

POLITECNICO DI TORINO

Corso di Laurea Magistrale in Ingegneria Gestionale



TESI DI LAUREA MAGISTRALE

L'impatto del Covid-19 sul traffico Internet in Italia

Relatore: Carlo Cambini

Candidato: Marta Castellina

ANNO ACCADEMICO 2020-2021

Indice

| | |
|--|----|
| Introduzione | 1 |
| Capitolo 1 - Il ruolo dell'ICT in periodo COVID-19 | 3 |
| 1.1 Disuguaglianze in periodo COVID-19 | 4 |
| 1.2 Il ruolo della mobilità | 9 |
| 1.3 Il ruolo dei media: tra persuasione e fake news..... | 13 |
| 1.4 I cyber attacchi..... | 14 |
| 1.5 L'ICT in periodo COVID-19 in Italia..... | 15 |
| Capitolo 2: Analisi descrittiva | 20 |
| 2.1 Il dataset | 20 |
| 2.2 Andamenti temporali delle otto tipologie di traffico | 27 |
| 2.3 Province più e meno colpite | 35 |
| 2.4 Traffico nelle province più e meno colpite..... | 39 |
| 2.5 Suddivisione in zone | 45 |
| 2.6 Considerazioni di fine capitolo | 57 |
| Capitolo 3: Analisi di regressione | 58 |
| 3.1 Preparazione del dataset..... | 58 |
| 3.2 Definizione del modello..... | 59 |
| 3.3 Interazioni con i periodi..... | 62 |
| 3.4 Interazioni con le zone d'Italia | 67 |
| 3.5 Interazione con le province più colpite | 70 |
| 3.6 Considerazioni di fine capitolo | 73 |
| Conclusioni | 74 |
| Appendice..... | 76 |
| Indice delle tabelle | 83 |

| | |
|-------------------|----|
| Bibliografia..... | 85 |
| Sitografia..... | 88 |

Introduzione

Lockdown. Questa, secondo gli esperti del Collins English Dictionary, è stata la parola simbolo dell'anno 2020: lo scoppio della pandemia di Covid-19 ha fatto registrare un numero totale di 250 milioni di casi di utilizzo del vocabolo, numero nettamente superiore rispetto al 2019, in cui erano stati tracciati solamente 4 mila casi.

Quest'unico termine, spiegano i linguisti del Collins, è in grado di accomunare, come non avveniva dai tempi delle Grandi Guerre, tutte le generazioni e le popolazioni di ogni parte del mondo, che allo stesso modo hanno dovuto far fronte a un'emergenza di carattere globale.

Non è la prima volta nel corso della storia che si verifica una pandemia, ma il grande discrimine tra l'epidemia di Covid-19 e quelle verificatesi nel passato è il modo in cui, grazie ai progressi tecnologici degli ultimi anni, le persone hanno potuto trasferire online parte delle proprie vite. Così, seppur spinti a rispettare politiche di tipo *stay at home*, è stato possibile mantenere una parvenza di normalità: a prescindere dalle diverse opinioni, favorevoli o contrarie, non si può negare il ruolo fondamentale che l'ICT ha avuto nel portare scuola, lavoro e tempo libero nelle case di tutti.

Negli Stati Uniti così come in Italia, le compagnie telefoniche si sono ritrovate a fronteggiare il più grande stress test di sempre, con un tasso di crescita esponenziale per quanto riguardava il traffico giornaliero sulle linee Internet.

Per tali motivi, lo scopo del presente lavoro di tesi è quello di indagare in che misura il Covid-19 abbia impattato sui consumi del traffico Internet in Italia.

Il primo capitolo tratterà il ruolo dell'ICT, in tutte le sue numerose sfaccettature, nei primi mesi del 2020. Grazie all'analisi della letteratura esistente, sono state identificate diverse tematiche: la prima riguarderà le disuguaglianze emerse, sia quelle in ambito socio-economico, sia quelle del *digital divide*; si focalizzerà poi sul ruolo della mobilità e su come le varie strategie adottate dai diversi Paesi siano state più o meno rispettate; sarà poi analizzato il ruolo dei mass media e saranno accennati i cambiamenti che si sono verificati

nell'universo dei cyber attacchi. Per ultimo sarà presentato un approfondimento sulla situazione italiana di tali temi.

Il secondo capitolo illustrerà il dataset utilizzato, mostrando alcune statistiche descrittive: inizialmente saranno analizzati i consumi medi giornalieri per utente per diverse tipologie di traffico e successivamente sarà effettuata la stessa analisi nel caso delle province più e meno colpite dal Coronavirus, dapprima a livello nazionale e poi separatamente per le tre zone di Nord, Centro e Sud Italia.

Nel terzo capitolo saranno presentati i risultati ottenuti dalle analisi di regressione, che sono state effettuate separatamente per le fonti di traffico che hanno mostrato nelle statistiche descrittive una relazione con l'andamento dell'epidemia. Focus dell'analisi sarà la correlazione tra i consumi medi giornalieri e, dapprima, l'interazione dei casi con i periodi più importanti nell'orizzonte temporale in esame e in secondo luogo con le diverse zone italiane di appartenenza.

Capitolo 1 - Il ruolo dell'ICT in periodo COVID-19

In un momento in cui il distanziamento sociale è uno dei migliori modi per arginare la diffusione del Coronavirus, il settore delle comunicazioni risulta essere l'asse portante per garantire la continuazione di quella che prima era la normalità: informazione, scuola, lavoro e intrattenimento si sono spostati online, costituendo uno stress test per le compagnie di telecomunicazione, non abituate a un simile traffico giornaliero.

Rispetto ad aprile 2019, infatti, nel 2020 si è assistito a un incremento del 7% di utenti, corrispondente a quasi 300 milioni di nuovi user. Akamai, una delle principali piattaforme per la distribuzione di contenuti web (tra i cui clienti figurano anche colossi come Facebook, Amazon, FedEx e Microsoft), ha registrato, solo nel periodo febbraio-marzo 2020, una crescita del 30% di traffico Internet, con picchi anche di 167 Tbps, notevolmente maggiori rispetto agli 82 Tbps di marzo 2019.

Il ricorso all'informazione è stato significativo, soprattutto nelle prime fasi della diffusione del virus, quando ancora regnava molta incertezza: il New York Times riferisce, per esempio, un incremento di lettori online del 50% nel mese di marzo 2020 e nello stesso articolo mostra come a beneficiare di questo periodo siano state in particolar modo le testate giornalistiche locali. È pur vero, però, che una simile tendenza è stata rilevata anche per i siti che producono contenuti di disinformazione: le fake news sono aumentate in modo considerevole (+19% rispetto al 2019), e il 40% di quelle prodotte a marzo 2020 riguardavano il Coronavirus. La stessa Organizzazione Mondiale della Sanità ha definito questo fenomeno come *infodemia* e ha dovuto provvedere ad adottare delle iniziative volte a limitarne la diffusione su larga scala, anche grazie alla collaborazione con le principali piattaforme online.

È evidente l'utilizzo massiccio della rete per ricorrere a forme alternative di scuola e lavoro: così come l'e-learning ha sostituito le lezioni in presenza – nel mese di aprile 2020, infatti, quasi il 90% degli studenti di tutto il mondo ha dovuto affrontare un modello di didattica a distanza – così, nei casi in cui sia stato possibile, il lavoro è diventato in forma telematica, il cosiddetto *remote working*.

Con una simile crescita di utenti online non è mancato, ovviamente, un altresì netto aumento di cyber attacchi. In linea con tutte le epoche di crisi – a partire da quella economica del 2008 – il COVID-19 non ha certo rappresentato un’eccezione: i cyber criminali hanno utilizzato a loro vantaggio la vulnerabilità degli individui, al punto da spingere una delle più importanti aziende in materia, Panda Security, a definire il Coronavirus come la più grande minaccia in ambito cybersecurity di sempre.

1.1 Disuguaglianze in periodo COVID-19

Si possono distinguere due tipologie di disuguaglianze, già esistenti prima della pandemia, ma acuite ulteriormente da essa: quelle digitali – altresì note come *digital divide* – e quelle di carattere socio-economico.

Il *digital divide*

La disuguaglianza nell’accesso a internet può avere, principalmente, due fonti: geografica ed economica.

La prima tipologia riguarda soprattutto quei Paesi con una vasta estensione territoriale, che faticano a garantire una copertura soddisfacente: le zone urbane beneficiano, di norma, di una buona connessione, mentre in quelle rurali mancano le infrastrutture per poter fornire un servizio adeguato. In Europa – fatta eccezione per l’UK, che secondo il regolatore delle telecomunicazioni Ofcom è riuscito a fornire una copertura di banda larga al 99,5% della popolazione – la maggior parte dei Paesi riscontra tale problema, seppur in misura molto ridotta rispetto ad altre parti del mondo come gli USA. Non a caso in America questo tema era già particolarmente rilevante molto prima dello scoppio dell’epidemia di COVID-19. Gli ultimi governi hanno spostato la loro attenzione in ambito tech anche su questo tema, che attualmente è diventato centrale nei programmi elettorali dei candidati alle presidenziali: Biden promette di investire 20 miliardi di dollari per espandere la banda larga anche nelle zone rurali; Trump, dall’altra parte, garantisce il suo impegno per assicurare a ogni cittadino una connessione a Internet ad alta velocità.

Sicuramente complice di questa situazione è il fatto che per le società di telecomunicazioni sia più profittevole investire nelle aree urbane – che mediamente in America contano 2000 persone per miglio quadrato – rispetto a quelle rurali, che invece ne hanno solamente 10.

Tuttavia, il motivo principale per cui finora si è sempre temporeggiato per non effettuare una simile operazione è dato dai costi elevati che genererebbe: è stato stimato che, per garantire a tutti i cittadini americani una connessione in fibra ottica, lo stato dovrebbe effettuare una spesa compresa tra gli 80 MLD \$ e i 150 MLD \$.

Se prima, però, gli abitanti delle zone rurali potevano recarsi nelle scuole o nelle biblioteche per utilizzare la rete, con le politiche di *stay at home* e con la chiusura dei locali pubblici si sono ritrovati in gran parte isolati.

Per fronteggiare le problematiche sorte in queste aree durante la pandemia, molte compagnie di telecomunicazione hanno concesso al pubblico il proprio hotspot WiFi gratuitamente, mentre altre hanno fornito dei pacchetti famiglia comprensivi di contratti ad hoc. Questo perché in alcuni casi, più che l'accesso alla rete in sé, il vero ostacolo è costituito dai prezzi elevati dei contratti con le compagnie telefoniche. Il report *The Great Digital Divide* di Capgemini evidenzia come già prima della pandemia il 69% delle persone senza accesso a Internet vivesse in assoluta povertà e come il 48% di esse desiderasse avere accesso alla rete, che in questo periodo ancor di più permette di rimanere connessi con il resto del mondo.

