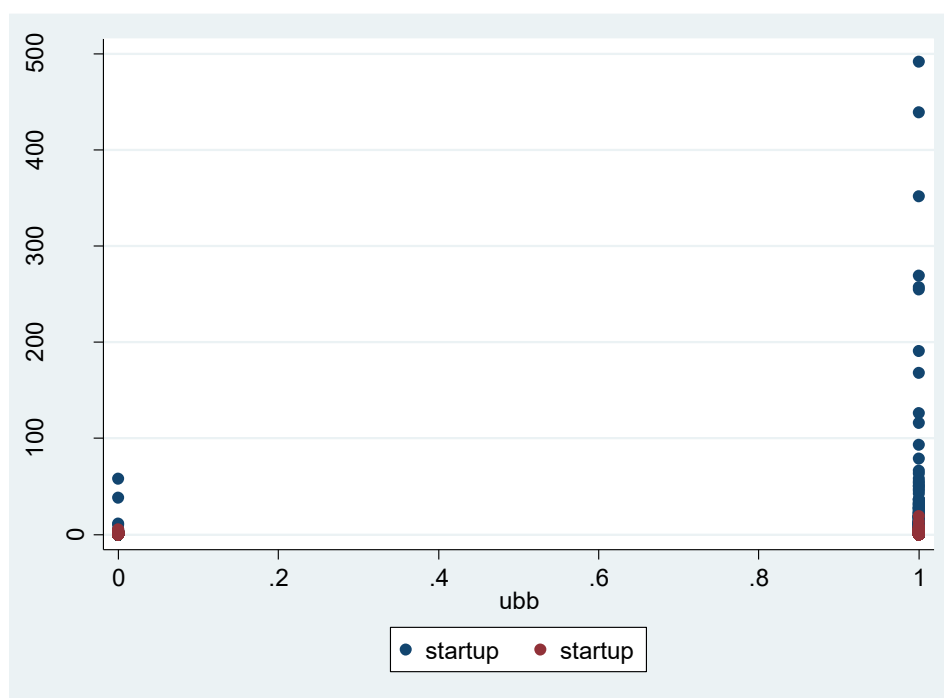


FIGURA 3-2 TASSI DI CRESCITA ETEROGENEI TRA COMUNI DIVERSI

Per proseguire l'analisi empirica, si posizionano su un grafico (Figura 3-2) i diversi comuni discriminando in base all'eventuale presenza o meno di un'università all'interno di essi. L'asse delle ordinate indica il numero di startup innovative iscritte, mentre sulle ascisse è presente la dummy relativa alla copertura con banda larga ultraveloce. I comuni in blu sono caratterizzati dalla presenza di almeno un ateneo, viceversa per i comuni di color amaranto.

Risulta evidente come un fattore, caratteristico di ogni territorio, renda i comuni molto eterogenei tra loro. L'inevitabile conseguenza è che le unità territoriali seguano trend diversi tra loro. Per proseguire lo studio, si fa riferimento alla precedente equazione riguardante gli effetti fissi: $Y_{X,t} = \alpha_X + \beta_1 UBB_{X,t} + u_{X,t}$. Discriminando le due unità territoriali, U (in cui è presente almeno un ateneo) e S (in cui non è presente), essi avranno intercetta diversa e costante nel tempo, ma probabilmente anche la pendenza non sarà la stessa. Presumibilmente l'effetto, all'interno del comune U, sulla variabile dipendente seguirà una retta più ripida mentre la situazione in S sarà stimabile con una retta dal coefficiente angolare sarà minore.

Quando ciò accade, gli effetti fissi a livello comunale non sono più rappresentabili con rette parallele. Una possibile soluzione è eseguire un'interazione tra time dummy e variabili di controllo che non variano nel tempo (Tabella 3-10). Ciò permette appunto di controllare per trend temporali condizionati a variabili di riferimento costanti (ovvero il grado di urbanizzazione, la presenza di università e la dummy relativa alle aree bianche). Una volta ottenute tali variabili, si includono nelle regressioni per poi analizzare nuovamente i coefficienti della variabile principale d'interesse.

TABELLA 3-10 CONTROLLO CON TREND TEMPORALI CONDIZIONATI A VARIABILI DI RIFERIMENTO

	(1) log(startup)	(2) log(startup)
UBB	0.0630*** (0.003)	0.0561*** (0.003)
OLT distance	0.0001*** (0.00003)	0.0001*** (0.00003)
gdu	0 (.)	0 (.)
università	0 (.)	0 (.)
aree bianche	0 (.)	0 (.)
patents	0.0000003 (0.000001)	-3.28e-08 (0.000001)
FTTH		0.197*** (0.0252)
_cons	0.0558*** (0.00259)	0.0490*** (0.00273)
Year-Region FE	X	X
Trend temporali condizionati a variabili time invariant	X	X
N	55111	55111
R ²	0.618	0.623

Standard errors in parentheses

Source: auto.dta

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Prima di procedere, si ricordano brevemente gli effetti registrati in precedenza. Nella [Tabella 3-4](#), nelle ultime due colonne si erano registrati effetti della variabile di interesse compresi tra il 7,61% e 6,66%. Per quanto riguarda la tecnologia FTTH, si era constatato che la copertura generasse un incremento del 32,6% delle startup innovative iscritte ogni anno al Registro delle Imprese.

Considerando che le caratteristiche idiosincratice di ogni comune, non influenzino solo l'intercetta, bensì anche la pendenza, i valori ottenuti assumendo rette parallele si riveleranno sovrastimati.

In quest'ultima analisi per la stima OLS ([Tabella 3-10](#)), si sono pertanto incluse ventuno nuove variabili generate dall'interazione tra le variabili di controllo e le time dummies corrispondenti all'arco temporale 2013-2019.

Ciò incrementa la fiducia nella direzione e con una certa accuratezza nella magnitudine dell'effetto della fibra: anche in queste specificazioni, la rete UBB genera ancora un effetto positivo e significativamente diverso da zero sulla variabile dipendente. Nella regressione [1] in cui non è incluso FTTH, si registra un impatto pari al 6,3%. L'inclusione della fibra "più potente" riduce l'impatto di 0,69 punti percentuali, il quale si assesta al 5,61%. La variabile FTTH risulta anch'essa altamente significativa ($p\text{-value} < 1\%$) e il relativo coefficiente è circa 3,5 volte superiore al valore relativo alla banda larga ultraveloce. La distanza OLT è fortemente significativa, sebbene l'impatto si assesti sullo 0,01%. Confrontando i risultati ottenuti nella regressione [2], rispetto alla situazione senza controlli con trend temporali condizionati a variabili di riferimento, risulta che l'effetto della fibra sia diminuito di un punto percentuale, mentre l'FTTH di quasi 13.

Includendo gli effetti fissi inerenti al comune, a livello anno-regione e i trend temporali condizionati alle variabili di riferimento, l'effetto risulta ancora fortemente positivo e si attesta intorno al 6%. Nei paragrafi successivi, tali evidenze saranno confrontate con quanto si ricaverà dall'event study e dalla stima a variabili strumentali (Tabella 3-13).

3.1.3 Event study

La diffusione della banda larga ultraveloce nei vari comuni italiani è un processo eterogeneo e diversificato nel tempo: alcuni comuni hanno ricevuto la fibra prima e altri in momenti successivi. Considerando la casualità intrinseca all'istante in cui si verifica il trattamento, in letteratura è frequente l'utilizzo del metodo *event study* per stimarne l'impatto causale sulla variabile dipendente.

Con lo stimatore canonico OLS, si è effettuato un confronto tra due gruppi di comuni, discriminando in base al dispiegamento o meno delle reti UBB, senza però assumere il problema dell'endogeneità (ovvero si è ipotizzato che la fibra non fosse correlata a fattori dell'errore econometrico che influenzino a sua volta la nascita di startup innovative).

Nell'analisi difference-in-differences (DD) sono dunque inclusi due periodi di tempo ("pre" e "post") e due gruppi di elementi ("trattamento" e "controllo"). Anno per anno grazie al progredire del dispiegamento, nuovi comuni vengono raggiunti e di conseguenza possono cambiare l'insieme di appartenenza.

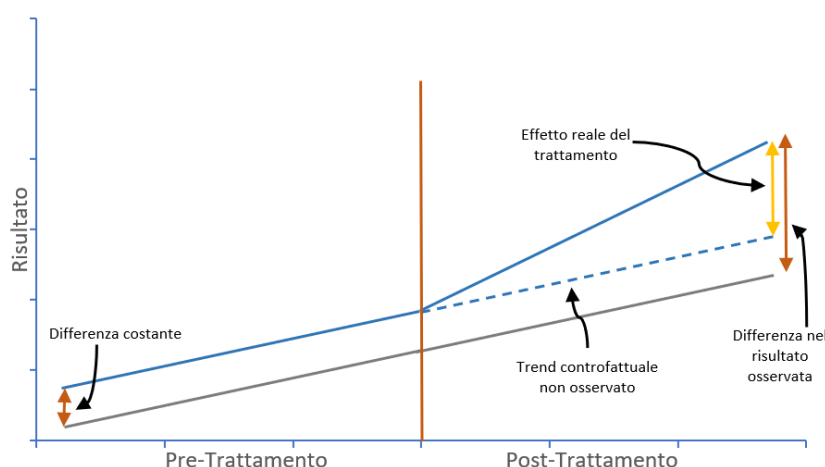


FIGURA 3-3 DD - INTERPRETAZIONE GRAFICA

Non essendo possibile fare affidamento su molte informazioni quando la tempistica del trattamento varia, si ammette che il campione sia *as good as random* condizionalmente agli effetti fissi di tempo e comune (Bertrand, Duflo e Mullainathan [2004]).

Il modello OLS contempla solo situazioni in cui nessuna delle unità sia trattata nel primo periodo, o tutte beneficino dello stesso stato di trattamento, aprendo così in modo fuorviante lo spazio per il confronto. In particolare, nel caso in cui gli effetti causali presentino una dinamica, la differenza tra i risultati post-trattamento riflette il cuneo tra gli effetti a lungo e a breve termine, poiché alcune unità sono state trattate prima. Tale differenza viene sottratta da una analoga per il secondo periodo, creando così una distorsione a breve termine. Quindi anche assumendo la fibra condizionata a effetti fissi esogena alla variabile dipendente, la stima potrebbe comunque essere scorretta nel caso in cui gli effetti si accumulino nel tempo. Lo stimatore OLS del parametro β_1 attribuisce peso maggiore agli effetti di breve termine, mentre sottostima gravemente la portata dell'impatto a lungo termine. A. Goodman-Bacon [2020] ha dimostrato che lo stimatore DD a due vie a effetti fissi è una media ponderata di tutti i possibili stimatori che confrontano i gruppi temporali tra loro. I pesi sono proporzionali alle dimensioni del gruppo e alla varianza della dummy del trattamento in ogni coppia, la quale risulta essere maggiore per le unità trattate al centro del panel. Se l'effetto legato alla variabile d'interesse si verificasse dopo un numero limitato di anni, il coefficiente potrebbe quindi essere sottostimato perché gli effetti di lungo periodo impattano marginalmente nella stima.

La media ponderata per la varianza dei parametri degli effetti del trattamento talvolta può presentare pesi negativi (Abraham e Sun [2018], Borusyak e Jaravel [2018]). Quando gli effetti del trattamento non cambiano nel tempo, la DD produce una media ponderata per la varianza tra gruppi e tutti i pesi sono positivi. I coefficienti negativi sorgono qualora gli effetti non siano costanti nel tempo. Quando le unità già trattate fungono da controlli, le variazioni dei loro effetti di trattamento nel tempo saranno poi sottratte dalla stima empirica. Ciò non implica un fallimento del modello, ma porta alla luce il rischio di riassumere gli effetti variabili nel tempo attraverso un singolo coefficiente: l'impatto medio stimato derivante dal trattamento potrebbe essere al di fuori del guscio convesso dei veri effetti dinamici. Per alleviare parzialmente il problema, si può includere un folto gruppo di controllo mirando ad ampliare l'insieme dei buoni confronti.

Il modello event study sfrutta una variabile dummy DmT pari a 1 per tutti gli anni successivi al trattamento. Essa è espressa in funzione del tempo relativo $T = t - t_0$, dove t è l'anno preso in considerazione, mentre t_0 è l'anno in cui il comune m ha effettivamente ricevuto la fibra. Nel caso in considerazione T include valori compresi tra -6 e 4.

Si eseguono dunque alcune regressioni in cui le variabili d'interesse sono identificate con la sommatoria sui vari anni in modo da ottenere un coefficiente per ogni periodo analizzato.

Dato che non tutti i comuni ricevono la fibra nello stesso momento anche il tempo relativo cambia. L'equazione di regressione è la seguente:

$$Y_{m,t} = \beta_0 + \sum_T \gamma_T \mathbb{I}\{T = t - t_0\} + \beta_2 X_{r,t} + \beta_3 X_{3m,t} + \beta_4 X_{4p,t} + \beta_5 X_{5m,t} + \gamma_m + u_{m,t}$$

Dove la variabile $X_{3m,t}$ rappresenta la distanza OLT, $X_{4p,t}$ il numero di brevetti a livello provinciale, $X_{5m,t}$ la tecnologia FTTH. Come controlli aggiuntivi costanti nel tempo si sono sfruttate le variabili relative a GDU, presenza di università e aree bianche. Si include poi

l'effetto fisso inerente al comune (γ_m) e l'effetto temporale a livello anno-regione ($X_{r,t}$). Come suggerito da Borusyak & Jaravel [2018] a causa della potenziale collinearità, è necessario non includere nella regressione le dummies relative all'inizio del periodo osservato e al tempo prima dell'avvenuto trattamento ($T = -6, T - 1$).

La variabile dipendente è il numero di nuove startup innovative annue, mentre la variabile d'interesse è la fibra. Le prime due specificazioni esaminate erano state rappresentate nelle colonne [4] e [5] della [Tabella 3-4](#). Negli ultimi due modelli si sfrutta quanto ottenuto nella [Tabella 3-10](#), controllando con trend temporali condizionati alle variabili prese a riferimento. Si ricorda inoltre che nella seconda e quarta specificazione è inclusa la copertura con tecnologia FTTH.