La disponibilità di una connessione a Internet veloce ha una chiara correlazione con la capacità di un individuo di isolarsi nel rispetto di policy di tipo *stay at home* (L. Chiou e C. Tucker [2020]): utilizzando i dati di localizzazione provenienti dai cellulari, si è osservato come a febbraio 2020 negli USA (in un momento in cui, quindi, non si era ancora verificata l'impennata di contagi che ha poi caratterizzato i mesi successivi) i dispositivi associati a regioni ad alto reddito avessero maggiori probabilità di spostarsi da casa durante il giorno. A marzo 2020, contrariamente, è stata rilevata una forte correlazione tra la probabilità che il dispositivo restasse a casa e il fatto che si trovasse in una regione ad alto reddito.

Una spiegazione della differenza nella mobilità delle varie regioni può esser data dalla forte correlazione esistente tra l'accesso a Internet ad alta velocità e un reddito familiare medio-alto: la combinazione di questi due fattori sembra favorire maggiormente il rispetto delle

politiche *stay at home*. Partendo da un campione di 20 milioni di dispositivi mobili, collezionati in modo anonimo tramite i dati di localizzazione, sono stati creati alcuni gruppi in base al reddito familiare e i tipi di connessione sono stati suddivisi tra cablata, in fibra ottica o di tipo DSL. Già studi precedenti (Miller e Tucker [2011]; Tucker e Yu [2019]) avevano dimostrato la relazione tra alto reddito e infrastrutture per connessioni ad alta velocità. La particolarità dello studio di Chiou e Tucker è rappresentata dall'introduzione del fattore che monitora la capacità di isolarsi, aggiungendo, infine, una variabile dummy rappresentante le direttive dello stato (esistenti o meno), che risultano non produrre effetti significativi a meno che non siano presenti anche un alto reddito o una connessione veloce.

È necessario a questo punto precisare che Chiou e Tucker hanno utilizzato solo la rete fissa in quanto la rete mobile non può essere considerata un suo sostituto perfetto per tre motivi:

- I piani tariffari in questo caso hanno, solitamente, dei limiti in termini di dati a disposizione;
- Allo stesso modo sono presenti limiti per quanto riguarda i dati trasmessi tramite hotspot o tethering;
- Molte attività sono pensate per esser svolte su un computer.

Le disuguaglianze socio – economiche

È possibile rilevare anche disuguaglianze in ambito economico e sociale: suddividendo la popolazione in base al reddito, all'etnia o al livello di istruzione, si può notare come il Coronavirus abbia avuto, mediamente, effetti più o meno rilevanti a seconda della categoria di appartenenza.

È senz'altro significativo in questo ambito lo studio di K.A. Couch, R.W. Fairlie e H. Xu [2020], che ha analizzato l'impatto del COVID-19 sulla disoccupazione dei gruppi etnici minoritari.

Le restrizioni volte a intensificare il distanziamento sociale hanno portato alla chiusura di un notevole numero di attività che non potevano esser svolte in maniera agile. Conseguenza immediata di ciò è stato un incremento vertiginoso del tasso di disoccupazione, che negli USA ha raggiunto, per la prima volta dopo quasi cento anni, i livelli della Grande Depressione del '29. Da un'analisi più approfondita, si è però potuto

notare come questo aumento non sia stato omogeneo in tutte le categorie analizzate, ma abbia anzi interessato soprattutto le minoranze. Già durante il secolo scorso il rapporto di disoccupazione tra neri e bianchi era 2:1, con una maggiore volatilità per quanto riguarda i primi: nel momento in cui il ciclo di business si indebolisce, solitamente sono loro a essere subito licenziati. Tuttavia si può osservare che, in realtà, il loro livello di occupazione si è mantenuto relativamente alto rispetto a quello della Grande Depressione. Altrettanto non si può dire, invece, della categoria dei *Latinx* (neologismo utilizzato per identificare persone di etnia latinoamericana negli USA), che più di tutti hanno subito le conseguenze di questo periodo.

Si è quindi interessati a capire il motivo per cui minoranze diverse abbiano riscontrato effetti diversi. Per fare ciò è stata utilizzata una decomposizione non lineare per identificare fattori collegati a lavoro, abilità e geografia, che sono poi stati considerati per le tre categorie in esame: quella dei bianchi (utilizzata come categoria base, rispetto a cui rapportare le altre due), quella dei neri e quella dei *Latinx*. I risultati hanno mostrato che i neri hanno una distribuzione industriale favorevole, che in parte li ha protetti dalla disoccupazione, mentre i *Latinx* hanno una distribuzione occupazionale meno favorevole e minori capacità in linea con le richieste presenti sul mercato del lavoro.

Una parziale spiegazione può esser collegata alla distribuzione geografica dei due gruppi: il primo vive soprattutto nel Sud degli USA, dove vive solo una piccola percentuale di tutte le altre etnie, mentre il secondo risiede principalmente a Ovest, dove è presente anche gran parte della comunità asiatica americana, che crea quindi una competizione in ambito lavorativo.

Un'altra correlazione rilevante in ambito di disuguaglianze sociali è quella tra la probabilità che un individuo americano infetto da Coronavirus sviluppi o meno delle complicazioni a seconda dell'etnia di appartenenza, del grado di scolarizzazione e del reddito (E. E. Wiemers, S. Abrahams et al. [2020]). Già studi precedenti (Blumenshine et al. [2008]; Kumar e Quinn [2012]) avevano dimostrato il ruolo di fattori socio-economici relativamente all'esposizione a pandemie di influenza e possono essere utilizzati per capire meglio l'effetto del COVID-19. Suddividendo il problema tra possibilità di contrarre il virus e probabilità che, una volta contratto, sopraggiungano delle complicazioni, emerge facilmente come le persone più povere siano maggiormente esposte a questi rischi: non

solo hanno minor accesso alla sanità (che garantisce alle persone più ricche, invece, di ricevere le cure in modo tempestivo), ma hanno anche minori possibilità di rispettare il distanziamento sociale a causa dei lavori che svolgono, delle loro abitazioni, dei quartieri in cui vivono (di solito a elevata densità abitativa) e della dipendenza dal trasporto pubblico. Lo stato socio-economico influenza, inoltre, fattori quali l'esistenza di patologie pregresse, lo stress e degli stili di vita non sani, che portano poi a delle complicazioni.

Sono state create 4 categorie per analizzare l'esposizione al contagio: ispanici, non-ispanici neri, non-ispanici bianchi e i rimanenti. L'analisi relativa alle probabilità di complicazioni una volta infetti è stata condotta in modo separato per diversi livelli di scolarizzazione (liceo o inferiore, qualche anno di università, laurea o superiore) e per diverse fasce di reddito. I risultati ottenuti evidenziano che i non-ispanici neri corrono maggiori rischi di contrarre il virus rispetto agli altri gruppi e che gli appartenenti ai livelli inferiori di scolarizzazione e alle fasce più basse di reddito sono quelli più esposti alle possibili complicazioni.

In merito alla disuguaglianza economica si è anche osservato uno sbilanciamento nel numero eseguito di test per COVID-19 (S. Schmitt-Grohé, K. Teoh, M. Uribe [2020]). L'implicazione è ovvia: tramite questa analisi si vuole vedere la correlazione tra reddito e accesso al sistema sanitario negli USA.

Da una lettura sommaria dei dati si potrebbe pensare che, siccome il numero totale di test somministrati è lo stesso sia nelle regioni ad alto reddito, sia in quelle a basso reddito, non vi sia alcun tipo di ineguaglianza nella somministrazione: utilizzando la curva di Lorenz – che è un indice di concentrazione/omogeneità con cui si può misurare in che modo un bene trasferibile sia diviso tra la popolazione – si osserva che il reddito è distribuito in modo eterogeneo, ma altrettanto non si può dire del numero di test eseguiti, che segue invece la linea di 45°, indicatore di omogeneità.

È stata quindi effettuata un'analisi di regressione che ha determinato una forte correlazione positiva tra il reddito e il numero di test negativi: è evidente a questo punto che non ci sia un equo accesso ai test in quanto, a parità di test effettuati, ci si aspetterebbe lo stesso numero di esiti negativi. Questo può essere spiegato con il fatto che i più poveri aspettino di sviluppare sintomi seri per richiedere l'assistenza sanitaria, fatto non sorprendente in una delle sanità più costose al mondo. Infatti, come si può leggere sul New

York Times del 16 giugno 2020, negli Stati Uniti i costi per singolo test per Coronavirus oscillano tra i 100 \$ e i 2315 \$, variabilità data dalla mancata regolamentazione degli ospedali da parte del Governo, che, durante una pandemia non fa che allargare ulteriormente la forbice in termini di reddito tra ricchi e poveri.

1.2 Il ruolo della mobilità

I primi casi di Coronavirus sono stati registrati nella città cinese di Wuhan a dicembre 2019 e già l'11 marzo 2020 Tedros Adhanom Ghebreyesus, direttore generale dell'Organizzazione Mondiale della Sanità, ha dichiarato pandemia il focolaio internazionale di infezione da COVID-19. In soli 3 mesi, quindi, il virus si è diffuso con un'incredibile rapidità e per frenare l'impennata dei contagi la maggior parte dei Paesi ha adottato differenti politiche di tipo *stay at home* o, nei casi più gravi, ha imposto veri e propri lockdown. L'ICT è stato un importante alleato per permettere il rispetto degli interventi non farmaceutici, in particolare il distanziamento sociale, che però rappresenta solamente una soluzione sub-ottima a causa delle esternalità presenti: da un lato gli individui sono interessati ai benefici privati che esso apporta, ma dall'altro non si curano dei suoi benefici sociali, come per esempio la riduzione del rischio di congestione degli ospedali. Dopotutto, il rispetto delle norme è un tipico problema di azioni collettive (Ostrom [1991]), in quanto si fronteggia un *common good*: per l'individuo è costoso agire in modo corretto e non riesce a percepire il suo contributo effettivo, anche a causa del problema del free rider, che può essere superato grazie al capitale civico (Guiso et al. [2011]).

Strategie di lockdown, diffusione del virus e mobilità

Paesi diversi hanno adottato strategie diverse per fronteggiare l'epidemia ed è stata evidenziata la correlazione di tali policy con la diffusione del contagio e i pattern di mobilità (N. Askitas et al. [2020]).