Il passo seguente è stimare i coefficienti associati alle dummies, rappresentarli graficamente e analizzarne la dinamica nel tempo sulla nascita di startup innovative.

I grafici di seguito riportati ([Figura 3-4](#), [Figura 3-5](#), [Figura 3-6](#) e [Figura 3-7](#)) presentano andamenti simili e si riscontrano vari indizi di potenziale endogeneità. Tali evidenze sono state rilevate in tutte le regressioni in quanto prima dell'anno t_0 , l'intervallo di confidenza non include lo zero in tre punti, corrispondenti ai periodi -5, -4 e -3.

È evidente, dunque, che già prima del dispiegamento della banda larga ultraveloce, fossero presenti dei trend positivi (anche se con ritorni negativi) che avevano contribuito ad aumentare il numero di startup innovative fino a tre anni prima dell'inizio del dispiegamento. Dopodiché si rileva un effetto cumulativo per costruzione che si è perpetuato nei periodi successivi e in precedenza non stimabile correttamente attraverso il coefficiente OLS.

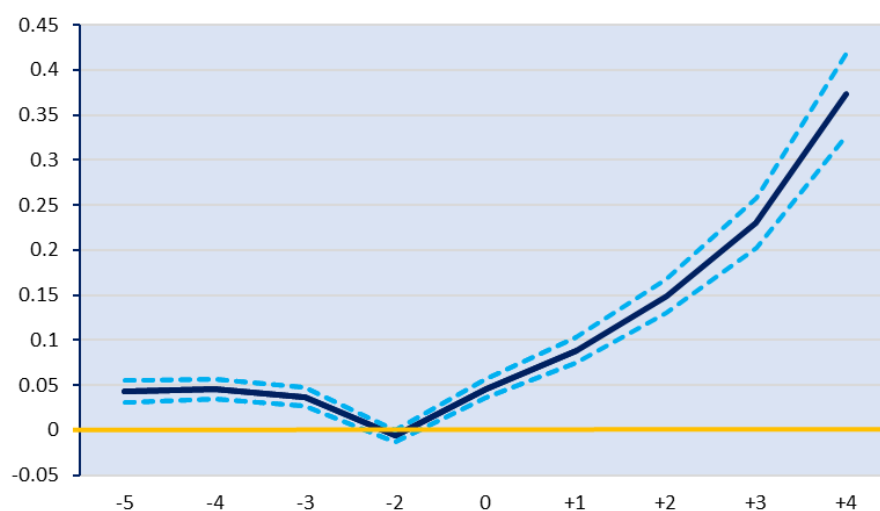


FIGURA 3-4 1° EVENT STUDY

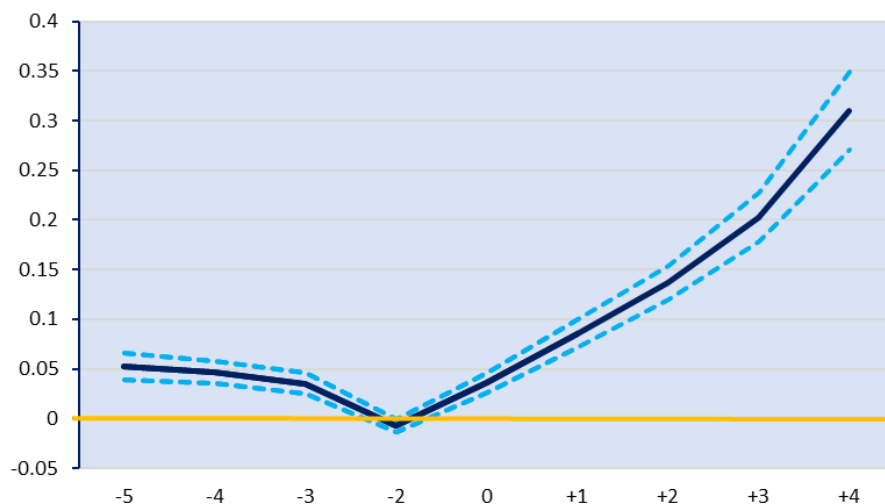


FIGURA 3-5 2° EVENT STUDY

Includendo i trend temporali condizionati alle variabili di riferimento, la dinamica ‘pre’ non perde di significatività ma è ancora statisticamente diversa da zero. Il test di esogeneità non è nuovamente rispettato e ciò manifesta un’ulteriore prova teorica della necessità di approfondire l’analisi attraverso lo studio con variabili strumentali (3.2 [Instrumental Variables](#)).

Le stime OLS corrispondenti alle regressioni riportate in [Figura 3-6](#) e [Figura 3-7](#), descritte in [Tabella 3-10](#), erano già positive e di valore rispettivamente pari a 0,063 e 0,056. L’effetto reale può essere determinato dalla media post treatment dei coefficienti ottenuti applicando il metodo event study. Si scopre dunque che l’OLS stia sottostimando l’impatto, in quanto i valori reali rilevati graficamente risultano pari a 0,074 e 0,068. Si può concludere dunque che il metodo dei minimi quadrati ordinari sia *down word biased* e si attende dunque che l’analisi a variabili strumentali esploda tale coefficiente poiché sottostimato sia a causa dell’endogeneità sia per le meccaniche del metodo OLS. È stato quindi possibile studiare gli effetti incrementali over time che precedentemente erano nascosti. Osservando il discostamento tra l’effetto ricavato dall’event study e la stima OLS si può ipotizzare un’approssimazione del fattore di correzione, il quale identifica di quanto la stima precedente sia sbagliata, indipendentemente dall’endogeneità ma a causa della meccanica intrinseca al metodo.

Risolvendo: $\begin{cases} 0,074 = 0,063 * (1 + \lambda) \\ 0,068 = 0,056 * (1 + \lambda) \end{cases}$, si ricava λ di valore compreso tra 0,175 e 0,214.

Un divario maggiore si rileva per quanto concerne i due grafici riportati in precedenza ([Figura 3-4](#) e [Figura 3-5](#)), in quanto a causa di un gruppo di controllo avente un numero limitato di variabili (poiché non erano inclusi i trend condizionati) i problemi derivanti dalla meccanica OLS non sono alleviati e stimano coefficienti 0,066 e 0,076 ([Tabella 3-4](#)) sottostimando gravemente l’effetto reale compreso tra 0,18 e 0,21. In queste specificazioni il fattore di correzione λ assumerebbe un valore ben maggiore, pari a circa 1,7.

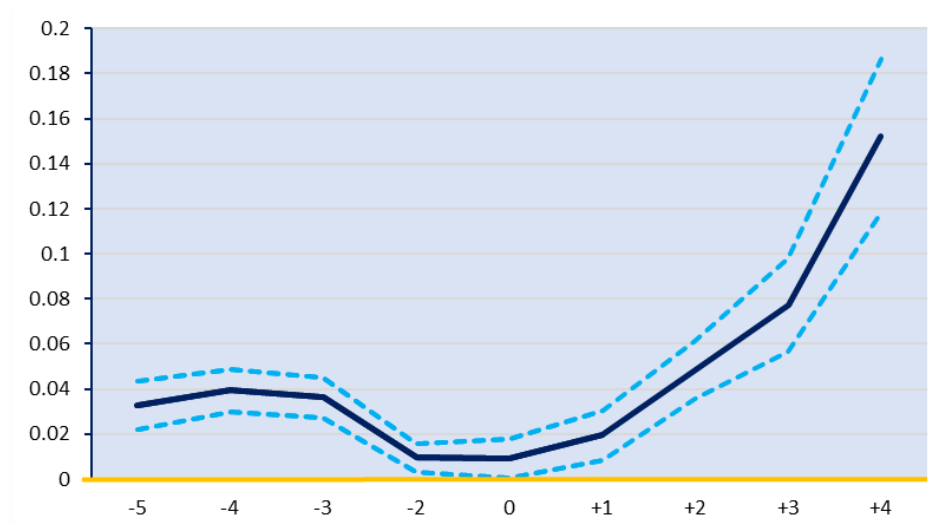


FIGURA 3-6 3° EVENT STUDY

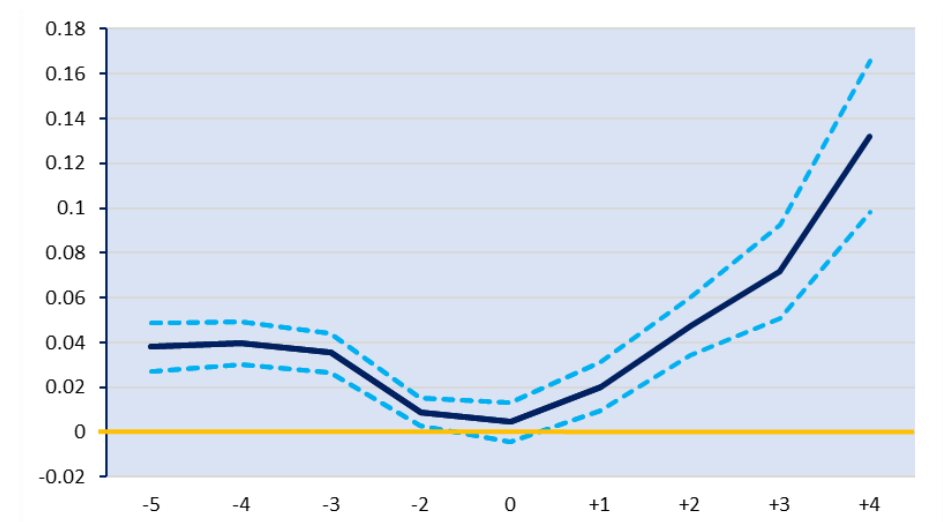


FIGURA 3-7 4° EVENT STUDY

Le curve verdi rappresentano i coefficienti e i rispettivi limiti dell'intervallo di confidenza delle dummies per ogni periodo analizzato nel caso in cui la variabile dipendente sia il numero di nuove startup innovative iscritte nel settore dei servizi. Le curve di color rosso mattone rappresentano invece le analoghe per quanto riguarda le nuove attività appartenenti agli altri settori (agricoltura/pesca, industria/artigianato, commercio e turismo).

Le regressioni corrispondenti ai primi due grafici (Figura 3-11, Figura 3-8) sono state descritte in Tabella 3-5. In Figura 3-10 e Figura 3-9 sono inclusi anche i trend temporali condizionati al grado di urbanizzazione, la presenza di università e aree bianche.

Non si registrano significative variazioni rispetto al caso generale e i periodi -5, -4 e -3 sono ancora indici di endogeneità.

Confrontando i coefficienti ricavati precedentemente attraverso la stima OLS con la media dei coefficienti post treatment registrati nell'event study, si constata un errore di maggior portata assoluta per quanto riguarda il settore dei servizi. Attraverso il metodo dei minimi

quadrati si era ottenuto un effetto compreso tra 0,052 e 0,060, il quale risulta minore della metà dell'effetto reale ricavabile graficamente pari 0,12 nella [Figura 3-11](#) e 0,14 nella [Figura 3-8](#). Il fattore di correzione ricavato nella casistica generale, per le specificazioni corrispondenti ([Figura 3-4](#), [Figura 3-5](#)), non è in questa circostanza efficiente ma sottostima l'errore. Anche per gli altri settori si verifica un rilevante scostamento relativo tra i diversi stimatori, in quanto sebbene la stima sia ancora vicina allo zero come ricavato col metodo OLS, essa nascondeva effetti incrementali over time. Lo stimatore indicava un valore compreso tra 0,0262 e 0,0217 relativamente lontano a 0,055 e 0,066 ricavabili graficamente. Per gli ultimi due grafici valgono le stesse considerazioni poiché non si riscontrano differenze significative.

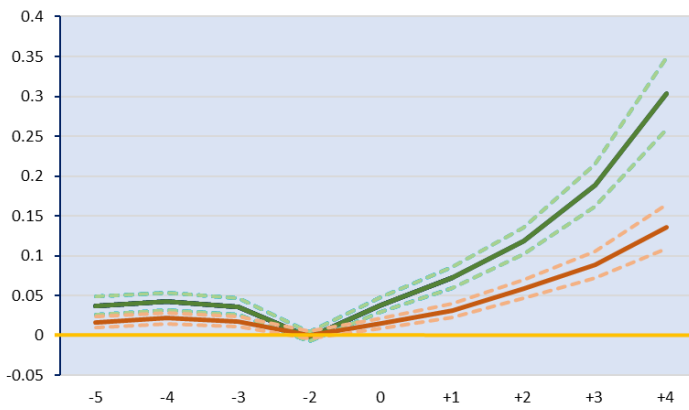


FIGURA 3-11 1° EVENT STUDY SUDDIVISO PER SETTORE

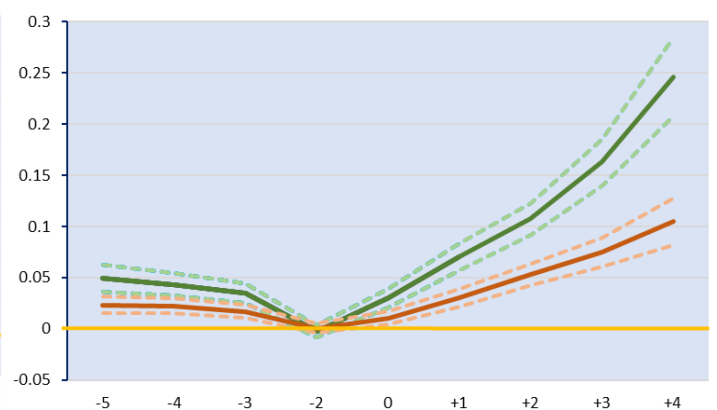


FIGURA 3-8 2° EVENT STUDY SUDDIVISO PER SETTORE

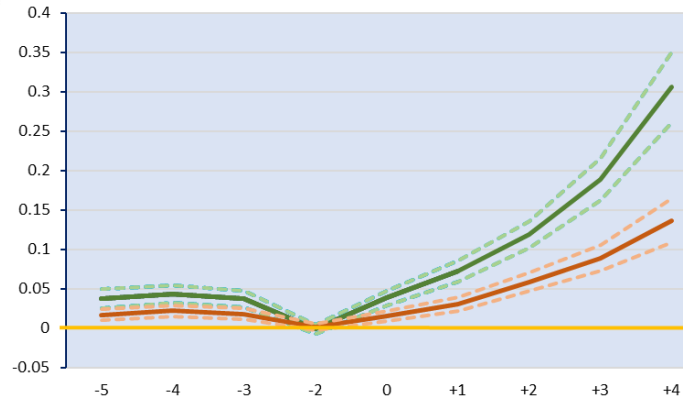


FIGURA 3-10 3° EVENT STUDY SUDDIVISO PER SETTORE

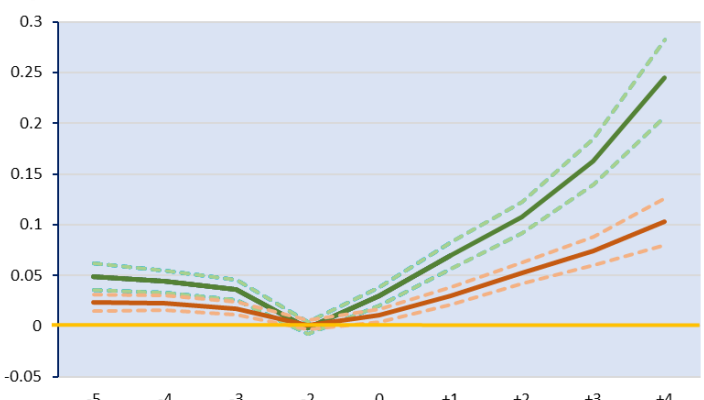


FIGURA 3-9 4° EVENT STUDY SUDDIVISO PER SETTORE

I risultati ottenuti in questa sezione rappresenteranno un utile punto di riferimento futuro durante l'analisi a variabili strumentali: è possibile ricavare che in realtà la correzione dovuta all'inclusione dello strumento sia minore di quanto possa sembrare da un primo confronto con la stima OLS corrispondente. È frequente che la stima IV sia significativamente diversa da quanto ottenuto col metodo dei minimi quadrati, ma il relativo intervallo di confidenza comprenda invece la media post treatment dei coefficienti derivanti dall'event study, e dunque non sarà possibile sostenere che le due stime siano statisticamente diverse tra loro.

3.2 Instrumental Variables

3.2.1 Giustificazione teorica

Nel modello OLS si riscontrato un problema di selezione, in quanto i comuni in cui si registra un maggior numero di startup innovative, hanno probabilità superiori di adottare la tecnologia a banda larga ultraveloce. Esiste una correlazione tra le caratteristiche proprie delle unità territoriali e il dispiegamento di tali reti e ciò implica che l'adozione della tecnologia non sia esogena al risultato di interesse.

Applicando il metodo event study (3.1.3 Event study), si è ottenuta un'ulteriore conferma della potenziale endogeneità intrinseca alla variabile d'interesse. In periodi antecedenti al dispiegamento della fibra, è stato infatti rilevato un “pre” trend che contribuiva significativamente alla crescita delle startup innovative. Pertanto, si è ritenuto necessario approfondire ulteriormente l'analisi empirica sfruttando le variabili strumentali. Sebbene il coefficiente OLS sia probabilmente distorto, i risultati precedenti sono comunque essenziali per sviluppare accuratamente il modello. Inoltre, anche in assenza di endogeneità, essi avrebbero sottostimato meccanicamente l'effetto poiché il momento in cui avviene il dispiegamento della fibra non è uniforme nelle diverse aree geografiche e gli effetti di lungo periodo impattano marginalmente nella stima OLS.

Le variabili strumentali hanno il potenziale di risolvere l'endogeneità e il relativo modello è rappresentabile dalla funzione seguente.

$$Y_{m,t} = \beta_0 + \beta_1 UBB_{m,t} + \beta_2 Z_{m,t} + u_{m,t}$$

Lo scopo è ancora stimare β_1 , ovvero l'effetto dell'introduzione della fibra sul numero di startup innovative all'interno di un'unità territoriale, includendo però le caratteristiche racchiuse nella variabile strumentale $Z_{m,t}$. Se non fosse stata presente la correlazione tra l'errore e la variabile di riferimento, la stima OLS sarebbe risultata presumibilmente la più corretta. Quando ciò non avviene, il coefficiente stimato non è centrato sul valore reale ed escludendo lo strumento, la prima assunzione dei minimi quadrati ordinari non risulterebbe rispettata. La distribuzione di u condizionata alla variabile UBB non avrebbe media nulla, dato che $UBB_{m,t}$ risulta correlato con $u_{m,t}$, e quindi il valore atteso $E(u_{m,t} | UBB_m, \dots, UBB_{mT}, \alpha_m)$ risulterebbe diverso da zero.

Analiticamente, la $cov(UBB, u)$ sarà ricavabile dall'aspettativa del prodotto tra UBB e l'errore, e tale valore, nel caso in cui la variabile d'interesse sia endogena, sarà diverso da 0.

La stima OLS di β_1 risulta pertanto distorta e inconsistente.

$$E(\beta_1^{OLS} - \beta_1) \neq 0$$

Per ottenere stimatori coerenti di β_0 e β_1 quando UBB e u sono correlati, è quindi necessario aggiungere alcune informazioni. A tale scopo, si introduce una nuova variabile osservabile (definita “strumento”), che soddisfi due condizioni fondamentali.

Innanzitutto, la *validità*. Lo strumento z non deve essere correlato con il termine di errore econometrico ($Cov(z, u) = 0$). Con questa assunzione la nuova variabile sarà definita esogena, ovvero senza alcun effetto parziale sulla variabile dipendente e non correlata con

variabili omesse. La seconda condizione riguarda la *rilevanza*: lo strumento deve essere (fortemente) correlato con la variabile endogena ($UBB_{m,t}$), $Cov(z, UBB) \neq 0$.

Dunque, si può definire z come variabile strumentale per UBB e ricapitolando, dovrà permettere di esprimere la diffusione broadband senza essere connessa con la nascita delle startup innovative.

La condizione che z sia correlata con UBB (nella popolazione) può essere verificata empiricamente, dato un campione casuale. Il modo più semplice per farlo è stimare una semplice regressione tra UBB e z .

$$UBB = \pi_0 + \pi_1 z + v$$

Allora, poiché $\pi_1 = \frac{cov(z, UBB)}{var(z)}$, l'ipotesi fondamentale vale se e solo se π_1 è diverso da zero. Quindi, si dovrà essere in grado di rifiutare l'ipotesi nulla $H_0: \pi_1 = 0$ contro l'alternativa bilaterale ($\pi_1 \neq 0$), a un livello di significatività sufficientemente piccolo (5% o 1%).

Per il panel studiato, una variabile strumentale z , non correlata a qualsiasi altro fattore non osservato (1), che influisce sul numero di nuove startup innovative, e (2) altamente correlata al dispiegamento delle reti UBB, potrebbe essere la distanza OPB. Considerazioni a riguardo saranno espresse nella sezione [3.2.2 Scelta dello strumento](#).

Si considera di nuovo il modello strutturale, supponendo ora di avere *due* variabili esogene escluse dall'equazione precedente: z_2 e z_3 .

Se le due variabili risultassero entrambe correlate con UBB , si potrebbe semplicemente usare ciascuna di queste come strumento. In quel caso però avremmo due stimatori IV, e nessuno di questi sarebbe efficiente. Dato che z_1, z_2 e z_3 non sono correlate con u_1 , anche ogni combinazione lineare delle variabili esogene risulterà non correlata con l'errore, e quindi ognuna di queste può essere una valida variabile strumentale. Per trovare la migliore IV, si identifica la combinazione lineare è più altamente correlata con UBB . È dunque ricavabile dall'equazione in forma ridotta seguente.

$$UBB = \pi_0 + \pi_1 z_1 + \pi_2 z_2 + \pi_3 z_3 + v_2,$$

$$\text{Dove } E(v_2) = 0, cov(z_1, v_2) = 0, cov(z_2, v_2) = 0, cov(z_3, v_2) = 0.$$

Allora, la migliore IV per UBB è la combinazione lineare delle z_j , che si denomina UBB^* :

$$UBB^* = \pi_0 + \pi_1 z_1 + \pi_2 z_2 + \pi_3 z_3$$

Affinché questa variabile non sia perfettamente correlata con z_1 è necessario che almeno uno di π_2 o π_3 sia diverso da zero. Questa condizione è definita *l'ipotesi chiave di identificazione*, essenziale una volta assunto che le z_j siano tutte esogene. È possibile testare l'ipotesi nulla $H_0: \pi_2 = 0$ e $\pi_3 = 0$ sfruttando una statistica F . Per proseguire il commento teorico, si divide UBB in due parti. Il primo è UBB^* ; non correlato al termine di errore. Il secondo pezzo è v_2 , possibilmente correlato con u_1 e di conseguenza implica una distorsione. Date le informazioni sulle z_1 , si può calcolare UBB^* per ogni osservazione, a condizione di conoscere i parametri della popolazione π_j . Purtroppo, ciò non accade mai nella pratica. Tuttavia, come si è visto nella sezione precedente, si può sempre stimare la forma ridotta tramite OLS. Così si regredisce UBB su z_1, z_2, z_3 ottenendo:

$$\widehat{UBB} = \hat{\pi}_0 + \hat{\pi}_1 z_1 + \hat{\pi}_2 z_2 + \hat{\pi}_3 z_3$$

A questo punto, è necessario verificare che z_2 e z_3 siano congiuntamente diverse da 0 a un livello di significatività ragionevolmente piccolo (non superiore al 5%).

Una volta ottenuto $\widehat{UBB}$, si utilizza come IV per UBB . Con strumenti multipli, tale variabile è comunemente chiamata stimatore dei minimi quadrati a due stadi (2SLS). La ragione deriva dal fatto che una volta inclusa la nuova variabile nelle specificazioni, le stime IV $\widehat{\beta}_0, \widehat{\beta}_1$ e $\widehat{\beta}_2$ sono *identiche* alle stime OLS derivanti dalla regressione di y su $\widehat{UBB}$ e z_1 .

Una corrente della letteratura economica interpreta la regressione precedente come segue. Il fitted value, $\widehat{UBB}$, è la versione stimata di UBB^* , non è correlata con l'errore. Pertanto, la 2SLS "epura" la variabile d'interesse dalla sua correlazione con u_1 prima di eseguire la regressione OLS. Possiamo dimostrarlo inserendo $UBB = UBB^* + v_2$.

$$y = \beta_0 + \beta_1 UBB^* + \beta_2 z_1 + u_1 + \beta_1 v_2$$

L'errore composito $u_1 + \beta_1 v_2$ ha media zero ed è non correlato con UBB^* e z_1 , motivo per cui la regressione di y su $\widehat{UBB}$ e z_1 è efficace e il coefficiente stimato dà l'impatto causale dell'adozione della tecnologia sulla performance aziendale.

Nella maggior parte dei casi si dovrebbe evitare di eseguire il secondo stadio manualmente, poiché gli errori standard e le statistiche di test ottenuti in questo modo non sono validi poiché il termine di errore include v_2 , mentre, gli errori standard coinvolgono solo la varianza di u_1 .

Riprendendo l'equazione $y = \beta_0 + \beta_1 UBB + \beta_2 z_1 + \dots + \beta_k z_{k-1} + u_1$, l'ipotesi necessaria affinché la 2SLS abbia le proprietà desiderate per i grandi campioni è che ogni variabile z_j sia non correlata con u_1 . Per assicurarne la coerenza è poi richiesta almeno una variabile esogena, non presente e parzialmente correlata con UBB . Infine, affinché i soliti errori standard 2SLS e le statistiche t siano asintoticamente valide, si deve assumere l'omoschedasticità: la varianza dell'errore strutturale u_1 , non può dipendere da nessuna delle variabili esogene.

L'applicazione dei metodi con variabili strumentali a sezioni trasversali indipendenti non solleva nuove difficoltà. Come per le stime OLS, si includono spesso variabili dummies corrispondenti ai periodi in modo da consentire l'inclusione degli effetti temporali aggregati. Esse sono esogene perché il passaggio del tempo non ha alcun effetto parziale sulla variabile dipendente e non risulta correlato con la variabile omessa, agendo dunque come strumenti. La stima delle variabili strumentali può quindi essere combinata con i metodi dei dati panel per stimare coerentemente i parametri in presenza di effetti non osservati ed endogeneità insite a una o più variabili esplicative variabili nel tempo.

3.2.2 Scelta dello strumento

L'individuazione della variabile strumentale non è un processo semplice poiché richiede di analizzarne approfonditamente la natura e il conseguente rispetto delle assunzioni fondamentali precedentemente descritte.

Facendo leva sulla necessità dell'infrastruttura complementare per compire il dispiegamento delle reti, è deducibile che la fibra si diffonda precedentemente nelle zone vicine ai nodi di backbone nazionali. Di conseguenza, la minor distanza è associata a livelli medi di copertura più alti. Pertanto, se tale variabile non risultasse correlata ad altre variabili omesse, rispetterebbe sia la rilevanza che la validità.

La statistica t sulla distanza OPB , nella regressione in cui la variabile dipendente è la copertura con banda larga ultraveloce, risulta pari -17,43 (Figura 3-12) e significativamente diversa da 0 a livello dell'1%.

Questa regressione lineare rappresenta solo un primo controllo; uno studio più esaustivo della rilevanza sarà

```
. regress ubb OPB_distance, r
```

Linear regression

Number of obs = 55111
F(1, 55109) = 303.65
Prob > F = 0.0000
R-squared = 0.0063
Root MSE = .44474

	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]
OPB_distance	-.0013998	.0000803	-17.43	0.000	-.0015572 -.0012423
_cons	.3353161	.0040715	82.36	0.000	.3273359 .3432962

FIGURA 3-12 REGRESSIONE UBB SU OPB DISTANCE

presentato nella sezione 3.2.3 *Stima empirica* e lo strumento più accurato sarà identificato con l'iterazione tra una dummy che assume valore 1 dal 2015 in poi (dopo l'introduzione delle infrastrutture UBB) e la distanza tra ogni comune e il nodo OPB più vicino. L'effetto di tale variabile sul dispiegamento della banda larga ultraveloce risulterà significativo e compreso tra lo 0.1% e lo 0.2%; come evidenziato successivamente nella Tabella 3-14.