Le politiche di intervento che sono state considerate sono quelle riguardanti i viaggi internazionali, il blocco dei trasporti, la cancellazione di eventi pubblici, le restrizioni sugli assembramenti, la chiusura delle scuole, dei posti di lavoro, le restrizioni sulla mobilità

interna e gli inviti a restare a casa. Per quanto riguarda la mobilità, che può essere una misura della conformità alle politiche vigenti, sono stati considerati luoghi diversi: negozi, farmacie e supermercati, parchi, stazioni di transito, posti di lavoro e aree residenziali.

Gli studi precedenti (Gao et al. [2020]) avevano dimostrato che in molte contee americane in cui le restrizioni fossero solo raccomandate ma non imposte, la mobilità non era affatto diminuita. Dopo i primi casi, negozi, supermercati e farmacie hanno visto una diminuzione in termini di densità di persone, mentre è aumentata quella nelle aree residenziali e nei parchi. Dall'analisi di correlazione rispetto all'aumento del numero dei contagi si può osservare che la maggior evidenza di contenimento si ha per quanto riguarda la cancellazione degli eventi pubblici e le restrizioni sugli assembramenti, entrambe misure volte a ridurre un contatto massivo difficile da tenere sotto controllo, che risulta più facile nelle scuole e nei posti di lavoro, nonostante anche in quel caso ci siano considerevoli gruppi di persone insieme. Lo stesso risultato è ottenuto quando l'analisi è ripetuta per quanto riguarda la riduzione della mobilità.

Appare quindi evidente che le migliori politiche da adottare sembrano essere la cancellazione di eventi pubblici e il rafforzamento delle restrizioni sugli assembramenti. In secondo luogo c'è la chiusura delle scuole e dei posti di lavoro e le strette sui trasporti pubblici, quest'ultima conseguenza però delle altre politiche.

Il capitale civico

In assenza di un vaccino o di altre cure mediche, il comportamento umano è essenziale per contenere la rapida diffusione del contagio (Van Bavel et al., [2020]) e per tale motivo i politici di tutto il mondo si sono dovuti appellare alla responsabilità civica dei propri cittadini: tra gli esempi europei spiccano Angela Merkel, che il 18/03/2020 ha definito questo periodo "la più grande sfida per la Germania dopo la seconda guerra mondiale" e ha sottolineato come la chiave per affrontarlo fosse agire collettivamente come società; Emmanuel Macron, che il 16/03/2020 si è appellato al senso di responsabilità e solidarietà del popolo francese, confidando nel fatto che la regola migliore da applicare fosse proprio quella che i cittadini si imponevano da soli; Giuseppe Conte, che il 26/04/2020 ha evidenziato l'importanza della condotta responsabile di ognuno, aggiungendo infine quella

che è diventata una delle frasi simbolo del lockdown italiano: “Se ami l’Italia, mantieni la distanza”.

Tutti e tre i leader citati fanno riferimento al capitale civico: è richiesto ai singoli cittadini di considerare il costo sociale delle proprie azioni individuali. In una pubblicazione del 1993 Putnam lo definisce come “la disponibilità ad agire collettivamente e perseguire attività a favore della società”. Può esser misurato in tanti modi: il numero di donatori di sangue in una regione, l’affluenza alle urne (perché la partecipazione alle votazioni politiche è fortemente correlata con la volontà di contribuire ai *public goods*), il tasso di lettura dei giornali.

Le aree con un maggior capitale civico tendono a essere connesse meglio, sia dal punto di vista economico, sia dal punto di vista sociale (Knack & Keefer [1997]; Tabellini [2010]). Non stupisce quindi che siano le zone in cui il virus si è diffuso con più facilità nelle fasi iniziali, mentre quando le normative sono diventate più stringenti e le informazioni sono state maggiormente diffuse, è proprio in queste città che si è registrato il più evidente arresto nella crescita giornaliera dei contagi (A. Bartscher et al. [2020]; J.M.Barrios et al. [2020]). In queste aree già prima di eventuali lockdown obbligatori la mobilità risulta notevolmente diminuita e anche il numero giornaliero di morti per COVID-19 ha un trend decrescente.

La maggior evidenza del ruolo del capitale civico si ha in quei territori in cui le politiche attuate sono state meno stringenti: è minore il contributo individuale in situazioni in cui ci siano misure obbligatorie come il lockdown totale. Nonostante ciò, però, l’Italia è stata comunque oggetto di numerosi studi, non solo in quanto primo Paese occidentale a esser stato fortemente colpito dall’epidemia, ma anche perché la letteratura esistente in materia spesso l’ha considerata nelle analisi, rendendo così più facile il confronto con i risultati ottenuti. È stata effettuata una simulazione (R.Durante et al., 2020) secondo cui se tutte le province italiane avessero avuto il capitale civico di quelle nel primo quartile della distribuzione ci sarebbero stati solo 2700 morti a maggio, anziché i più di 30000 registrati.

Anche per quanto riguarda la mobilità individuale, che è uno dei principali modi con cui il capitale civico può contribuire alla riduzione dei contagi, è stato condotto uno studio. Utilizzando il numero di viaggi settimanali a testa, è stato effettuato un confronto tra quelli effettuati prima e dopo i primi casi di COVID-19 (che in Italia si sono verificati alla fine della

terza settimana di febbraio). Nelle due settimane seguenti (i.e. la quarta di febbraio e la prima di marzo) si è osservata una significativa diminuzione di viaggi nelle aree ad alto capitale civico, tendenza mantenuta anche dopo la fine del lockdown. È però necessario fare una precisazione: è certamente notevole la differenza nella mobilità tra regioni ad alto e a basso capitale civico, tuttavia è un dato che va contestualizzato. Occorre, infatti, considerare anche le politiche specifiche delle regioni e, soprattutto, il PIL pro capite, in quanto le regioni ad alto reddito di norma hanno più possibilità di svolgere lavori in modalità agile.

Il rispetto delle misure anti-COVID-19 in relazione alla loro durata

Di fronte a misure forti di distanziamento, quali politiche di *stay at home* o di lockdown, i Paesi hanno scelto strategie diverse nella definizione della loro durata: la Cina ha adottato tali policy per un tempo indefinito, mentre i Paesi occidentali come Italia, Francia, Germania e USA hanno stabilito fin da subito una data di fine, che però in molti casi è stata posticipata a causa del continuo aumento di casi. In Italia, per esempio, il Presidente del Consiglio Giuseppe Conte il 9 marzo ha firmato il Dpcm con cui tutta l'Italia, già zona rossa, è entrata in un lockdown che inizialmente sarebbe dovuto durare fino al 3 aprile, ma successivamente la data di termine è stata estesa prima al 13 aprile e poi al 4 maggio.

È stato dimostrato (G. Briscese et al. [2020]) che nei casi in cui non sia specificata una data di fine (come successo in Cina), aumenta la percezione della severità della ragione che ha motivato le limitazioni, ma al contempo aumentano i disagi psicologici; invece, nel caso in cui vi sia una data termine nota, si tende a pensare a un'emergenza temporanea e per tale motivo l'estensione delle misure può portare alla diminuzione della fiducia nelle autorità e al rifiuto a un'ulteriore collaborazione. Ciò avviene per l'effetto di *goal gradient* (A. Bonezzi et al., 2011), secondo il quale più si è lontani dall'obiettivo meno ci si impegna per raggiungerlo. Il ruolo dell'aspettativa diventa quindi cruciale: coloro che sono sorpresi negativamente con un inaspettato prolungamento del confinamento sono meno propensi a continuare lo sforzo rispetto a coloro che invece già lo avevano intuito. Inoltre, coloro che rispettano tutte le misure necessarie (quelle in esame sono state le seguenti: evitare posti affollati, evitare di incontrare amici, parenti, mantenere le distanze di sicurezza dagli altri, limitare le volte in cui si va al supermercato, lasciare casa solo quando è necessario) sembra essere maggiormente colpito dal mismatch sulle aspettative rispetto a chi, invece, non

osserva tutte e 6 le misure in esame, riducendo quindi in loro la volontà di mantenere lo sforzo.

1.3 Il ruolo dei media: tra persuasione e fake news

Il caso Fox News

I media hanno svolto un ruolo importante nella diffusione di notizie sul Coronavirus, avendo anche un effetto persuasivo sul rispetto delle policy adottate dai governi. Quotidiani e telegiornali hanno visto un aumento di pubblico notevole, per tale motivo sarebbe auspicabile che, durante una pandemia, trasmettessero informazioni chiare, reali e soprattutto allineate ai pareri degli esperti in materia.

Ciò non si può dire di Fox News, il canale di news più importante della tv cablata americana, che ha da subito minimizzato il pericolo creato dalla diffusione del virus (Ecarma [2020]), influenzando così anche il pubblico che lo seguiva. La reciproca influenza che ha con Donald Trump è ormai nota dalla pubblicazione del libro *Hoax: Donald Trump, Fox News, and the Dangerous Distortion of Truth* (Brian Stelter, [2020]), che inizia proprio citando il Dr. Marc Siegel, l'esperto in ambito medico intervistato da Fox News, che il 6 marzo 2020 dichiarava che "nel peggior caso possibile, sarà solamente un'influenza".

Particolarmente importante è quindi lo studio effettuato sulla persuasione generata dalle emittenti televisive, in particolare Fox News e CNN, sul rispetto del distanziamento sociale e delle politiche *stay at home* (A.Simonov et al. [2020]). È stato rilevato che, sebbene molti Stati avessero già predisposto le proprie dichiarazioni di emergenza, i telespettatori che seguivano Fox News hanno aumentato la loro propensione al distanziamento sociale solo a partire dal 13 marzo 2020, giorno in cui il presidente Trump ha dichiarato l'emergenza nazionale. La CNN, invece, da subito ha invitato all'interno dei programmi medici esperti che seguissero le linee dell'OMS. In questo caso non compare nessuna evidenza di correlazione tra telespettatori e minor sforzo nell'impegno per il *social distancing*.

L'infodemia

A un aumento della domanda di informazione corrisponde un aumento anche nell'offerta, che tuttavia si diversifica su due fronti: quella buona e quella cattiva. Se quella buona è più

che mai utile in periodi di crisi come quello attuale, quella cattiva, che non ha alcuna base medico-scientifica, non fa che infondere preoccupazione, panico e addirittura, in certi casi, isteria. In tale ambito, i principali mezzi di *fake news* sono i giornali locali, e, soprattutto, i social network (C.R. Mejia et al., 2020). Cure miracolose, teorie complottiste e notizie volte a screditare personaggi di spicco (si pensi a Bill Gates, che è diventato il bersaglio numero 1) sono i temi principali di cui trattano. Uno studio internazionale della University of New South Wales in Australia ha stimato circa 800 decessi in tutto il mondo, 5876 ricoveri e infortuni gravi, tutti a seguito di tentativi artigianali di curare il Coronavirus (per esempio 60 persone hanno perso la vista dopo aver bevuto metanolo, come suggerito in un articolo online). Come se non bastasse, uno studio del King's College di Londra ha dimostrato una relazione negativa tra la credenza a teorie complottiste sul virus e il rispetto delle linee guida per la salute pubblica.

È relativamente facile guadagnare con le fake news: un alto numero di condivisioni e interazioni ingannano gli algoritmi dei principali social, che le categorizzano come contenuti di tendenza e li propongono a un pubblico più vasto, andando così ad aumentare ulteriormente il numero di visualizzazioni sulle loro pagine. Possono quindi richiedere prezzi più alti agli inserzionisti, che spesso utilizzano delle aste online per accaparrarsi gli spazi pubblicitari migliori, affidandosi però anche in questo caso ad algoritmi che non riescono a distinguere un contenuto vero da uno falso, e rischiando così di gettare discredito sulla propria azienda.

Proprio per tutelare i cittadini, le piattaforme online hanno messo in atto varie misure volte a contenere la diffusione dell'*infodemia*: Whatsapp ha ridotto il numero di inoltri consentiti, Google ha introdotto dei link alle pagine delle autorità sanitarie quando è effettuata una ricerca legata al Coronavirus e i principali social network hanno sviluppato meccanismi volti a rimuovere al più presto dopo la pubblicazione quei contenuti che possono essere pericolosi.

1.4 I cyber attacchi

Con l'aumento della fruizione tecnologica è aumentata anche l'esposizione al cybercrime.

Uno dei principali motivi per cui i cyber-criminali effettuano il maggior numero di attacchi durante le situazioni più critiche è sicuramente lo stato emotivamente instabile delle vittime, che diventano quindi facili prede.

È stato stimato che le minacce di cybersecurity costeranno al mondo, in tutto, circa 6 trilioni di dollari nel 2021. Il numero di attacchi è cresciuto di 5 volte dopo la diffusione dei primi casi di COVID-19 secondo l'OMS. Nel mirino degli hacker non solo i singoli utenti, ma anche e soprattutto le strutture sanitarie, che già da qualche anno sono state sottoposte a più data breach di qualunque industria finanziaria (C.M. Williams et al. [2020]). I database delle strutture sanitarie sono particolarmente importanti perché contengono informazioni sensibili sui pazienti, che possono essere rivenduti sul dark web a 10-20 volte tanto rispetto all'ammontare presente sulla carta di credito di un utente.

Allo stesso modo, anche le università e gli enti principali di ricerca sono stati sottoposti a molti attacchi. Il principale mezzo sono state le mail di phishing contenenti parole come "WHO", "Vaccino" o "Donazione", che avevano al loro interno un malware oppure estraevano le credenziali dello user.

Il caso più eclatante è sicuramente l'attacco ransomware (un codice malevolo che blocca il pc della vittima e richiede un riscatto per sbloccarlo) all'ospedale di Dusseldorf, che ha portato al decesso di una paziente a causa del ritardo delle cure.

1.5 L'ICT in periodo COVID-19 in Italia

Siccome l'analisi sviluppata e descritta nei capitoli successivi utilizzerà dati italiani, è necessario, infine, un focus sul settore dell'ICT in Italia.

Il ruolo dell'AGCOM nella gestione dell'emergenza sanitaria nazionale

L'Italia non è sicuramente in controtendenza rispetto al resto del mondo: l'utilizzo della rete è aumentato in breve tempo e l'Autorità per le Garanzie nelle Comunicazioni si è occupata di permettere al più vasto numero possibile di utenti l'accesso ai servizi e alle reti di comunicazione. Per fare ciò, sono state adottate soluzioni temporanee per evitare il rischio di congestione della rete e, allo stesso tempo, sono stati istituiti incentivi

regolamentari per favorire connessioni stabili e di qualità attraverso gli investimenti in banda ultralarga.

Sono stati predisposti 4 Tavoli, uno per ogni area, al fine di permettere una mediata comunicazione tra i vari stakeholder: comunicazioni elettroniche, servizi postali, media tradizionali e piattaforme online. Sono stati trattati vari temi, ma i più importanti emersi sono stati sicuramente l'accesso alle reti, il potenziamento in ambito sicurezza, la consegna di pacchi per l'e-commerce e il controllo sui fornitori di servizi di media audiovisivi.

L'ecosistema delle comunicazioni

È ormai un dato di fatto che la pandemia abbia causato una grave flessione del Prodotto Interno Lordo (del 12,4% nel secondo trimestre 2020) e dei consumi (11,5%), aumentando del 18,6 % la propensione al risparmio delle famiglie.

In questo scenario anche l'ecosistema delle comunicazioni ha risentito delle conseguenze: solo nel primo trimestre il suo valore risulta inferiore quasi del 6% rispetto a quello del 2019 in tre dei suoi macrosettori (Telecomunicazioni, Media e Servizi Postali), presentando invece una controtendenza riguardo ai *VOD* – contenuti audiovisivi online a pagamento –, che hanno registrato una crescita più che rilevante: solo a marzo +60% rispetto a marzo 2019.

Il settore dei media è stato fortemente impattato dal netto decremento della vendita degli spazi pubblicitari, che da sempre costituiscono un importante introito, conseguenza sia della diminuzione dei consumi delle famiglie sia del blocco delle attività produttive.

Le telecomunicazioni hanno riscontrato una contrazione inferiore rispetto al generale andamento dell'economia grazie alla maggior domanda di connettività, che però è stata mediata dall'impatto negativo del PIL sul reddito degli individui e sui nuovi investimenti.

Sulla base degli andamenti appena descritti, è stato stimato che alla fine del 2020 la perdita complessiva nel settore di comunicazioni potrebbe essere compresa tra i 3 e i 5 miliardi (ossia tra il -6% e il -10% rispetto al 2019).

Informazione e disinformazione

Come già visto precedentemente in questo periodo è fondamentale disporre di un'informazione corretta e oggettiva e per tale motivo, con la delibera n. 129/20/CONS, i fornitori di servizi media sono stati richiamati al fine di garantirla.

Come nel resto del mondo è aumentata la domanda in tale ambito, tanto da portare il 50% degli ascoltatori nella fascia oraria delle 20 a seguire i principali telegiornali nazionali e locali. Inoltre, una notevole quantità di utenti unici ha scaricato app di notizie e ha sottoscritto abbonamenti premium di editoriali online (+7% rispetto a marzo 2019).

Seppur meno colpita di altri Paesi, anche l'Italia è stata vittima dell'*infodemia*: più del 30% di utenti Italiani ha consultato siti di disinformazione (+11% rispetto a marzo 2019). Proprio per questo motivo l'Autorità ha messo a punto un sistema di monitoraggio continuativo della qualità dell'informazione, oltre ad aver instaurato un confronto con i gestori delle piattaforme online dei gruppi Facebook, Google, Amazon e Bytedance per stabilire una linea comune d'azione volta a fronteggiare la disinformazione.

Rete e infrastrutture

Con l'aumento della fruizione di Internet, la pressione sulle infrastrutture è stata senza precedenti (durante le settimane di lockdown il traffico dati è aumentato del 75% rispetto allo stesso periodo nel 2019) e ha portato di conseguenza a un rallentamento delle prestazioni in termini di connessione, passando, in download, dai 61 Mbps standard ai 56 Mbps. Solo durante la prima settimana c'è stata una crescita del 33% nel traffico in rete (tra i maggiori in Europa), incremento che secondo Paolo Campoli, leader del segmento di Service Provider di Cisco Emea, in tempi normali si sarebbe verificato in un anno o un anno e mezzo. Il chief technology & information officer di Tim, Michele Gamberini, afferma che l'azienda ha riportato aumenti dell'80% sulla linea fissa e del 40% sulla linea mobile, soprattutto nella fascia oraria dalle 9:30 alle 16:30, di solito poco congestionata. Per completare il quadro, Vodafone segnala il ritorno del traffico voce, con le telefonate in crescita del 40% e Fastweb registra il boom (+300%) nei giochi online. In questo scenario si è rivelata importantissima l'azione effettuata dai principali operatori di streaming: Netflix, Prime Video e Youtube hanno diminuito il bit-rate, fornendo quindi agli utenti immagini in minor definizione, ma alleggerendo di molto la congestione della rete.

Il digital divide

Tre sono i principali fattori di discriminazione in ambito tecnologico-digitale in Italia: età, reddito ed educazione. Nel 2019, infatti, il 33,8% delle famiglie non possedeva un computer o un tablet e la percentuale saliva al 70,6% nelle famiglie composte di soli anziani e scendeva al 14,3% nelle famiglie con almeno un minorenni. Per quanto riguarda il livello di istruzione, solo il 7,7% delle famiglie in cui vi sia almeno un laureato non ha un computer in casa. Inoltre è determinante il fattore reddito, discriminante principale a causa dei costi talvolta elevati per una connessione a Internet, che non rientra tra i beni di prima necessità.

A differenza del problema degli Stati Uniti, in cui è difficile fornire una copertura accettabile per tutto il territorio, in Italia la copertura territoriale consentirebbe all'88,9% delle famiglie l'accesso a Internet con velocità maggiori o uguali a 30 Mbps, tuttavia è solo il 37,2% di esse che possiede una simile connessione, fatto particolarmente evidente nelle regioni del sud Italia. Quindi, per avere una crescita in digitalizzazione del Paese sarà necessario, oltre a pianificare investimenti infrastrutturali mirati (*technology push*), stimolare la domanda, rappresentata dalla capacità di spesa delle famiglie, più che dalla necessità di una connessione (*demand pull*).

Non è ancora stata colmata del tutto, però, la differenza tra aree urbane e rurali, e in particolare le aree di montagna, in cui spesso la copertura è ben al di sotto della media, se non del tutto inesistente.

Le piattaforme online

Le piattaforme online costituiscono un unicum nello scenario della tele-comunicazione, guadagnando tra gennaio e luglio 700 mila sottoscrittori secondo EY: Netflix, Prime Video e anche il nuovo Disney+ sembrano essere gli unici ad aver beneficiato di questo periodo.