Le ipotesi fondamentali sono necessarie per stimare coerentemente i parametri, in particolare nell'identificazione di β_1 . In questo contesto, ricavare un parametro significa scrivere il coefficiente della variabile principale di interesse in termini di momenti della popolazione, stimabili utilizzando un campione di dati. Successivamente, β_1 può essere individuato in termini di covarianze tra lo strumento e la variabile dipendente.

$$\text{Cov}(z, y) = \beta_1 \text{Cov}(z, UBB) + \text{Cov}(z, u)$$

Ora, sotto le due ipotesi: $\text{Cov}(z, u) = 0$ e $\text{Cov}(z, UBB) \neq 0$, è possibile risolvere per β_1 come

$$\beta_1 = \frac{\text{Cov}(z, y)}{\text{Cov}(z, UBB)}$$

È interessante notare che quest'algebra fallisca se lo strumento e la variabile d'interesse non fossero correlate, cioè se $\text{Cov}(z, UBB) = 0$. L'equazione risultante mostra che β_1 è identificabile dalla covarianza della popolazione tra z e y divisa per la covarianza della popolazione tra z e UBB . Dato un campione casuale, si stimano le quantità della popolazione tramite analoghi del campione. Dopo aver annullato le grandezze del campione al numeratore e al denominatore, si ottiene lo stimatore delle variabili strumentali di β_1 :

$$\hat{\beta}_1 = \frac{\sum_1^n (z_i - \bar{z})(y_i - \bar{y})}{\sum_1^n (z_i - \bar{z})(UBB_i - \bar{UBB})}$$

Lo stimatore IV di β_0 è semplicemente uguale a $\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{UBB}$, il quale somiglia alla stima dell'intercetta OLS, tranne che lo stimatore della pendenza (β_1), è in questo caso lo stimatore IV. Chiaramente quando $z = UBB$ si ottiene lo stimatore OLS di β_1 . In altre parole, quando la variabile dipendente è esogena, può essere usata proprio come IV, e il relativo stimatore IV sarà allora identico allo stimatore OLS.

Una caratteristica particolare è che, quando UBB e u sono correlati, così che la stima delle variabili strumentali risulta effettivamente necessaria, essa non sarà essenzialmente mai imparziale. Lo stimatore IV può avere una distorsione sostanziale nei piccoli campioni, e per questo motivo si utilizza quasi esclusivamente con campioni ad elevata numerosità (55111 nel caso studiato), in cui avrà distribuzione approssimativamente normale.

Per eseguire l'inferenza su β_1 , è necessario identificare un errore standard, utilizzabile per calcolare le statistiche t e gli intervalli di confidenza. Tendenzialmente si impone l'ipotesi

di omoschedasticità, proprio come nel caso di OLS. Tale assunzione è adesso dichiarata condizionata alla variabile strumentale z , e non più alla variabile esplicativa endogena UBB . Insieme alle precedenti ipotesi si aggiunge pertanto:

$$E(u^2|z) = \sigma^2 = Var(u)$$

Si può poi dimostrare che, se sono rispettate le ipotesi fondamentali e vale la precedente uguaglianza, allora la varianza asintotica di β_1 è uguale a

$$\frac{\sigma^2}{n\sigma_{UBB}^2\rho_{UBB,z}^2}$$

Come per lo stimatore OLS, la varianza asintotica²³ dello stimatore IV diminuisce fino a zero al ritmo di $\frac{1}{n}$, dove n è la dimensione del campione.

L'equazione precedente è interessante perché tutte le quantità incluse sono stimabili in modo coerente dato un campione casuale. Per valutare σ_{UBB}^2 , si calcola semplicemente la varianza campionaria di $UBB_{m,t}$; per quanto riguarda $\rho_{UBB,z}^2$, è possibile eseguire la regressione di $UBB_{m,t}$ su $z_{m,t}$. Infine, utilizzando i residui IV, si riesce a ottenere σ^2 .

$$\hat{u}_{m,t} = y_{m,t} - \hat{\beta}_0 - \hat{\beta}_1 x_{m,t}$$

Uno stimatore coerente di σ^2 è dunque

$$\sigma^2 = \frac{1}{n-2} \sum_{m=1}^n \sum_{t=2013}^{2019} \hat{u}_{m,t}^2,$$

dove è comune includere la correzione dei gradi di libertà, nonostante l'effetto sia trascurabile al crescere della dimensione del campione.

L'errore standard asintotico di $\hat{\beta}_1$ è la radice quadrata della varianza asintotica stimata, che è data da $\frac{\sigma^2}{SST_{UBB} \cdot R_{UBB,z}}$, in cui SST_{UBB} rappresenta la somma totale dei quadrati di $UBB_{m,t}$ e $R_{UBB,z}$ la forza della relazione lineare tra la variabile indipendente e lo strumento nel campione.

Una seconda ragione per cui è interessante si trova nella possibilità di confrontare le varianze asintotiche tra gli stimatori IV e OLS, quando UBB e u non sono correlati. Sotto le ipotesi di Gauss-Markov, la varianza dello stimatore OLS è $\frac{\sigma^2}{SST_{UBB}}$, mentre la formula comparabile per lo stimatore IV è $\frac{\sigma^2}{SST_{UBB} \cdot R_{UBB,z}}$. Poiché $R_{UBB,z}$ è sempre inferiore a 1, la varianza IV è sempre maggiore della varianza OLS, quando OLS è valido. Se UBB e z sono solo leggermente correlati, $R_{UBB,z}$ può essere piccolo, e questo può tradursi in una varianza di campionamento molto grande per lo stimatore IV. Più altamente correlati sono z e UBB , più $R_{UBB,z}$ è vicino a 1, e di conseguenza più piccola sarà la varianza dello stimatore. Nel caso in cui $z = UBB$, $R_{UBB,z} = 1$, e si ottiene la varianza OLS, come previsto.

Per quanto riguarda l'*R-squared*, tendenzialmente si utilizza la formula standard: $R^2 = 1 - \frac{SSR}{SST}$, dove SST è la somma totale dei quadrati di y . Tale misura della stima IV può quindi essere negativa perché è possibile che SSR sia più grande di SST . Tuttavia, non si ritiene tale grandezza molto utile, poiché quando la variabile endogena e u sono correlati, non è

²³ Nella formula sopra riportata, σ_{UBB}^2 rappresenta la varianza della popolazione di UBB e σ^2 della popolazione di u , mentre $\rho_{UBB,z}^2$ è il quadrato della correlazione della popolazione tra UBB e z . Quest'ultimo termine indica quanto altamente correlati sono UBB e z .

possibile scomporre la varianza di y in $\beta_1^2 Var(UBB) + Var(u)$, e quindi l'*R-squared* non possiede un'interpretazione naturale. Se l'obiettivo fosse quello di produrre il più grande *R-squared*, sarebbe conveniente utilizzare sempre OLS poiché i metodi IV sono destinati a fornire stime migliori dell'effetto della banda larga ultraveloce sulla variabile dipendente solo quando UBB e u sono correlati; mentre la bontà dell'adattamento non rappresenta un fattore.

Inizialmente si è valutata la distanza OLT come possibile candidata a variabile strumentale. Tuttavia, nonostante la forte rilevanza generata dalla correlazione tra la distanza OLT e la diffusione dell'UBB, l'operatore di telecomunicazioni sostiene effettivamente costi minori dispiegando la fibra nei comuni vicini a una sede OLT e ciò comprometterebbe la validità di questo potenziale strumento.

Pertanto, poiché tale ipotesi non è trascurabile, come nel modello precedente, si include la distanza OLT tra le variabili di controllo, e si identifica come variabile strumentale la distanza OPB, strumento ritenuto esogeno derivante dall'architettura della dorsale Kaleidon.

Sfruttando la specificazione creata per la stima OLS, si mira a ottenere l'analisi a variabili strumentali con equazioni confrontabili alle precedenti. La variabile dipendente è ancora *log(startup)*, mentre le covariate principali sono time dummies interagite con la distanza al nodo OPB più vicino. Inoltre, sono inclusi i trend temporali condizionati a controlli costanti nel tempo ed effetti fissi a livello di comune e anno-regione.

In tutte le specificazioni presentate successivamente, la distanza OLT risulta fortemente significativa ed endogena. Tale evidenza conferma che in realtà la distanza OLT non sarebbe stata uno strumento valido, e implica che i comuni lontani da un OLT siano caratterizzati da numerosità inferiore di startup innovative.

Le sedi OLT tendono a sua volta ad essere installate prima nei CO più vicini ai nodi OPB, poiché la vicinanza ne semplifica la diffusione. Ciò comporta che le aree, la cui distanza al nodo OPB è minore, manifestino maggiori probabilità di ricevere la fibra. Si attende dunque una correlazione negativa, ovvero che nel primo stadio della regressione IV, lo strumento sia significativo (rilevanza) e presenti coefficiente di segno negativo ([Tabella 3-14](#)).

Adesso si analizza in ordine la direzione (presumibilmente) derivante dalla correzione con variabili strumentale e conseguentemente se ne discute l'esogeneità.

La restrizione di esclusione sarà violata se ci fossero tendenze sottostanti, correlate sia con lo strumento identificato sia con i risultati di interesse.

Nella sezione successiva si verifica empiricamente la validità della variabile, focalizzandoci, in un primo momento, sul potenziale impatto delle variabili omesse sulla direzione della correzione IV.

Alcuni shock potrebbero impattare significativamente sulla diffusione della fibra e allo stesso tempo sul numero di startup innovative. Ad esempio, nel caso in cui all'interno di un comune, a seguito di finanziamenti statali, fosse aperto un incubatore tecnologico, ciò comporterebbe un sostanziale incremento del numero di startup e contemporaneamente aumenterebbe la domanda di banda. Un effetto simile potrebbe giustificare una sovrastima di OLS rispetto a IV.

Altre variabili omesse potrebbero invece avere un effetto opposto, in quanto è possibile che il dispiegamento delle reti in certe zone potrebbe non derivare da dinamiche interne al mercato. Se per combattere il divario economico caratterizzante la penisola italiana, il governo concentrasse gli investimenti pubblici al sud Italia o in zone rurali, aumenterebbe

la probabilità di ricevere la fibra in aree caratterizzate da scarsa presenza di startup. Però nel caso in cui questi piani nazionali fossero stati sanciti prima dell'introduzione della banda larga ultraveloce, ciò risulterebbe time invariant e l'effetto fisso inerente al comune ne catturerebbe l'impatto potenziale. Driver tangibili della forte correlazione tra lo strumento e le variabili omesse potrebbero essere i bandi di intervento nelle aree bianche emessi dal 2015 al 2019.

Un'altra circostanza plausibile è il caso in cui l'introduzione della fibra derivasse da una volontà politica, derivante dalla forte e attuale propaganda sul digitale. Finanziare le nuove tecnologie potrebbe inoltre diminuire i fondi dedicati alle startup, nel caso in cui le risorse fossero limitate e sostituite. Anche quest'ultimo shock risulterebbe correlato positivamente alla diffusione di banda e negativamente alle startup innovative. Pertanto, si attende che i bias presenti nella stima OLS tendano a sottostimare il reale effetto della variabile d'interesse.

Per approcciare le IV si eseguono quattro regressioni (Tabella 3-11) includendo lo strumento e sfruttando il modello OLS strutturato precedentemente. Si rileva un impatto della fibra molto alto, il quale ci porta a dubitare riguardo l'esogeneità dello strumento: potrebbe catturare qualcosa che ne enfatizzi l'effetto. Tuttavia, con un buon grado di fiducia, si ritiene rispettata l'esogeneità nelle stime IV in cui sono incluse le interazioni (colonne [3] e [4], condizionate al GDU, aree bianche e la presenza di università) poiché tali variabili riescono a controllare per trend differenziali basati su variabili time invariant. Inoltre, va ricordato che anche le evidenze ottenute dall'event study, indirizzassero verso coefficienti di portata significativamente maggiore rispetto quanto ottenuto nella stima OLS. Nonostante ciò, per maggior correttezza si ritiene comunque utile una verifica riguardante la bontà della variabile strumentale.

Probabilmente il problema dal punto di vista econometrico potrebbe derivare dalla concentrazione delle startup nelle grandi città, la cui la distanza dal nodo nazionale più vicino è presumibilmente pari a 0 o comunque molto piccola. Nonostante i nodi OPB siano stati posizionati nel 2012 (prima del periodo analizzato), una certa correlazione tra l'ubicazione e le dinamiche delle startup innovative potrebbe ancora esistere, dato che le aree scarsamente popolate sono caratterizzate sia da un basso incremento di attività sia da un'elevata distanza dal nodo OPB più vicino.

Lo strumento potrebbe quindi catturare ulteriori effetti che impattano significativamente sulla variabile dipendente. Per approfondire questo aspetto si eseguono quattro regressioni omettendo i comuni capoluoghi di provincia (Tabella 3-12), minimizzando dunque la possibilità che lo strumento assorba endogeneità derivanti dall'essere una grande città. Rispetto alla situazione precedente, si riscontra un'effettiva diminuzione dell'effetto della banda larga compresa tra il 30 e 39%. Tale evidenza implica ulteriori preoccupazioni riguardo la validità della variabile strumentale e dunque ne giustifica un ulteriore approfondimento (Figura 3-13, Figura 3-14, Figura 3-15, Figura 3-16, Figura 3-17). Si riscontra che il dispiegamento della fibra in un comune, non capoluogo di provincia, implica un incremento di circa il 22% delle startup innovative nascenti. L'FTTH subisce una perdita di significatività prima parziale ([2]) e poi totale includendo i trend temporali ([3]). Escludendo le maggiori città italiane, era comunque prevedibile tale effetto poiché già col modello OLS, i maggiori impatti per la fibra erano stati registrati nelle aree ad alta urbanizzazione (Tabella 3-6), non bianche (Tabella 3-7), aventi almeno un polo universitario (Tabella 3-8) e caratterizzate da un elevato reddito locale (Tabella 3-9); e un'elevata percentuale dei comuni aventi tali caratteristiche saranno presumibilmente capoluoghi di provincia.

È possibile che le direzioni delle distorsioni asintotiche siano diverse per IV e OLS. Per esempio, supponendo che $\text{Corr}(x, u) < 0$, $\text{Corr}(z, x) > 0$, e $\text{Corr}(z, u) > 0$, allora lo stimatore IV avrà un bias verso l'alto, mentre lo stimatore OLS verso il basso. La distorsione asintotica dello stimatore IV è inferiore a quella di OLS solo se $\frac{\text{Corr}(z, u)}{\text{Corr}(z, x)} < \text{Corr}(x, u)$. Se la $\text{corr}(z, x)$ è piccola, allora una correlazione apparentemente piccola tra z e u può essere ingigantita e rendere IV peggiore di OLS.

Dalla [Tabella 3-11](#) si evince inoltre che tutte le regressioni al primo stadio siano forti con una F-statistica elevata, il che significa che il bias da strumento debole non è un problema. L'F-test cattura infatti la rilevanza dello strumento, ovvero quanto è correlato con la variabile endogena.

TABELLA 3-11 MODELLO IV INIZIALE

	(1) log(startup)	(2) log(startup)	(3) log(startup)	(4) log(startup)
UBB	0.378*** (0.0382)	0.353*** (0.0359)	0.328*** (0.0675)	0.315*** (0.0663)
OLT distance	0.000685*** (0.0000989)	0.000553*** (0.0000874)	0.000588*** (0.000114)	0.000531*** (0.000110)
Gdu	0 (.)	0 (.)	0 (.)	0 (.)
università	0 (.)	0 (.)	0 (.)	0 (.)
aree bianche	0 (.)	0 (.)	0 (.)	0 (.)
patents	-0.00002 (0.000001)	-0.0000025** (0.000001)	-0.0000026** (0.000001)	-0.0000027** (0.000001)
FTTH		0.267*** (0.0315)		0.138*** (0.0243)
Year FE				
Municipality FE	X	X	X	X
Year-Region FE	X	X	X	X
Trend temporali condizionati a variabili time invariant			X	X
First stage F-test	481.47	477.41	120.95	199.28
N	54980	54980	54980	54980
R ²	-0.172	-0.121	0.002	0.022

Standard errors in parentheses

Source: auto.dta

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

TABELLA 3-12 MODELLO IV INIZIALE CON CAPOLUOGHI DI PROVINCIA OMESSI

	(1)	(2)	(3)	(4)
	log(startup)	log(startup)	log(startup)	log(startup)
UBB	0.2303*** (0.0169)	0.225*** (0.0170)	0.223*** (0.0386)	0.219*** (0.0389)
OLT distance	0.000317*** (0.0000479)	0.000295*** (0.0000471)	0.000319*** (0.00006)	0.000307*** (0.000061)
Gdu	0 (.)	0 (.)	0 (.)	0 (.)
università	0 (.)	0 (.)	0 (.)	0 (.)
aree bianche	0 (.)	0 (.)	0 (.)	0 (.)
patents	-0.00002 (0.000001)	-0.0000045 (0.000007)	-0.0000028 (0.000001)	-0.0000017 (0.000001)
FTTH		0.067** (0.0197)		0.0361 (0.0190)
Year FE				
Municipality FE	X	X	X	X
Year-Region FE	X	X	X	X
Trend temporali condizionati a variabili time invariant			X	X
First stage F-test	460.021	451.21	118.15	115.67
N	54224	54224	54224	54224
R ²	-0.089	-0.081	-0.055	-0.051

Standard errors in parentheses

Source: auto.dta

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

3.2.3 Stima empirica

Dopo aver riscontrato una possibilità endogeneità della variabile strumentale, si va ad analizzare più approfonditamente le correlazioni alla base di essa. La speranza è riuscire a dimostrare che lo strumento sia correlato alla nascita delle startup innovative, soltanto indirettamente, attraverso il dispiegamento della banda larga ultraveloce.

Si verifica dunque se prima dell'inizio della diffusione della fibra (avvenuto nel 2015), è comunque presente una correlazione significativa tra la variabile dipendente e la distanza tra ogni comune e il nodo OPB più vicino. A tale fine, si effettuano cinque regressioni a livello generale in cui le covariate principali sono time dummies interagite con la distanza OPB, includendo sia effetti fissi a livello di comune che anno-regione. A causa della

collinearità tra le variabili, l'arco di tempo analizzato sarà dal 2014 al 2019 escludendo così la dummy riferita al 2013. Una volta ottenuti i risultati, essi saranno rappresentati graficamente, congiuntamente ai relativi intervalli di confidenza. Si vorrebbe ottenere che l'interazione della distanza OPB con la dummy 2014 non sia significativamente diversa da zero, in modo da poter concludere che prima della fibra non ci fosse nessuna correlazione. Se invece già nel 2014 si verificasse una relazione, ad esempio se le grandi città avessero un maggior numero di startup indipendentemente dalle reti UBB, ciò implicherebbe di riscontrare che lo strumento cattura queste caratteristiche che causano la sovrastima dell'effetto. Di conseguenza, non sarebbe rispettata la condizione fondamentale della validità. Nel caso in cui il test non funzionasse in nessuna specificazione, si dovrebbe pertanto concludere che la stima IV non risolve il problema di selezione inerente alla OLS, e anzi peggiora ulteriormente la qualità dei risultati ottenuti.

Le correlazioni, dopo l'introduzione della banda larga, non minacciano invece la bontà dello strumento poiché le startup dipendono dalla fibra che a sua volta è fortemente correlata con lo strumento (rilevanza).

Inizialmente, si costruisce un modello in cui le variabili indipendenti sono rappresentate esclusivamente dalle interazioni tra la distanza OPB e le time dummies ($D_{m,t}$).

$$Y_{m,t} = \lambda_0 + \lambda_1 D_{m,2014} + \lambda_2 D_{m,2015} + \lambda_3 D_{m,2016} + \lambda_4 D_{m,2017} + \lambda_5 D_{m,2018} + \lambda_6 D_{m,2019} + \gamma_m + u_{m,t}$$

Come per le stime precedenti la variabile dipendente $Y_{m,t}$ è pari a $\log(startup + 1)$. L'effetto fisso inerente al comune (γ_m) e a livello di anno-regione ($X_{r,t}$) saranno sempre inclusi, anche nelle regressioni successive.

Nella [Figura 3-13](#) si riportano i risultati ottenuti con la prima specificazione. Sebbene dal 2015 la distanza OPB abbia un impatto altamente significativo (p-value < 1%) e negativo sulla nascita di startup innovative, si registra un effetto non statisticamente diverso da zero nel 2014. Come si può vedere dal grafico, l'intervallo di confidenza al 95% relativamente al primo anno analizzato non include lo zero. Il p-value relativo all'interazione specifica è pari al 39%. Senza ulteriori controlli si temeva una correlazione della variabile dipendente con il nodo OPB anche prima del 2015. Tuttavia, l'evidenza ricavata rappresenta un'ulteriore conferma sulla validità dello strumento.

Inoltre, è possibile trarre importanti informazioni riguardo la direzione della stima a variabili strumentali. Se dal momento in cui si diffonde la banda larga ultraveloce, si verificasse una correlazione negativa dello strumento sia con la variabile dipendente che con la fibra, allora nell'IV si registrerebbe un impatto positivo del dispiegamento delle reti sulle iscrizioni di startup innovative.

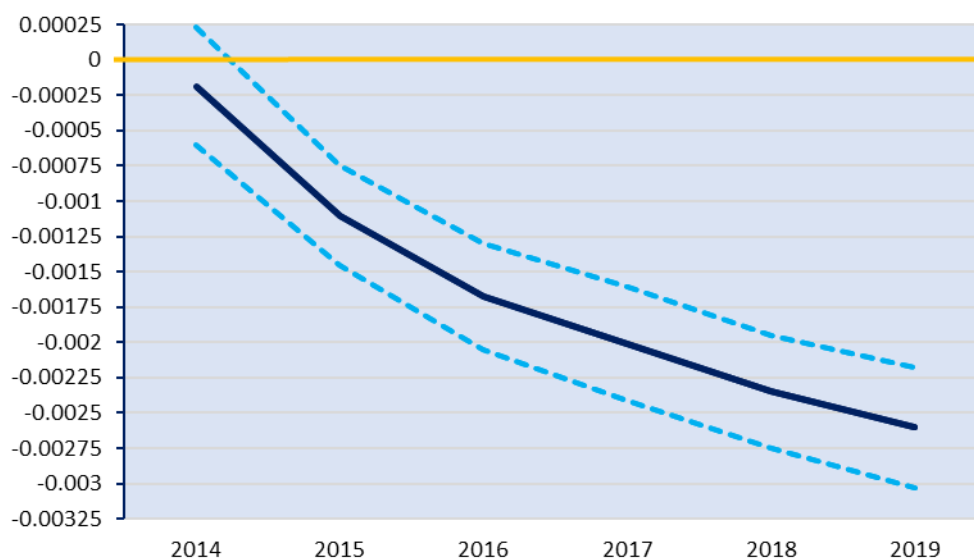


FIGURA 3-13 1° TEST ESOGENEITÀ

Per irrobustire l'analisi si aggiungono, una ad una, le altre iterazioni precedentemente generate, riguardanti il grado di urbanizzazione, le aree bianche e la presenza di università. A causa della collinearità tra le variabili, anche per queste time dummies, le informazioni relative al 2013 sono omesse. Aggiungendo incrementalmente i controlli, si attende un accostamento del coefficiente d'interesse verso lo zero e dunque una conseguente perdita di significatività. Infine, si individueranno i set di controlli che garantiranno una forte validità dello strumento e si includeranno nella stima IV.

Nella [Figura 3-14](#) si aggiungono le covariate relative al grado di urbanizzazione. Il coefficiente relativo al 2014 diminuisce da -0.0006 a -0.00037 e il relativo p-value aumenta di 30,8 punti percentuali, raggiungendo il 70,8%. L'anno successivo la distanza OPB risulta statisticamente diversa da zero con un livello di significatività maggiore rispetto al caso precedente (p-value pari a 3,9%). Dal 2016 in poi invece si verifica una forte e negativa correlazione tra lo strumento e la variabile dipendente.

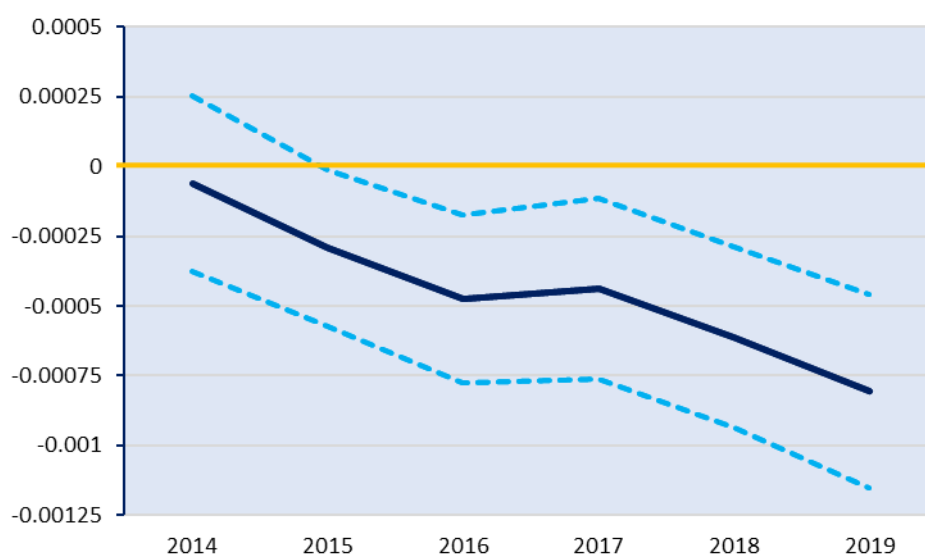


FIGURA 3-14 2° TEST ESOGENEITÀ

Integrando tra i controlli anche le time dummies riguardanti la presenza di università, l'impatto della distanza OPB subisce un lieve incremento (Figura 3-15). Sebbene la stima attesa del coefficiente specifico del 2014 sia di valore identico a quanto registrato nella Figura 3-14, la probabilità di estrarre una statistica sfavorevole all'ipotesi nulla almeno quanto il valore effettivamente calcolato con i dati, supponendo che l'ipotesi nulla sia corretta, risulta in questo caso pari al 57,8%. Nel 2015 lo strumento risulta statisticamente diverso da 0 a livello dell'1%.

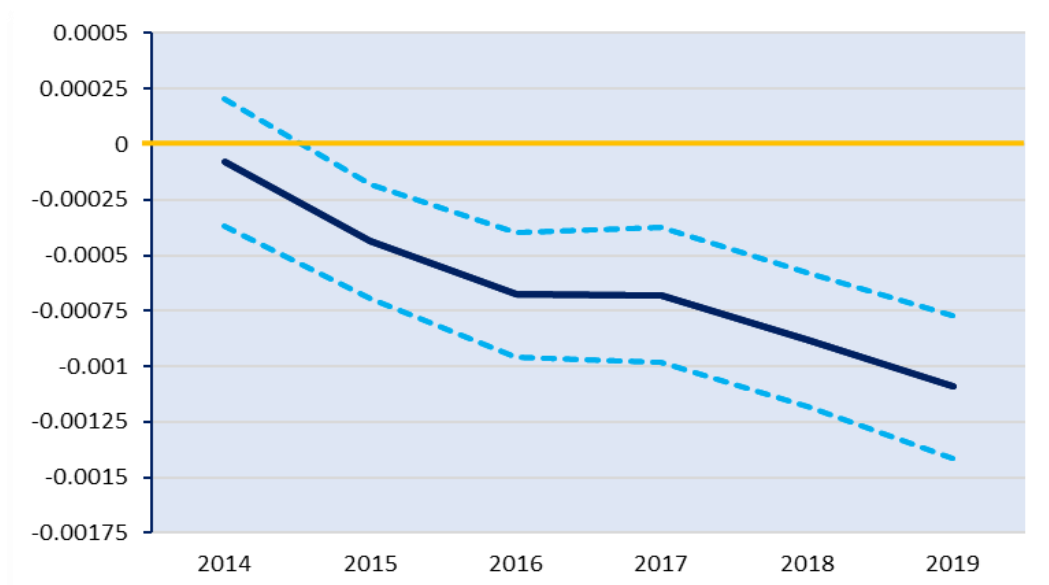


FIGURA 3-15 3° TEST ESOGENEITÀ

Includendo anche le iterazioni che caratterizzano le aree bianche (Figura 3-16), non si verificano rilevanti variazioni rispetto alla specificazione appena descritta.