I tassi di crescita sono ormai da tempo straordinari per i *Vod* (video on demand), che possono raccogliere enormi quantità di dati sui propri utenti, al fine di migliorare le indagini sulla clientela e migliorare i processi di produzione e fruizione. I dati sono ormai fonte di vantaggio competitivo non replicabile e i settori di attività dove esse operano sono caratterizzati da economie di scala, costi fissi e affondati, basso multi-homing ed elevati switching cost. Tutto ciò porta, indubbiamente, a un aumento del livello di concentrazione,

tendente a una situazione di monopolio. Tale rischio va ad aggiungersi al problema ormai noto della tassazione digitale, per cui l’Autorità sta da tempo studiando regole da azionare per garantire la loro piena *accountability* sul territorio nazionale.

Capitolo 2: Analisi descrittiva

Vista la complessità del dataset utilizzato, prima di effettuare l'analisi econometrica, è stata eseguita un'analisi descrittiva, in modo da ottenere una visione più dettagliata dei dati in possesso.

In un primo momento sono state prodotte alcune statistiche generali rappresentanti gli andamenti dei consumi medi giornalieri per diverse fonti di traffico. Utilizzando come riferimento le date più importanti nell'orizzonte temporale in esame, si è così potuto ottenere evidenza di come non tutte le categorie di traffico presentassero un legame con la pandemia.

Per raffinare l'analisi, si è proceduto con l'identificazione di quali, tra le province italiane, fossero state quelle colpite in modo più o meno rilevante in termini di casi positivi registrati di Covid-19. Tale raggruppamento è stato propedeutico per l'ultima parte, in cui sono stati studiati i diversi andamenti dei consumi medi giornalieri pro capiti nelle zone più e meno colpite identificate.

2.1 Il dataset

Il dataset principale utilizzato è stato fornito al Politecnico di Torino dalla società TIM SpA e conteneva i dati italiani del traffico giornaliero, tra il 16 aprile 2020 e il 30 novembre 2020, con una granularità a livello comunale.

Vista la sua posizione predominante nel panorama dei fornitori di rete telefonica, tali dati possono essere considerati come una buona approssimazione dell'andamento su scala nazionale: nel suo quarto Osservatorio sulle Comunicazioni del 2020, l'AGCOM ha pubblicato i dati relativi alle quote di mercato di tutti i principali operatori di telefonia e TIM SpA risulta essere un chiaro leader di settore. Su rete fissa, infatti, è indiscutibile la sua posizione predominante, che seppure in calo di un punto percentuale rispetto al 2019, continua a mantenere un netto distacco dalle altre, con una quota di mercato del 42,1% (Grafico 1).

Diversa è la situazione sulla rete mobile, dove il primo posto è da tempo conteso tra TIM e Vodafone: secondo quanto riportato, a settembre 2020 Vodafone prevale su TIM con una differenza di 0,1 punti percentuali (Grafico 2). Analizzando, però, i precedenti Osservatori pubblicati si può notare come questa differenza non sia così rilevante perché le due società si alternano nel ricoprire la prima posizione.

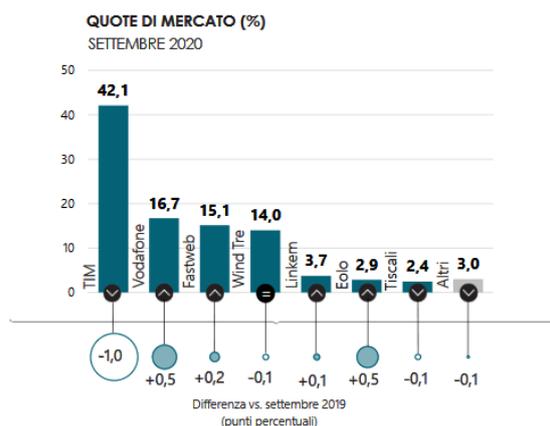


Grafico 1 - Quote di mercato della telefonia fissa.
Fonte: AGCOM

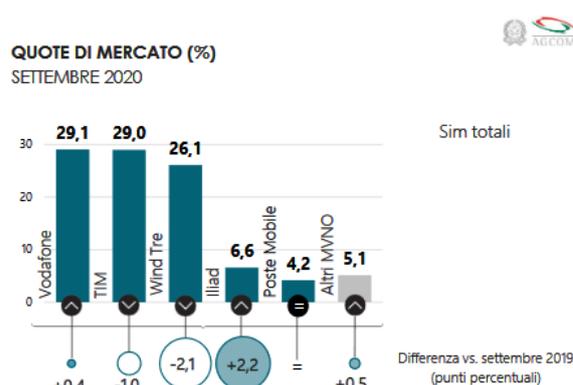


Grafico 2 - Quote di mercato della telefonia mobile
Fonte: AGCOM

Struttura del dataset

Il dataset iniziale presentava un totale di 5.445.787 record e 41 variabili, così suddivise:

- **Date:** data compresa tra il 16 aprile 2020 e il 30 novembre 2020;
- **Atype:** la tipologia di connessione, suddivisa tra fonti di traffico da rete fissa (ADSL e FTTC) e da rete mobile (2G, 3G, 4G);
- **Comune, Provincia e Regione** in cui è stata effettuata la rilevazione. Nel dataset erano presenti 7233 comuni distinti, 107 province o unità territoriali e 20 regioni;
- **ISTAT:** codice Istat del comune;
- **User type:** ulteriore differenziazione per il traffico su rete mobile. Sono presenti due tipologie di utenti: gli utenti sedentari (S) e quelli di passaggio (P), in base al fatto che abbiano o meno utilizzato Internet al di fuori del comune di residenza. Tutti i dati della linea fissa sono invece identificati come utenti sedentari;

- Ognuna delle 8 fonti di traffico (*Streaming, Gaming, Social Network, Instant Messaging, Browsing, Miscellaneous, Cloud, Software Update*) era suddivisa su 4 variabili: **rx** (il traffico in ricezione in Bps), **tx** (il traffico in trasmissione in Bps), **du** (il numero di utenti del campione della categoria), **tm** (il tempo di utilizzo della categoria).

Trattandosi di un database di grandi dimensioni, la prima operazione effettuata è stata quella di *data preprocessing* e conseguente *data cleaning*: il dataset iniziale conteneva il traffico di alcune città senza riportarne il codice Istat associato. Per tali record, quindi, mancava l'informazione relativa alla provincia di appartenenza, oppure l'associazione fatta risultava errata e, per tale motivo, il conseguente esito delle analisi avrebbe potuto essere falsato.

Per ovviare a questo problema è stato estratto il database aggiornato di tutti i comuni e le frazioni italiane presente sul sito dell'Istat, così strutturato:

- **Codice Regione:** codice della regione di appartenenza;
- **Codice dell'Unità territoriale sovracomunale:** codice dell'unità territoriale a livello intermedio tra il Comune e la Regione;
- **Codice della Provincia:** codice della provincia di appartenenza;
- **Progressivo del comune:** valore alfanumerico di tre caratteri associato a ogni comune all'interno della provincia di appartenenza;
- **Codice Comune formato alfanumerico:** codice del comune;
- **Denominazione italiana e straniera;**
- **Ripartizione geografica:** suddivisione del territorio italiano in Nord-ovest, Nord-est, Centro, Sud e Isole;
- **Denominazione Regione;**
- **Denominazione dell'Unità territoriale e sovracomunale:** denominazione della provincia, città metropolitana, libero consorzio di comuni o unità non amministrative;
- **Tipologia di Unità territoriale:** codice indicante la tipologia di unità territoriale (provincia, provincia autonoma, città metropolitana, libero consorzio di comuni, unità non amministrativa);

- **Flag Comune capoluogo di provincia/città:** valore booleano per indicare se il comune sia o meno capoluogo;
- **Sigla automobilistica;**
- **Codici comunali in diversi formati in anni precedenti al 2020;**
- **Codici NUTS1, NUTS2, NUTS3 del 2019 e del 2020:** nomenclature delle unità territoriali per varie statistiche.

Sono state estratte le informazioni relative ai codici regionali, provinciali e comunali, alle denominazioni, alla ripartizione geografica e alla sigla automobilistica, che hanno costituito un nuovo dataset utilizzato per effettuare un join, tramite query SQL, con il dataset iniziale fornito da Tim SpA, per ottenere così le informazioni mancanti.

Tipologie di connessione

Come detto in precedenza, le due tipologie di connessioni fisse disponibili nel dataset erano ADSL e FTTC: l'ADSL (Asymmetric Digital Subscriber Line) è stata per molto tempo la tecnologia più diffusa in Italia, principalmente grazie all' utilizzo della stessa infrastruttura in rame impiegata dalla linea telefonica. La definizione di "asimmetrica" è dovuta al fatto che fornisce due diverse velocità: in download può arrivare fino a un massimo di 20 Mbps e in upload può raggiungere 1 Mbps.

Diverso è il concetto alla base dell'FTTC (Fiber To The Cabinet), invece, che fa affidamento a un misto di cavi in fibra ottica e in rame, combinazione che permette di raggiungere una velocità massima di 200 Mbps in download e 20 Mbps in upload.

Sono stati analizzati gli andamenti dei valori di traffico medio giornaliero per utente per ciascuna delle otto tipologie di fonti di traffico e in tutti i casi il consumo in FTTC superava quello in ADSL (si riporta a titolo di esempio l'andamento nella tipologia *Streaming* in Grafico 3).

Di seguito è mostrato il codice utilizzato sull'editor di Power Query per questo passaggio. Per evitare ridondanza, si riporta solamente la tipologia *Streaming*, ma analogo procedimento è stato utilizzato per tutte le altre fonti di traffico.

```

#"Filtrate righe" = Table.SelectRows([...], each [atype] = "ADSL" or [atype] =
"FTTC"),
#"Raggruppate righe" = Table.Group([...], {"date", "atype"}, {"Dw streaming", each
List.Sum([rx_streaming]), type nullable number}, {"Up streaming", each
List.Sum([tx_streaming]), type nullable number}, {"Ut streaming", each
List.Sum([du_streaming]), [...]}),
#"Aggiunta colonna personalizzata" = Table.AddColumn("#Raggruppate righe",
"Streaming", each ([Dw streaming]+[Up streaming])/[Ut streaming])

```

L'idea alla base del codice è stata quella di selezionare solo i record con tipologia di rete ADSL o FTTC, e poi procedere, per ogni fonte di traffico, con il raggruppamento a livello di data e di tipologia di rete fissa, sommando, distintamente, tutto il traffico in ricezione (rx), in trasmissione (tx) e il totale degli utenti (ut). Nell'ultimo passaggio è stata creata una nuova colonna, costituita dal traffico medio giornaliero, in cui si è calcolato, sempre mantenendo lo stesso criterio di raggruppamento, un rapporto tra la somma di download e upload e la somma degli utenti calcolati allo step precedente.

Il consumo medio giornaliero per utente ottenuto dalla query può essere rappresentato nella seguente formula:

$$\text{Traffico medio giornaliero}_{data,a,t} = \frac{\sum_{i=1}^c rx_{i,data,a,t} + \sum_{i=1}^c tx_{i,data,a,t}}{\sum_{i=1}^c ut_{i,data,a,t}} \quad (2.1)$$

in cui *data*, *a* e *t* rappresentano, rispettivamente la data, il tipo di connessione fissa e la tipologia di traffico in esame. Per ogni *data*, *a* e *t*, sono sommati i valori di $rx_{data,a,t}$ – il traffico in download – e di $tx_{data,a,t}$ – il traffico in upload – e il tutto è diviso per la somma del numero degli utenti rilevati. *C* rappresenta l'unità territoriale base, ossia i comuni: le varie sommatorie vengono infatti effettuate su tutti i comuni presenti.

Come evidenziato anche dal quarto Osservatorio dell'AGCOM [2020], tale risultato non deve stupire: per la prima volta le linee FTTC superano quelle in rame con rispettivamente 43,9 e 41,4 punti percentuali sul totale delle reti fisse. Ciò è sicuramente dovuto alle migliori prestazioni garantite, che permettono di ottenere velocità superiori e tempi più rapidi, implicando un consumo mediamente maggiore: permettono, infatti, di poter usufruire correttamente di tutte le funzionalità *HD*, che spesso richiedono una connessione di banda molto più stabile di quella offerta dalle tecnologie ADSL.

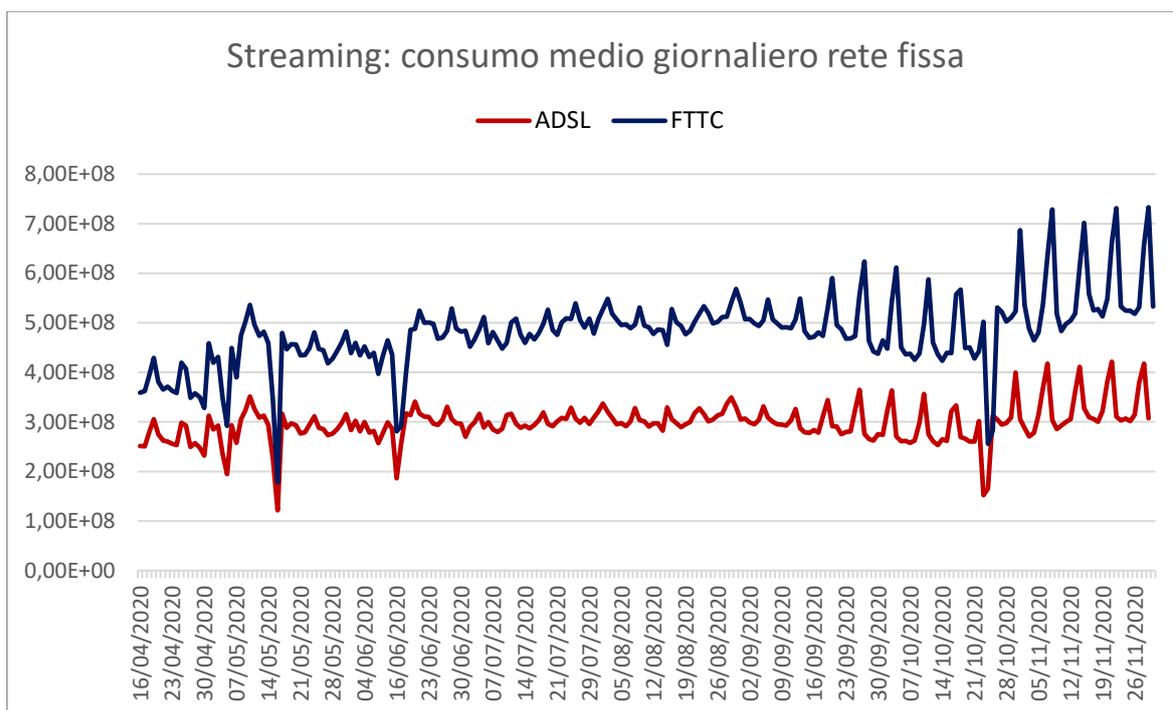


Grafico 3 – Streaming: traffico medio giornaliero su rete fissa

La rete mobile, invece, era suddivisa in tre tipologie: 2G, 3G e 4G. Anche in questo caso sono stati analizzati gli andamenti temporali del traffico medio giornaliero per le otto fonti disponibili, ma in tutte le tipologie è stata evidente la predominanza delle reti 4G (per il calcolo si può far riferimento alla (2.1), con l'unica differenza nella selezione iniziale, in cui Atype è stato posto uguale a 2G, 3G o 4G). Si riportano di seguito due grafici che mostrano, rispettivamente, l'andamento temporale (Grafico 4) e la ripartizione rispetto al totale (Grafico 5). Entrambi fanno riferimento al traffico medio giornaliero di Streaming, ma analoghi risultati sono stati ottenuti in tutte le altre tipologie.

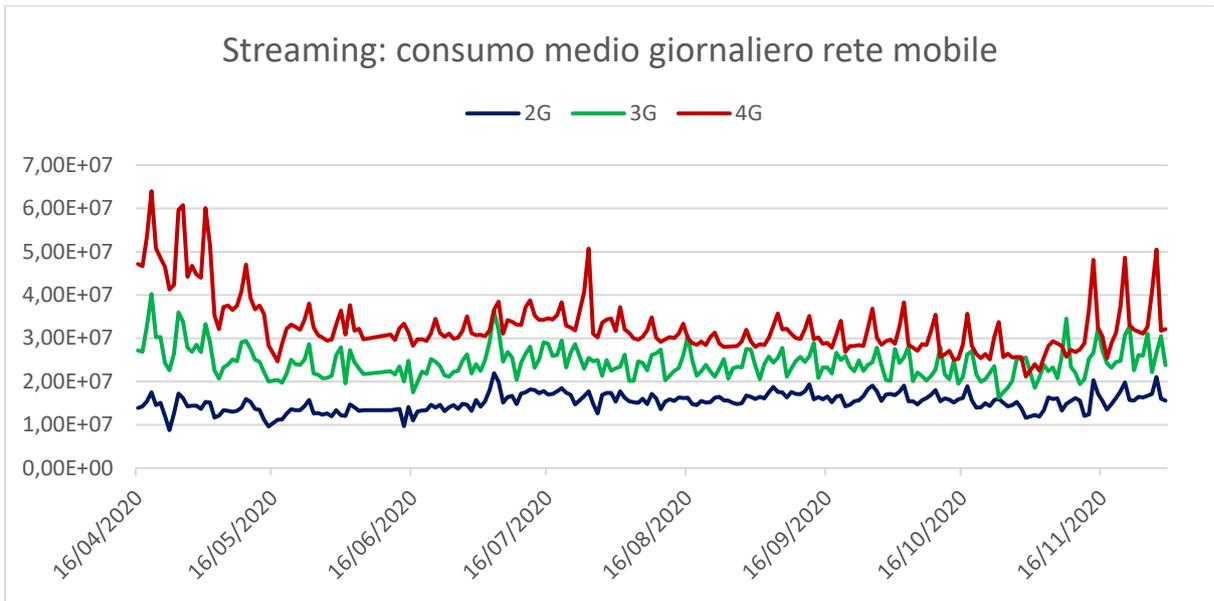


Grafico 4 - Consumo medio giornaliero di Streaming per tipologia

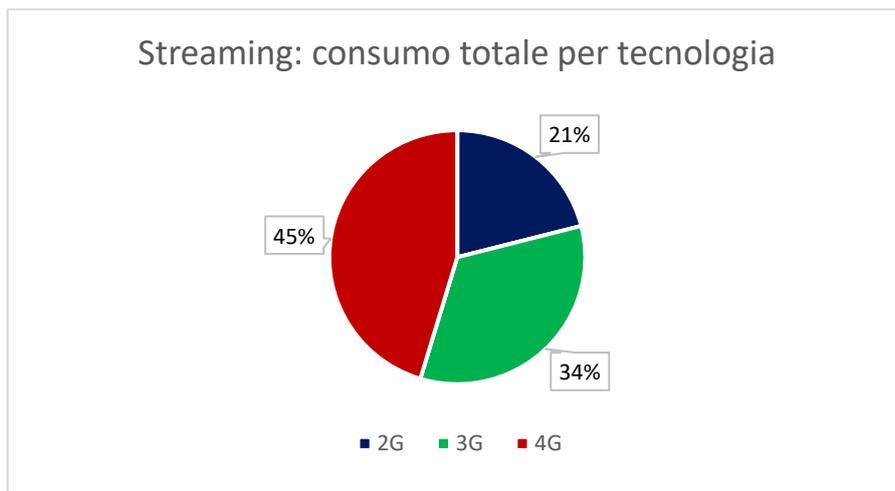


Grafico 5 - Streaming: consumo totale per tecnologia

Il motivo di questa differenza, che porta il 4G (4th generation) ad avere un consumo equivalente quasi alla metà del traffico totale, è dovuto alle sue caratteristiche intrinseche: basato su Protocolli Internet (IP), è il primo standard progettato prevalentemente per favorire le connessioni dati e offrire fino a 100 Mbps per l'accesso mobile.

Il 2G, al contrario, è una rete che era stata pensata per le comunicazioni vocali, offrendo una velocità massima di 250 Kbps; il 3G, invece, è stata una transizione tra le due, offrendo

discrete prestazioni sull'utilizzo sia di voce sia di dati, arrivando a garantire una velocità di 7 Mbps.

Non c'è da stupirsi quindi che il 4G sia la tecnologia prediletta, anche se il consumo delle altre tipologie è probabilmente dovuto a paesi più piccoli (che mostrano infatti un maggior consumo di 2G e 3G) rispetto ai grandi centri abitati: questo è dovuto al divario infrastrutturale che ancora oggi permane e di cui si rimanda a quanto riportato nel primo capitolo.

2.2 Andamenti temporali delle otto tipologie di traffico

A partire dai primi contagi del mese di febbraio 2020, in Italia sono stati individuati tre periodi temporali distinti: il primo (anche detta *prima ondata*), da febbraio a maggio 2020, è stato caratterizzato da un elevato numero di contagi e decessi principalmente nelle zone del Nord Italia; il secondo, definito *di transizione*, da maggio a metà settembre, in cui si è verificato un arresto nella diffusione dei contagi e una netta diminuzione nel numero dei decessi; il terzo, la cosiddetta *seconda ondata*, che ha visto nuovamente impennare la curva dei contagi e il numero dei decessi con una flessione nella pendenza rilevata solamente a fine novembre.