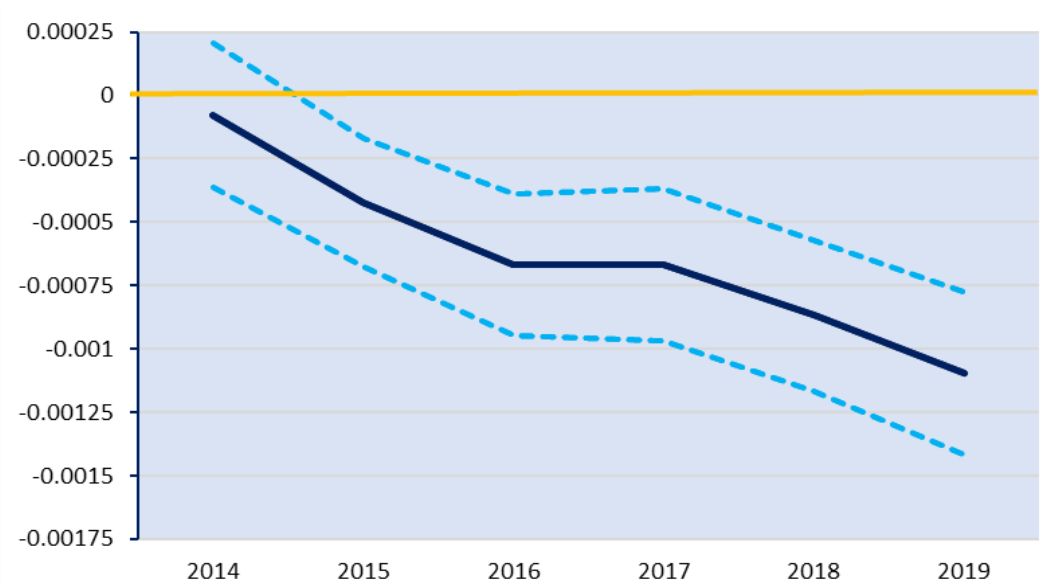


FIGURA 3-16 4° TEST ESOGENEITÀ

Infine, nella [Figura 3-17](#), introduciamo sette nuove variabili generate dall'iterazione tra popolazione²⁴ specifica di ogni comune e le time dummies corrispondenti all'arco temporale 2013-2019. In quest'ultima specificazione si verifica un'evidente perdita di significatività delle iterazioni relative allo strumento nella maggior parte degli anni analizzati.

Nel 2014, il coefficiente della distanza OPB è pari a -0.00013 e il relativo p-value raggiunge il 98,9%. Per mezzo di tale informazione, è pertanto possibile rispondere positivamente al test di esogeneità, poiché non essendo la distanza OPB con la dummy 2014 significativamente diversa da zero, si può concludere che prima della fibra non sia presente nessuna correlazione. Di conseguenza la validità dello strumento è rispettata.

Anche per gli anni successivi si registra un forte diminuzione di significatività per la distanza OPB, la quale risulta significativamente diversa da 0 a livello del 5% solamente nel 2019. A scopo descrittivo si riportano i livelli di significatività marginali nel periodo di diffusione della fibra: 26,4%, 5,5%, 79,7%, 14,9% e 0,5%.

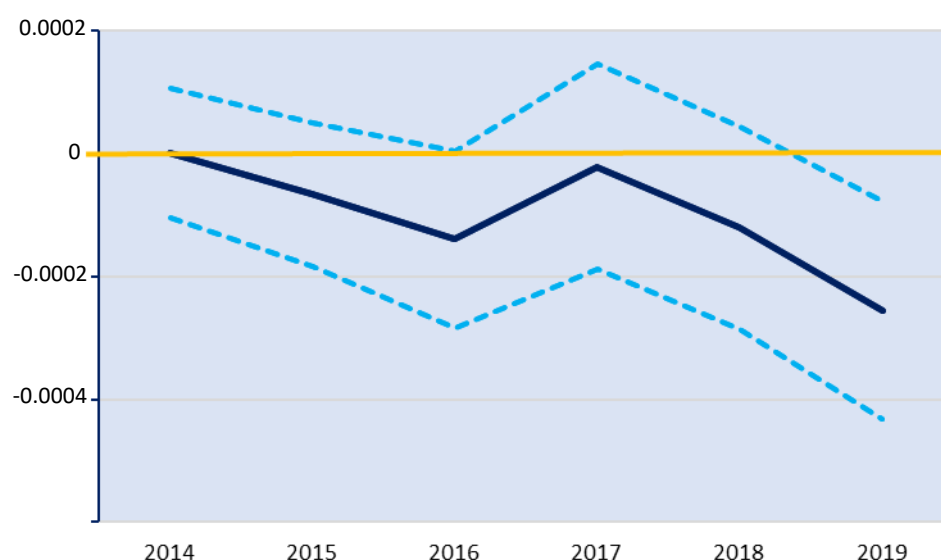


FIGURA 3-17 5° TEST ESOGENEITÀ

Dopo aver validato lo strumento, si registrano i risultati finali della stima IV nella [Tabella 3-13](#).

La variabile strumentale *post2015 OPB* è data dall'iterazione della distanza, tra ogni comune e il nodo OPB più vicino, con un dummy che assume valore 1 dal 2015 in poi. Sono inclusi inoltre gli effetti fissi a livello comunale e anno regione. I risultati sono controllati ulteriormente in modo non parametrico attraverso le tendenze sottostanti, interagendo le covariate di base con le dummy dell'anno di riferimento.

Nelle colonne [3] e [4] si sono aggiunti i trend temporali condizionati alla popolazione di ciascun comune. La variabile d'interesse UBB risulta perennemente significativa in tutte le regressioni. Nelle colonne [1] e [2], l'effetto risulta compreso tra il 31,5% e il 32,8%.

²⁴ Le informazioni relative alla popolazione residente a livello comunale sono state prelevate dal data warehouse del 15° Censimento generale, ad opera dell'ISTAT. La data di riferimento è il 9 ottobre 2011.

Includendo i controlli per la popolazione, l'implicazione derivante dall'essere una grande città viene presumibilmente catturata da tali variabili, comportando una diminuzione dell'impatto della banda larga di circa 20 punti percentuali. La diffusione della fibra all'interno di un comune genera pertanto un incremento del 13,6% delle iscrizioni annuali di startup innovative al Registro delle Imprese.

Confrontando le prime due colonne della tabella seguente con le stime OLS, in cui si è utilizzata la stessa specificazione (Tabella 3-10), si registrano risultati più di 5 volte superiori a quanto ricavato col modello dei minimi quadrati ordinari. Ciò suggerisce che la stima OLS sia troppo bassa ed è coerente con la distorsione presentata precedentemente. L'errore standard della stima IV è circa ventidue volte più grande dell'errore standard OLS.

Nella colonna [1], l'intervallo di confidenza del 95% per β_1 col metodo OLS è quindi molto più stretto di quello usando IV; infatti, l'intervallo IV comprende valori tra 0,196 e 0,460. Non contiene pertanto la stima OLS corrispondente (0,063) e ciò ci permette di affermare a un livello di significatività del 5% che la differenza tra le due sia statisticamente significativa. Confrontando la stima ottenuta con 0,074, ovvero la media post treatment dei coefficienti derivanti dal metodo event study (Figura 3-6), si può sostenere che le evidenze ottenute applicando i due metodi siano statisticamente diverse con un livello di fiducia del 5%.

Nella regressione [2], l'intervallo di confidenza IV assume valori da un minimo di 0,185 a un massimo di 0,445. Anche tale stima permette di rifiutare i corrispondenti ricavati precedentemente poiché i coefficienti pari a 0,056 (OLS) e 0,068 (event study, Figura 3-7) non risultano compresi.

La variabile relativa alla distanza OLT è sempre significativa (p-value < 5%), sebbene i valori dei coefficienti siano di ordini di grandezza inferiori rispetto all'impatto del dispiegamento della banda. Un incremento unitario della distanza corrisponde infatti a un aumento delle startup innovative nascenti compreso tra lo 0,0129% e lo 0,058%.

Contrariamente a quanto succede nelle stime OLS, anche la variabile relativa ai brevetti a livello regionale acquista significatività nelle regressioni [1] e [2]. L'effetto negativo è tuttavia difficilmente interpretabile. Si evince che, preso in considerazione un comune, a seguito di un incremento di 1000 registrazioni all'interno della provincia di appartenenza, le iscrizioni di nuove startup all'interno di quell'unità territoriale diminuiranno di circa lo 0,265%.

Infine, si osserva la variabile relativa a FTTH. Essa risulta altamente significativa e positiva nella regressione [2]. Confrontando tali valori con la relativa specificazione OLS (colonna [2] della Tabella 3-10) si osserva che tale coefficiente era stato sovrastimato, in quanto risulta inferiore di circa il 30%. Il relativo errore standard è rimasto più o meno costante ma osservando l'intervallo di confidenza a variabili strumentali [9.03, 18.56], si nota che la stima OLS (19.7) non è compresa. Pertanto, oltre a riscontrarne la diminuzione, si può affermare che tale differenza sia statisticamente significativa.

L'ultima regressione invece riporta un'evidenza divergente con quanto riscontrato nelle specificazioni precedenti. Il raggiungimento con la fibra ad alta qualità di un nuovo comune, risulta essere correlata negativamente alla nascita di nuove startup innovative. Nello specifico, un incremento unitario genera una diminuzione dello 0,43%. Tuttavia, il vero significato della variabile FTTH è l'impatto aggiuntivo di quest'ultima variabile sulla fibra, ovvero lo scostamento rispetto all'output derivante dalla regressione in cui è inclusa esclusivamente UBB.

TABELLA 3-13 MODELLO IV - POST CHECK ESOGENEITÀ

	(1)	(2)	(3)	(4)
	log(startup)	log(startup)	log(startup)	log(startup)
UBB	0.328*** (0.0675)	0.315*** (0.0663)	0.136*** (0.0495)	0.136*** (0.0502)
OLT distance	0.000588*** (0.000114)	0.000531*** (0.000110)	0.000129* (0.00006)	0.00013*** (0.00006)
Gdu	0 (.)	0 (.)	0 (.)	0 (.)
università	0 (.)	0 (.)	0 (.)	0 (.)
aree bianche	0 (.)	0 (.)	0 (.)	0 (.)
patents	-0.0000026** (0.000001)	-0.0000027** (0.000001)	-0.0000015 (0.000007)	-0.0000014 (0.000001)
FTTH		0.138*** (0.0243)		-0.0043*** (0.020)
Year FE				
Municipality FE	X	X	X	X
Year-Region FE	X	X	X	X
gduXyr	X	X	X	X
areebiancheXyr	X	X	X	X
universitàXyr	X	X	X	X
popolazioneXyr			X	X
First stage F-test	120.956	119.280	51.022	49.771
N	54980	54980	52192	52192
R ²	0.002	0.022	0.005	0.005

Standard errors in parentheses

Source: auto.dta

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Dunque, si conclude la seconda e più esaustiva analisi generale stimando che l'effetto dell'adozione della banda larga sui comuni è positivo, statisticamente significativo e si attesta intorno al 13,6% quando consideriamo gli effetti fissi e i trend temporali condizionati alle variabili di riferimento.

Come mostrato nella [Tabella 3-14](#), il coefficiente del primo stadio è significativo e con il segno atteso, e suggerisce che spostarsi di 1 km dal nodo OPB diminuisce la probabilità di ricevere connessioni ultraveloci di un valore compreso tra lo 0,12%-0,195%. Le differenze rispetto alle specificazioni precedenti sono che adesso la variabile dipendente è la copertura UBB, mentre la variabile d'interesse è lo strumento, ovvero *postXopb*. I valori corrispondenti dell'F-test sono riportati nella [Tabella 3-13](#). Lo strumento risulta forte, ovvero la rilevanza è sempre rispettata poiché il test F restituisce sempre valori ampiamente

superiori a 20. Una spiegazione dell'impatto positivo della distanza OPB sul dispiegamento della fibra potrebbe essere la seguente. Le sedi OLT tendono a essere installate prima nei CO vicini ai nodi della dorsale Kaleidon e successivamente in aree più distanti. Esiste infatti una correlazione negativa, pari a -0,0811, tra la distanza OPB e le variabili dummy relative alla presenza della sede OLT. Al contrario le distanze minime comuni-OLT e comuni-OPB sono positivamente correlate, con un valore di 0,22. Di conseguenza, poiché la distanza OLT risulta sempre correlata negativamente alla banda larga ultraveloce, i comuni relativamente più vicini ai nodi OPB hanno maggiori probabilità di ricevere tali tecnologie.

TABELLA 3-14 VERIFICA RILEVANZA VARIABILE STRUMENTALE

	(1) UBB	(2) UBB	(3) UBB	(4) UBB
postXopb	-0.00195*** (0.000199)	-0.00194*** (0.000199)	-0.00122*** (0.000189)	-0.00120*** (0.000188)
OLT distance	-0.00140*** (0.000118)	-0.00142*** (0.000118)	-0.00110*** (0.000111)	-0.00113*** (0.000112)
gdu	0 (.)	0 (.)	0 (.)	0 (.)
università	0 (.)	0 (.)	0 (.)	0 (.)
aree bianche	0 (.)	0 (.)	0 (.)	0 (.)
patents	0.00000492** (0.00000160)	0.00000484** (0.00000161)	0.00000580*** (0.00000165)	0.00000560*** (0.00000167)
FTTH		0 (.)		0.208*** (0.0290)
_cons	0.0490** (0.0181)	0.241*** (0.0179)	0.0706*** (0.0209)	0.0390 (0.0211)
Year FE				
Municipality FE	X	X	X	X
Year-Region FE	X	X	X	X
gduXyr	X	X	X	X
areebiancheXyr	X	X	X	X
universitàXyr	X	X	X	X
popolazioneXyr			X	X
N	55111	55111	52288	52288
R ²	0.729	0.729	0.732	0.733

Standard errors in parentheses

Source: auto.dta

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Si verifica adesso la validità dello strumento anche nel caso in cui la variabile dipendente sia osservata suddividendola per settore. L'obiettivo è verificare, come effettuato in precedenza per il caso generale, che l'iterazione tra distanza OPB e la dummy relativa al 2014 non abbia un effetto significativamente diverso da zero sulle iscrizioni di nuove startup innovative operanti nei servizi e negli altri settori, in modo da dimostrare l'assenza di correlazione prima dell'introduzione della fibra.