I tre periodi sopra elencati sono stati caratterizzati da obblighi e restrizioni più o meno stringenti e per tale motivo si è deciso di procedere nell'analisi indagando l'andamento temporale del consumo delle varie fonti di traffico, ponendo maggior interesse su alcune date importanti nell'orizzonte temporale in esame, quali:

- **4 maggio 2020:** con l'inizio della *fase 2*, è stato possibile il rientro a lavoro per alcune categorie di occupati ed è stato consentito spostarsi per incontrare i "congiunti";
- **18 maggio 2020:** è stata la data simbolo per la fine del lockdown nazionale incominciato a marzo. La maggior parte delle filiere produttive ha potuto riaprire, così come i bar e i ristoranti, che hanno ripreso a effettuare anche servizio al tavolo e non più solo l'asporto. Non era più necessaria l'autocertificazione per spostarsi ed era consentito incontrare una persona al di fuori del proprio nucleo familiare;
- **1 giugno 2020:** riapertura di tutte le attività commerciali;

- **15 giugno 2020:** inizio della *fase 3*, comprendente, tra le altre riaperture, anche quelle di aree giochi, cinema e teatri all'aperto;
- **13 ottobre 2020:** visto il nuovo incremento dei contagi, il Presidente del Consiglio Giuseppe Conte ha firmato il primo DPCM di quella che sarà in seguito definita *seconda ondata*, introducendo le prime restrizioni quali il limite al numero di persone durante cerimonie civili e religiose e il divieto di praticare sport di contatto se non in competizioni riconosciute dal Comitato Olimpico Nazionale Italiano (CONI);
- **24 ottobre 2020:** firma del DPCM che ha imposto la chiusura di palestre e piscine e incrementato la Didattica a Distanza e lo Smart Working;
- **3 novembre 2020:** firma del DPCM che ha previsto l'istituzione del coprifuoco nazionale dalle 22 alle 5, la DAD obbligatoria per le scuole superiori, la chiusura dei centri commerciali nel weekend e la riduzione al 50% della capacità dei mezzi. Inoltre, le regioni sono state suddivise in diverse categorie (zona rossa, arancione o gialla) a seconda del grado di criticità del livello dei contagi e all'indice R_0 , che l'Istituto Superiore di Sanità ha definito come "*il numero medio di infezioni secondarie prodotte da ciascun individuo infetto in una popolazione completamente suscettibile cioè mai venuta a contatto con il nuovo patogeno emergente*".

Rete fissa

Per ciascuna delle fonti di traffico presenti nel dataset (*Streaming, Gaming, Social Network, Instant Messaging, Miscellaneous, Browsing, Software Update e Cloud*) è stato preso in esame il consumo totale medio giornaliero in un modo simile a quanto fatto in precedenza, con l'unica differenza che il raggruppamento qui non è più stato fatto congiuntamente su data e tipologia di rete, ma solamente sulla data:

```

#"Filtrate righe" = Table.SelectRows([...], each [atype] = "ADSL" or [atype] =
"FTTC"),
#"Raggruppate righe" = Table.Group([...], {"date"}, {"Dw streaming", each
List.Sum([rx_streaming]), type nullable number}, {"Up streaming", each
List.Sum([tx_streaming]), type nullable number}, {"Ut streaming", each
List.Sum([du_streaming]), [...]}),

#"Aggiunta colonna personalizzata" = Table.AddColumn("#Raggruppate righe",
"Streaming", each ([Dw streaming]+[Up streaming])/[Ut streaming])

```

Quindi riassumibile nella formula:

$$\text{Traffico medio giornaliero}_{data,t} = \frac{\sum_{i=1}^c rx_{i,data,t} + \sum_{i=1}^c tx_{i,data,t}}{\sum_{i=1}^c ut_{i,data,t}} \quad (2.2)$$

Tra tutte le fonti in esame, solamente Social Network e Instant Messaging (rispettivamente Grafico 6 e Grafico 7) hanno evidenziato andamenti legati all'evoluzione dell'epidemia: il consumo è stato maggiore in quei periodi in cui le restrizioni imponevano di restare a casa (prima dell'inizio della fase 2) e successivamente all'inasprimento delle regole nella seconda ondata, seppur con una crescita più moderata, lontana dai livelli di aprile-maggio. Quest'ultima osservazione potrebbe essere ricondotta al diverso orizzonte geografico su cui sono state applicate le restrizioni: mentre per la prima fase il lockdown è stato effettuato a livello nazionale, nella seconda ondata a ogni regione sono state attuate norme più o meno stringenti in base all'andamento del numero di contagi.

Durante i mesi estivi, invece, la cosiddetta fase di transizione, il consumo è in entrambi i casi più basso.

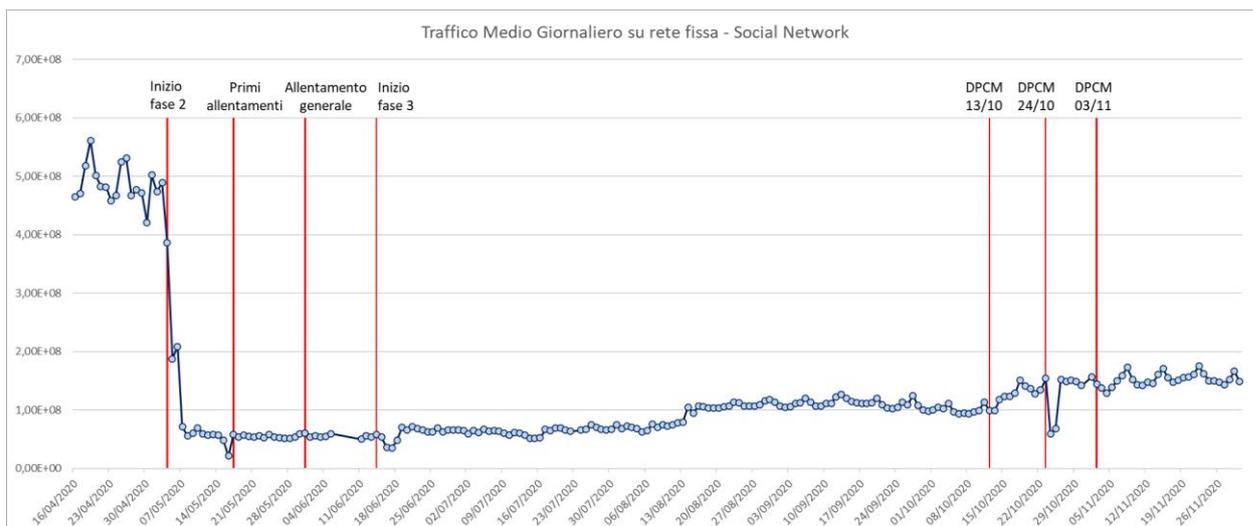


Grafico 6 – Social Network: traffico medio giornaliero su rete fissa

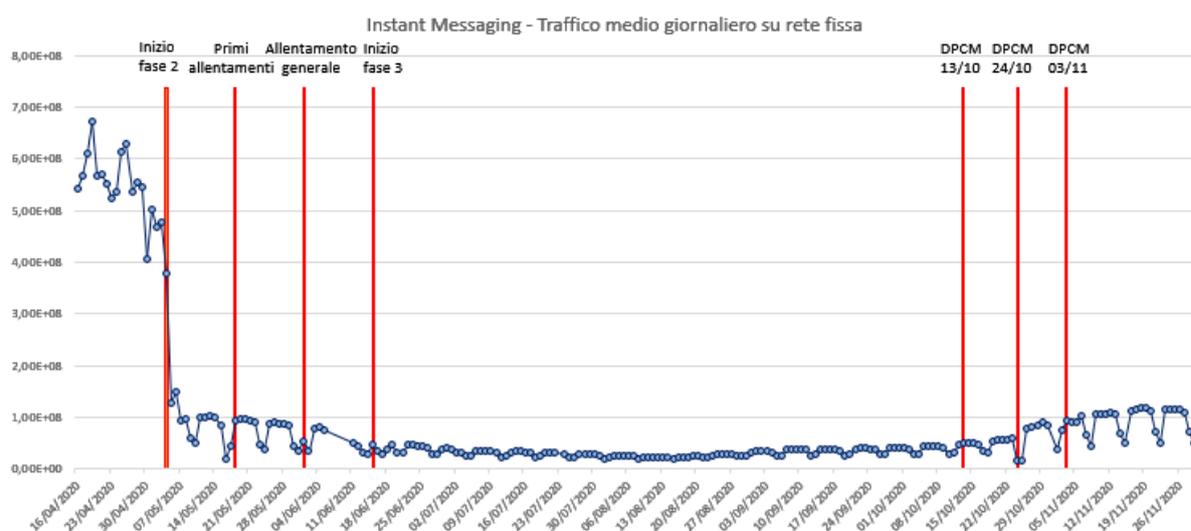


Grafico 7 – Instant Messaging: traffico medio giornaliero su rete fissa

In tutte le altre tipologie non sono emersi trend particolarmente rilevanti. Gli andamenti sembrano infatti essere costanti o addirittura opposti rispetto a quanto visto per le categorie precedenti. Per esempio, il traffico in Gaming (Grafico 8) aumenta a partire dall'inizio della *fase 2* e i valori registrati durante il lockdown nazionale sono i più bassi dell'orizzonte temporale in esame.

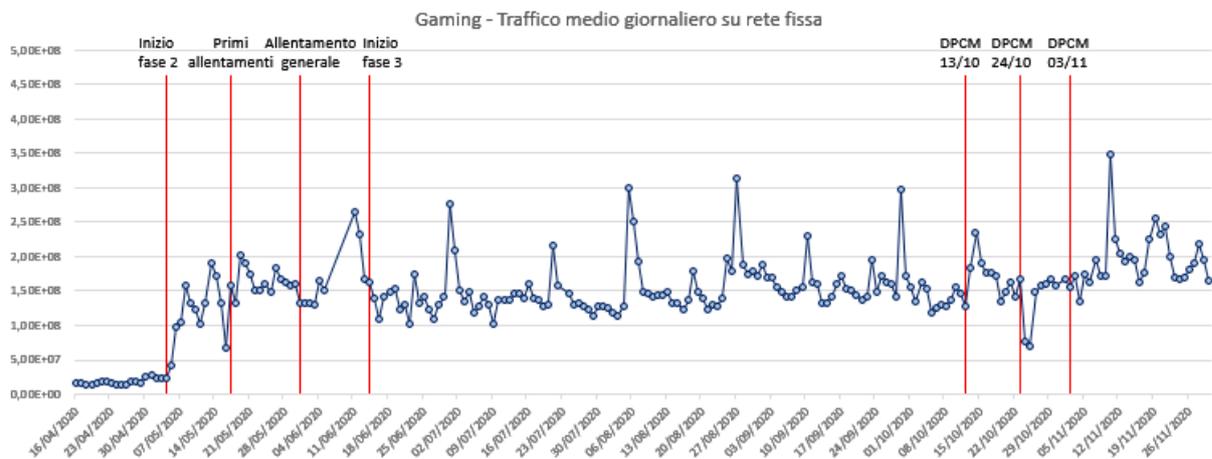


Grafico 8 – Gaming: traffico medio giornaliero su rete fissa

Rete mobile

L’analisi è stata ripetuta per la rete mobile, anche in questo caso con l’accortezza di cambiare il filtro iniziale sui record che presentassero la variabile Atype uguale a 2G, 3G o 4G. È stata aggiunta una differenziazione che sul fisso non poteva essere fatta: anziché considerare il traffico totale, si è deciso di analizzare distintamente il consumo medio giornaliero di un utente *sedentario* e quello di un utente *di passaggio*, effettuando così un confronto fra i due.