Nelle specificazioni rappresentate in [Figura 3-18](#) e [Figura 3-19](#), le covariate principali sono time dummies interagite con la distanza OPB. Sono inoltre inclusi effetti fissi a livello di comune, anno-regione e trend differenziali basati su gdu, aree bianche e università. In [Figura 3-19](#) e [Figura 3-21](#) e, sono inoltre aggiunti i nuovi controlli generati dall'iterazione tra la popolazione specifica di ogni comune e le time dummies corrispondenti.

Nei prossimi due modelli la variabile dipendente è il numero di nuove iscrizioni di startup innovative appartenenti al settore dei servizi. Già nella prima specificazione, la distanza OPB non risulta significativamente diversa da zero (p-value=65%) nel 2014. Come visto nel caso generale, includendo la popolazione la distanza interagita alle time dummies perde di significatività sia nell'anno d'interesse, il cui p-value cresce fino al 98.9%, sia nei periodi successivi. Ottenute queste evidenze, si può concludere che lo strumento funziona bene per spiegare gli effetti all'interno di questo settore.

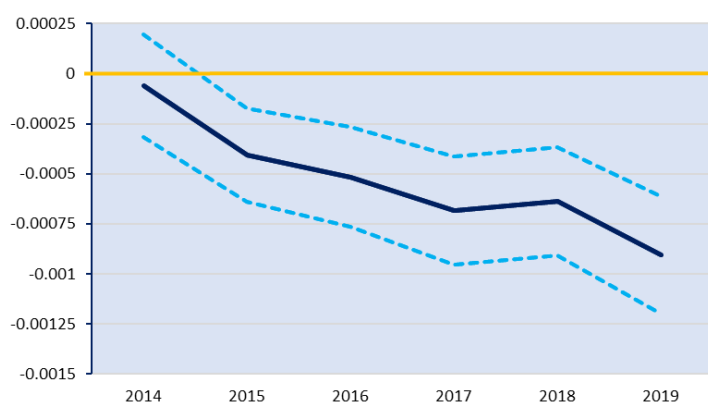


FIGURA 3-18 1° TEST ESGENEITÀ PER I SERVIZI

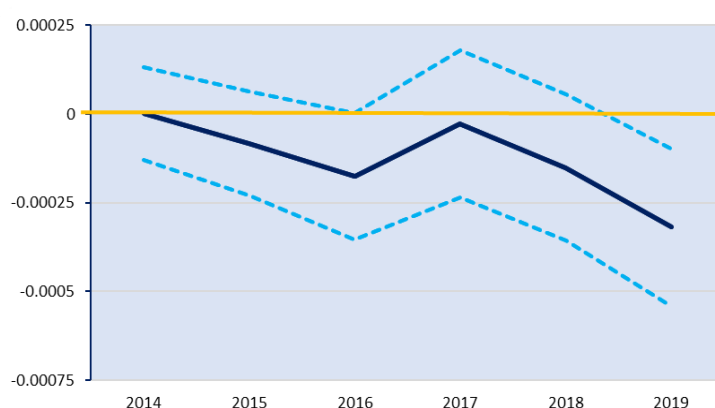


FIGURA 3-19 2° TEST ESGENEITÀ PER I SERVIZI

Per quanto riguarda gli altri settori non è possibile giungere alle medesime conclusioni. Tuttavia, nella prima specificazione ([Figura 3-20](#)) lo strumento sembra buono in quanto nel 2014 l'intervallo di confidenza comprende lo 0 e pertanto non si può rifiutare l'ipotesi nulla anche se il relativo p-value (pari a 6.5%) è relativamente basso.

Includendo la popolazione ([Figura 3-21](#)), si verifica un andamento opposto della variabile relativa alla distanza OPB interagita con la time dummy riferita al 2014, rispetto tutti i casi precedentemente analizzati. Prima della diffusione della banda larga ultraveloce, si registra una correlazione altamente significativa (p-value uguale al 2.4%). È possibile riscontrare che un aumento unitario della distanza tra un comune e l'OPB più vicino, incrementa il numero di startup di circa lo 0.025%. Pertanto, l'ipotesi fondamentale della validità non risulta rispettata in questo caso, e utilizzare comunque la stima IV comporterebbe risultati ancora più distorti e non risolverebbe il problema di selezione verificatosi nell'OLS.

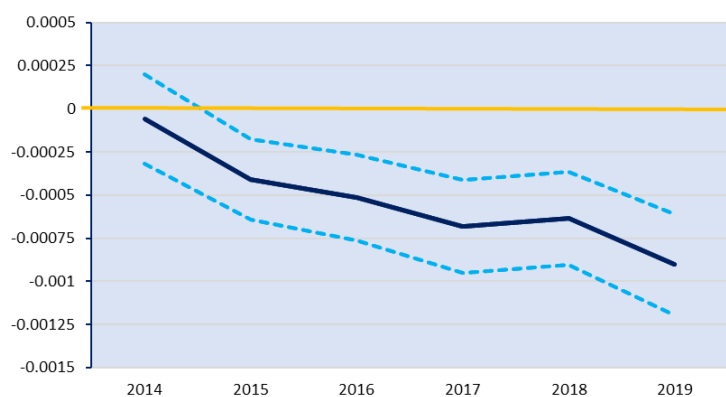


FIGURA 3-20 1° TEST ESOGENEITÀ PER 'ALTRI SETTORI'

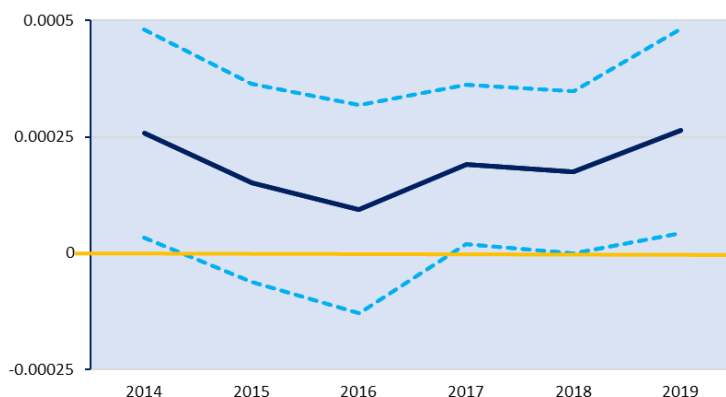


FIGURA 3-21 2° TEST ESOGENEITÀ PER 'ALTRI SETTORI'

Nella [Tabella 3-15](#) sono comunque riportate le evidenze ottenute dalla stima IV suddivise per settore. La copertura con banda larga ultraveloce risulta significativa ($p\text{-value} < 5\%$) solo per il settore dei servizi e l'effetto è di poco superiore al 6,9%. L'impatto aggiuntivo dell'FTTH sulla fibra invece non risulta mai significativo. Tuttavia, non si ritengono "buone" regressioni, in quanto la variabile strumentale non rispetta sempre la condizione fondamentale della validità. Per questo motivo si rinuncia ad approfondire ulteriormente i valori ricavati e si preferisce fare affidamento al modello OLS per quanto riguarda i risultati riscontrati in settori diversi.

TABELLA 3-15 MODELLO IV CON SUDDIVISIONE SETTORIALE

	(1)	(2)	(3)	(4)
	log(servizi)	log(altri)	log(servizi)	log(altri)
UBB	0.0693** (0.0252)	0.0055 (0.0160)	0.0691** (0.0254)	0.00551 (0.0160)
OLT distance	0.0000461 (0.0000518)	0.000013 (0.000032)	0.0000436 (0.000052)	0.000013*** (0.000032)
Gdu	0.00635 (0.00298)	0.002 (0.0031)	0.0065 (0.005)	0.00199 (0.0031)
università	0.0117 (0.0117)	0.0175 (0.011)	0.0118 (0.0171)	0.0175 (0.011)
aree bianche	0 (.)	0 (.)	0 (.)	0 (.)
patents	-0.000013* (0.0000006)	-0.00000035 (0.0000005)	-0.0000013* (0.0000007)	-0.00000035 (0.0000005)
FTTH			0.0103 (0.0159)	-0.0011 (0.009)
Year FE				
Municipality FE	X	X	X	X
Year-Region FE	X	X	X	X
gduXyr	X	X	X	X
areebiancheXyr	X	X	X	X
universitàXyr	X	X	X	X
popolazioneXyr			X	X
First stage F-test	211.243	211.243	111.71	111.71
<i>N</i>	54980	54980	54980	54980
<i>R</i> ²	0.023	0.019	0.023	0.019

Standard errors in parentheses

Source: auto.dta

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Conclusioni

Con questa tesi si è posto l'obiettivo di analizzare il legame tra il dispiegamento della fibra ottica e la nascita di nuove startup innovative, una variabile determinante per la crescita tecnologica di un paese. Attraverso l'innovazione è possibile raggiungere guadagni di produttività generalizzati per l'intera economia e sostenere la competitività a lungo termine. Dunque, studiarne i principali driver è essenziale per supportarne il progresso.

Sebbene l'impatto economico positivo della banda larga ultraveloce sia stato approfonditamente dimostrato, il legame tra le nuove reti NGA e l'innovazione non è stato oggetto di molte ricerche passate. Il tema trattato ha dunque un buon grado di novità e si auspica possa diventare un utile spunto di riflessione per ricerche future. Si constatano comunque alcuni aspetti coerenti e divergenti con la letteratura precedente.

Dai risultati si evince che la banda larga ultraveloce influenza positivamente l'entrata di startup innovative nel mercato italiano, con un effetto complessivo intorno al 13,6% ricavato dalla stima a variabili strumentali. Sebbene il concetto di startup non sia necessariamente associato a piccola impresa, generalmente esse hanno dimensioni limitate. Con quest'ultima constatazione, è possibile ricavare un aspetto contrastante con la letteratura precedente, nella quale era stato rilevato che solo l'ingresso delle grandi imprese fosse influenzato positivamente (Cambini & Sabatino [2020]) e una forte relazione positiva tra la dimensione delle imprese e i ritorni derivanti dall'ICT (Tambe e Hitt [2012]).

L'effettivo impatto sulla nascita di nuove startup innovative è complessivamente positivo ma in determinate suddivisioni l'investimento genera un incremento avente una portata significativamente minore. Alcuni comuni in cui si riscontrano tali evidenze appartengono alle aree bianche o non dispongono di un polo universitario. Nelle zone a fallimento di mercato potrebbe essere più costoso disporre di una rete di asset complementari, e dunque è presumibile che la produttività marginale dell'ICT sia minore del suo costo marginale. Nelle aree bianche, oltre a quanto appena detto, è probabile che sia scarsamente disponibile il capitale umano dell'alta istruzione necessario per trarre benefici dalla fibra (McCoy et al. [2016]) e aumentarne il grado di adozione (Bayo-Moriones & Lera-López [2007]).

Rimane invece poco chiaro il meccanismo per cui non si è riscontrato un effetto significativo dell'ICT avanzate nelle zone ad alta urbanizzazione, in quanto in linea con la letteratura passata si attendeva un impatto positivo derivante dallo sfruttamento di un'efficiente piattaforma digitale (raggiungibile grazie alla vicinanza ad attività complementari) e la conseguente penetrazione nelle economie locali maggiormente sviluppate. Ciò non deve allarmare poiché in tale contesto risulta forte l'impatto derivante dall'installazione della Fiber To The Home. Si rileva che nei grandi comuni sia necessario il raggiungimento della miglior tecnologia disponibile per ottenere effetti indiretti significativi sulle startup innovative.

Discriminando la variabile dipendente a seconda del settore a cui appartengono le startup innovative, si registrano maggiori impatti per le attività operanti nei servizi, sebbene l'influenza derivante dal dispiegamento della fibra risulti perennemente significativa anche nei comparti agricoltura/pesca, industria/artigianato, commercio e turismo.

Infine, è essenziale sottolineare alcuni limiti della strategia empirica utilizzata.

L'analisi si concentra esclusivamente su un particolare paese, particolarmente caratterizzato da alta frammentazione industriale e basse competenze digitali. Dunque, non è possibile ipotizzare la validità delle conclusioni ottenute in contesti esterni. Visto il forte impatto

aggiuntivo derivante dalla tecnologia Fiber To The Home, potrebbe poi essere interessante suddividere i dati in base alla velocità di connessione effettivamente raggiunta nei diversi comuni e verificare se queste differenze abbiano impatti eterogenei sulle startup innovative. Inoltre, sarebbe possibile perfezionare l'analisi integrando altri fattori in grado di influenzare la nascita di startup innovative. Alcune di queste variabili (a livello comunale) potrebbero essere il numero di imprese, il pil, l'età media della popolazione, il tasso di occupazione o la presenza di incubatori. Infine, per una miglior visione riguardante le startup, sarebbe auspicabile non studiarne esclusivamente l'entrata ma esaminare anche altre variabili come l'uscita dal mercato, il fatturato, indici di redditività o la probabilità di sopravvivenza.

Sintesi dei risultati ottenuti

In quest'ultima sezione è riassunta la strategia empirica utilizzata insieme alle principali evidenze ottenute riguardo la variabile principale d'interesse.

Nell'analisi si sfrutta l'introduzione scaglionata di infrastrutture a banda larga ultraveloce in Italia, integrando tali dati con informazioni sulle startup innovative a livello comunale. Nelle prime fasi si è valutato qualitativamente quali variabili non osservabili abbiano il potenziale di influenzare sia la nascita delle startup sia la diffusione della fibra. Oltre all'UBB (la cui natura sarà commentata in seguito), è probabile che anche l'ADSL risulti endogena: all'inizio dell'arco temporale studiato, nella maggior parte della penisola italiana è già stata dispiegata e una volta inclusi gli effetti fissi a livello di comune, il coefficiente catturerà variazioni positive soltanto nei comuni in cui precedentemente non era presente. In tal modo gli incrementi inclusi nella stima econometrica riguarderanno soltanto le piccole unità territoriali, caratterizzate al tempo stesso da arretratezza infrastrutturale digitale e scarsa numerosità di startup innovative. Per quanto riguarda la fibra, è invece possibile catturarne la variazione nel tempo e si è quindi deciso di focalizzare l'analisi esclusivamente su quest'ultima variabile. Sebbene anche la banda larga ultraveloce di qualità superiore, FTTH, sia potenzialmente endogena, si è assunto che driver della relativa diffusione siano caratteristiche idiosincratiche dei comuni e che tali correlazioni rimangano costanti nel tempo, in modo da catturare eventuali legami con variabili non osservabili includendo gli effetti fissi. Studiando il coefficiente risultante sarà così possibile valutarne l'impatto aggiuntivo rispetto all'effetto della variabile principale d'interesse.

Le specificazioni OLS migliori includono l'effetto temporale a livello di anno-regione e l'effetto fisso a livello comunale. È emerso un effetto altamente significativo ($p\text{-value} < 1\%$) della banda larga ultraveloce, il cui dispiegamento corrispondeva inizialmente a un aumento pari al 7,6% delle nuove startup innovative. Aggiungendo l'FTTH, l'impatto dell'UBB diminuisce fino al 6,66% poiché quest'ultima ingloba la frazione dell'impatto derivante dalla connessione di qualità superiore. Includendo ventuno nuove variabili generate dall'interazione tra le variabili di controllo e le time dummies corrispondenti all'arco temporale 2013-2019, l'effetto della variabile d'interesse diminuisce di circa 1 punto percentuale. L'impatto incrementale legato a FTTH varia da 32,6% a 19,7%.

Si è poi suddivisa la variabile dipendente in base al settore di appartenenza confrontando le startup operanti nel settore dei servizi con un'unica variabile comprendente le attività appartenenti all'industria/artigianato, commercio, agricoltura/pesca e turismo. La presenza della fibra, in un certo comune, genera un incremento di iscrizioni di startup operanti nei servizi del 6,04%. L'effetto diminuisce di poco meno di un punto percentuale nel caso si

includa anche la variabile relativa alla tecnologia FTTH. Per gli altri settori, la variabile UBB genera ancora un effetto significativo positivo anche se con impatto minore, infatti al dispiegamento delle reti corrisponde un aumento compreso tra il 2,62% e 2,17% delle iscrizioni annuali.

Suddividendo in base al grado di urbanizzazione, la copertura UBB non risulta sempre significativa. Nei comuni a medio-bassa urbanizzazione l'effetto della fibra è statisticamente diverso da zero (rispettivamente pari all'1,6% e 3,5%). Nelle unità territoriali densamente popolate perde invece di significatività ma si registra un forte impatto legato all'FTTH pari al 32,6%.

Nelle aree bianche, sia la fibra sia l'FTTH risultano statisticamente significative a livello dell'1% generando tuttavia un effetto minore rispetto al caso generale. Al contrario l'effetto di entrambe aumenta analizzando le realtà nei restanti comuni.

Sebbene sia in presenza che in assenza di un ateneo universitario si rilevi alta significatività della variabile UBB, l'effetto nei comuni, in cui è probabilmente presente una maggior conoscenza tecnica e scientifica, è compreso tra il 21,3% e 27,2%, ben maggiore del 4,29% e 4,06% registrati negli altri. Anche per la variabile FTTH vale la stessa considerazione e addirittura raggiunge il 58,2%.

Attraverso il metodo event study, si è mirato a risolvere due problemi intrinseci alla stima OLS: probabile endogeneità della banda larga ultraveloce e sottostima degli effetti dinamici che si accumulano nel tempo. La prima criticità è confermata in quanto si rileva prima dell'entrata della fibra qualche effetto statisticamente diverso da zero, ciò implica che l'entrata non sia random ma correlata a un fattore che impatta sulla variabile dipendente. I periodi -5, -4 e -3 sono indici di endogeneità e ciò manifesta un'ulteriore prova teorica della necessità di approfondire l'analisi attraverso lo studio con variabili strumentali. Inoltre, si constata che i risultati ottenuti fino a questo momento nascondono effetti incrementali over time in quanto le medie dei coefficienti post treatment, in tutte le specificazioni analizzate, risultano perennemente superiori alle corrispondenti stime OLS. Per mezzo delle variabili strumentali si cerca di risolvere il problema di selezione riscontrato, ovvero i comuni in cui si registra un maggior numero di startup innovative, hanno probabilità superiori di adottare la tecnologia a banda larga ultraveloce. Si affronta dunque tale endogeneità utilizzando come principale fonte di variazione esogena la distanza tra ogni comune e il suo nodo dorsale nazionale più vicino, e controllando attentamente le tendenze temporali sottostanti l'adozione della nuova tecnologia a banda larga ultraveloce. Discriminando la variabile dipendente rispetto al settore di appartenenza non si sono registrate particolari variazioni rispetto al caso generale. Tra la stima OLS e la media dei coefficienti post treatment derivanti dall'event study si verifica un simile scostamento relativo nelle due suddivisioni.

Dopo aver introdotto la variabile strumentale, si constata un effetto decisamente superiore rispetto a quanto ottenuta finora. Perciò si mette in dubbio l'esogeneità dello strumento, il quale potrebbe catturare qualcosa che enfatizzi l'effetto della variabile d'interesse. A tale fine, si sono effettuate cinque regressioni a livello generale in cui le covariate principali sono time dummies interagite con la distanza OPB. Si riesce dunque a ottenere una forte prova della validità dello strumento poiché l'interazione della distanza OPB con la dummy 2014 non risulta significativamente diversa da zero si può concludere che prima della fibra non ci fosse nessuna correlazione. Si è verificata la validità dello strumento anche nel caso in cui la variabile dipendente sia osservata suddividendola per settore. Si dimostra l'assenza di correlazione prima dell'introduzione della fibra per quanto riguarda i servizi e si può concludere che lo strumento esprima bene gli effetti inerenti a questa variabile dipendente. Per quanto riguarda gli altri settori, lo strumento sembra inizialmente buono ma includendo

la popolazione, prima della diffusione della banda larga ultraveloce, si registra una correlazione altamente significativa (p-value uguale al 2.4%). L'ipotesi fondamentale della validità non risulta rispettata in questo caso, e utilizzare sfruttare la stima IV comporterebbe di accettare risultati ancora più distorti rispetto a quanto verificatosi nell'OLS.

Nel caso generale, la variabile d'interesse UBB risulta sempre significativa in tutte le regressioni, anche includendo i trend temporali condizionati alla popolazione di ciascun comune. L'implicazione derivante dall'essere una grande città viene presumibilmente catturata da tali covariate, comportando una diminuzione dell'impatto della banda larga di circa 20 punti percentuali (da circa 32% a 13,6%). Confrontando con le stime OLS corrispondenti, si registrano effetti più di 5 volte superiori a quanto ricavato col modello dei minimi quadrati ordinari. Ciò suggerisce che la stima OLS sia troppo bassa ed è coerente con la distorsione presentata precedentemente nel metodo event study e a inizio analisi con le variabili strumentali. L'errore standard della stima IV è circa ventidue volte più grande dell'errore standard OLS, tuttavia, è possibile affermare che le stime siano significativamente diverse tra loro a un livello di fiducia del 5% e anche rifiutare l'ipotesi nulla che sia uguale a quanto ottenuto studiando i coefficienti post treatment derivanti dall'event study. Al contrario, la variabile relativa a FTTH, ancora altamente significativa era stata sovrastimata nella stima OLS e anche in questo caso si riscontra una differenza statisticamente significativa. Includendo i controlli per la popolazione si ottiene un'evidenza divergente con quanto riscontrato nelle specificazioni precedenti, ovvero una correlazione negativa tra la fibra ad alta qualità e la nascita di nuove startup innovative. L'interpretazione corretta di tale variabile è l'impatto aggiuntivo dell'FTTH sulla fibra, ovvero lo scostamento rispetto all'output derivante dalla regressione in cui è inclusa esclusivamente UBB e dunque il dispiegamento non implica una diminuzione delle startup innovative.

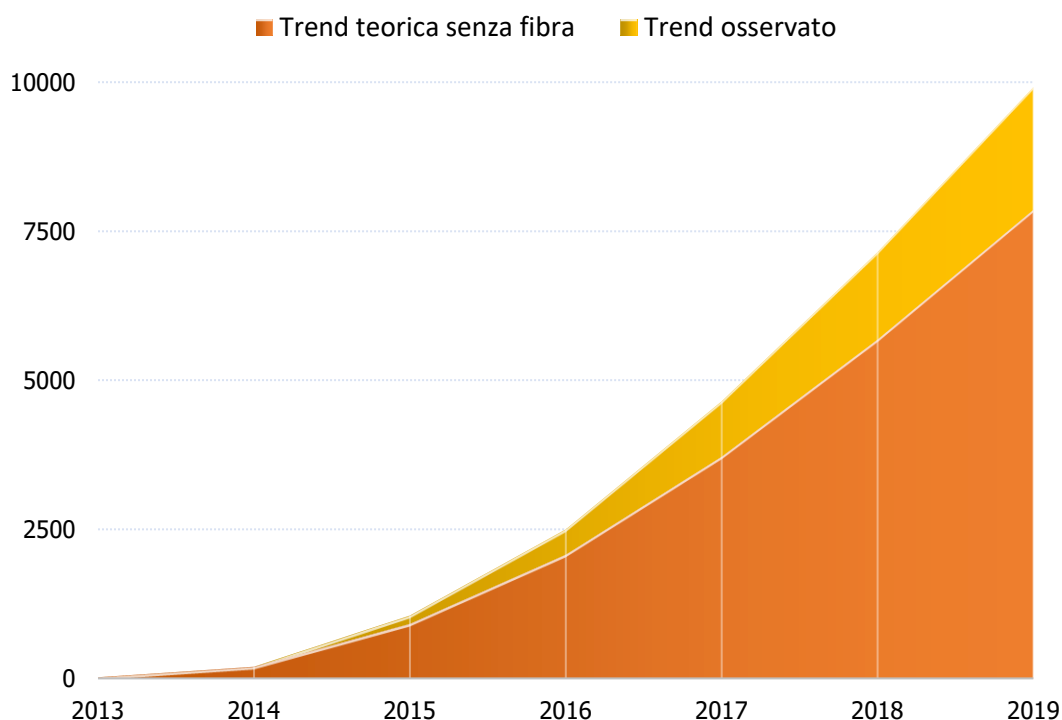
CONCLUSIONI E EVIDENZE OTTENUTE

	<i>UBB</i>	<i>FTTH</i>	Municipality FE	Year-Region FE	Trend temporali	PopolazioneXyr
OLS	0.063					
OLS	0.056	0.197				
Event Study	0.180					
Event Study	0.210					
Event Study	0.074					
Event Study	0.068					
IV	0.328					
IV	0.315	0.138				
IV	0.136					
IV	0.136	-0.0043				

CONCLUSIONI 2 EVIDENZE OTTENUTE CON SUDDIVISIONE SETTORIALE

	UBB		FTTH		Municipality FE	Year-Region FE	Trend temporali	PopolazioneXyr
	Servizi	Altri settori	Servizi	Altri settori				
OLS	0.060	0.026						
OLS	0.052	0.022	0,282	0,149				
Event Study	0.120	0.055						
Event Study	0.140	0.066						
Event Study	0.145	0.066						
Event Study	0.123	0.054						
IV	0.069	0.0055						
IV	0.069	0.0055	0,0103	-0,0011				

Per una miglior comprensione dell'effetto della fibra, si è riportato quanto ottenuto in un grafico controfattuale, evidenziando così la differenza tra il trend osservato e quanto sarebbe successo se non ci fosse stata la fibra. Si ritiene che la miglior stima sia quanto ottenuto con la variabile strumentale e descritto nelle colonne [3] e [4] della [Tabella 3-13](#), includendo dunque anche i trend temporali condizionati alla popolazione di ogni comune. In queste specificazioni, si riscontrava che la diffusione della fibra all'interno di un comune generasse un incremento del 13,6% delle startup innovative nascenti. Dal 2015 al 2019 l'incremento dovuto alla fibra è rispettivamente il 16,28%, 19,61%, 23,49%, 21,58% e 21,26% del totale delle startup innovative nascenti. Al termine del periodo studiato, si sono dunque complessivamente iscritte 9916 startup e si ritiene, con un buon grado di fiducia, che la banda larga ultraveloce sia stato un fattore determinante e responsabile della nascita di circa 2056 attività.



CONCLUSIONI 3 GRAFICO CONTRO FATTUALE

CONCLUSIONI 4 DIFFUSIONE FIBRA E STARTUP INNOVATIVE

13.60%	2013	2014	2015	2016	2017	2018	2019
<i>Diffusione startup</i>	18	176	860	1438	2141	2492	2761
<i>Diffusione fibra</i>	0	0	1030	2079	3713	3970	4323
<i>Incremento dovuto a fibra</i>	0	0	140	283	505	540	588
<i>Incremento teorico senza fibra</i>	18	177	720	1160	1645	1962	2178
<i>Incremento osservato</i>	18	177	860	1443	2150	2502	2766
<i>Trend teorica senza fibra</i>	18	195	915	2075	3720	5682	7860
<i>Trend osservato</i>	18	195	1055	2498	4648	7150	9916

Ringraziamenti

Vorrei dedicare queste ultime righe alle persone che hanno avuto un forte impatto su di me negli ultimi cinque anni.

Desidero innanzitutto ringraziare il professor Carlo Cambini, Relatore, ed il professor Lorien Sabatino, Co-relatore, per la disponibilità e la professionalità con cui mi hanno sapientemente guidato durante la stesura della tesi.

Ringrazio mia mamma, mio babbo e Alice per l'amore e il supporto che non mi hanno mai fatto mancare. Un grazie sincero a Marta, grande compagna in questa parte di vita, che continua a rendere le mie giornate migliori. Grazie a tutti i miei più cari amici, per la spensieratezza e la felicità che mi fanno provare. Grazie a tutti i compagni di corso che ho trovato sulla mia strada, ho appreso tanto da voi e vi porterò sempre nel cuore.

Infine, vorrei dedicare questo traguardo a mio nonno Alfredo, da sempre un mio grande tifoso, che oggi sarebbe stato tanto fiero di me.



**Politecnico
di Torino**



**Politecnico
di Torino**