Per fare ciò, è stata utilizzata una query che producesse come output il dataset raggruppato a livello di data e di *User Type*, e per entrambe le tipologie, S e P, calcolasse il traffico medio giornaliero con la stessa formula vista in precedenza (2.2).

Diversamente da quanto rilevato per il fisso, in questo caso quattro fonti di traffico hanno evidenziato un trend legato all’andamento della pandemia: Streaming (Grafico 9), Social Network (Grafico 10), Instant Messaging (Grafico 11) e Browsing (Grafico 12).

Le due linee seguono sempre lo stesso andamento, nonostante sia chiaro un consumo medio maggiore per gli utenti *sedentari* rispetto a quelli *di passaggio*. In generale si può affermare che i risultati ottenuti siano simili a quelli già analizzati con la rete fissa: i consumi sono maggiori nei periodi più significativi per la pandemia e diminuiscono durante l’estate.

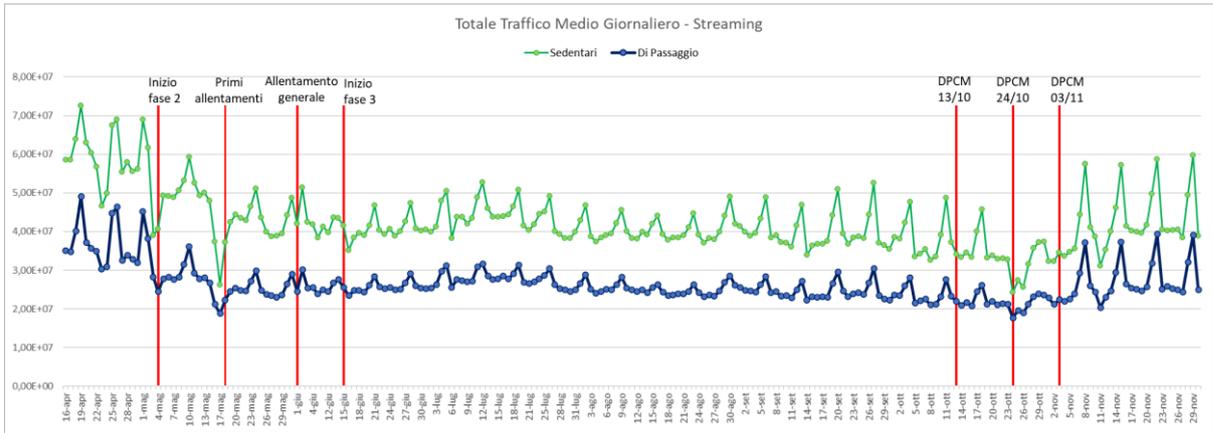


Gráfico 9 – Streaming: totale traffico medio giornaliero su rete mobile

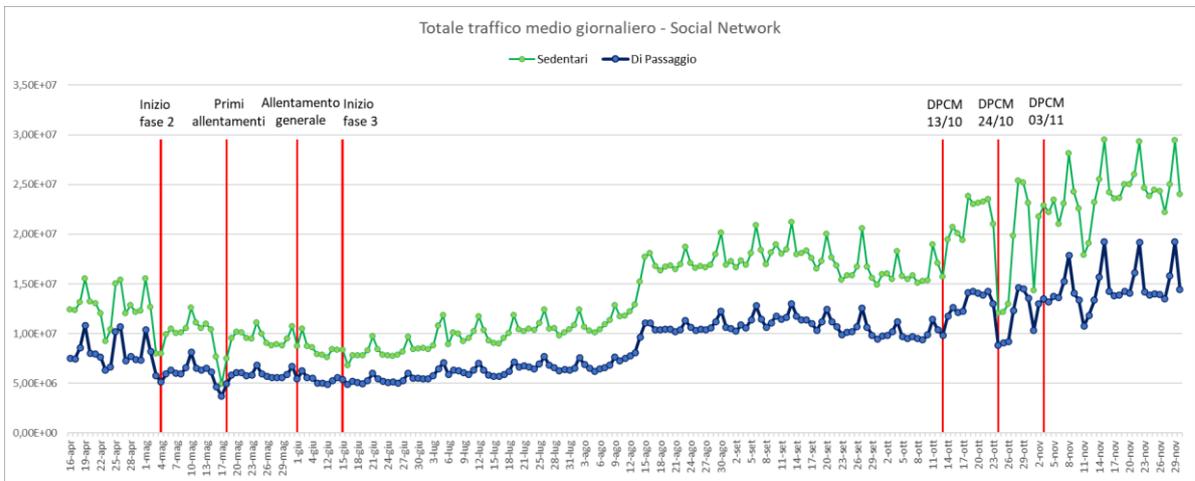


Gráfico 10 – Social Network: totale traffico medio giornaliero su rete mobile

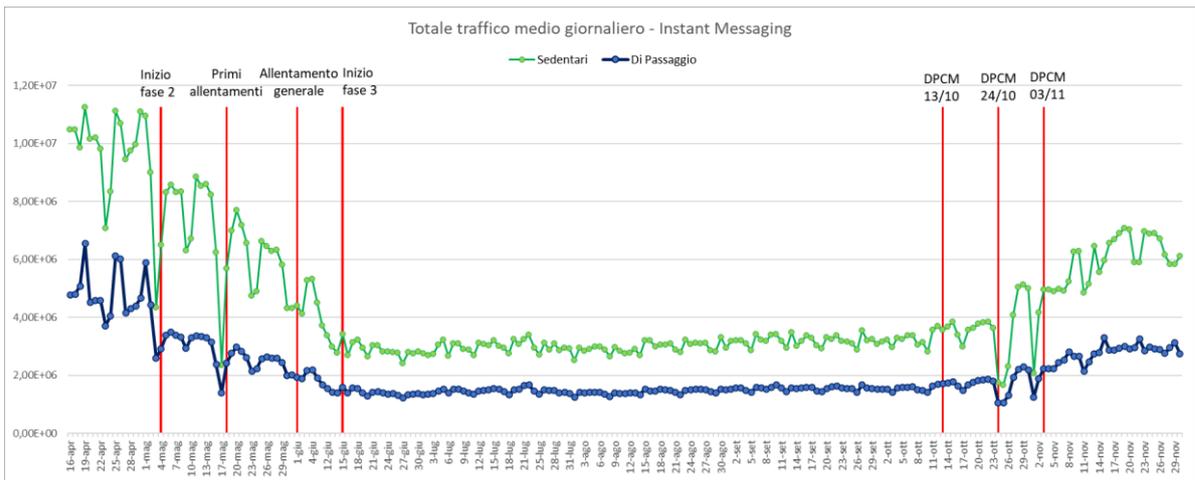


Gráfico 11 – Instant Messaging: traffico medio giornaliero su rete mobile

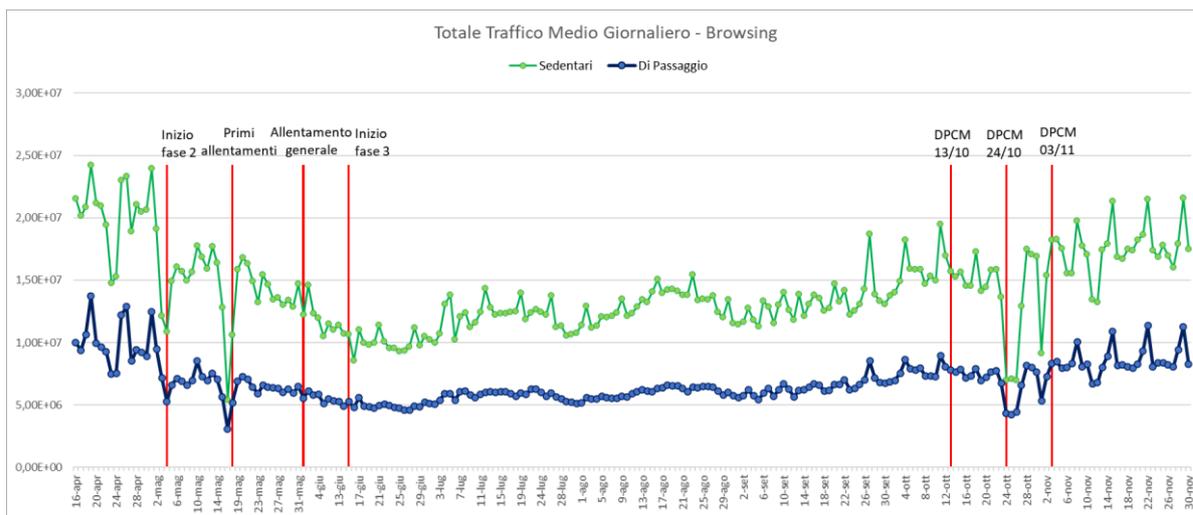


Grafico 12 – Browsing: traffico medio giornaliero su rete mobile

È necessaria una precisazione per quanto riguarda l'andamento del traffico su Social Network, sopra riportato nel *Grafico 10*. Come si può notare, il traffico medio giornaliero incomincia a crescere dal 15 agosto in poi, superando nettamente i livelli raggiunti nelle fasi delle maggiori restrizioni.

Per tutte le fonti si è effettuata anche un'analisi separatamente per traffico medio in ricezione e traffico medio in trasmissione e in questo caso si è rivelato particolarmente utile per indagarne meglio l'andamento atipico. Come si può notare dai due grafici di seguito riportati (*Grafico 13* e *Grafico 14*), l'andamento è dominato più dal traffico in download che da quello in upload, che presenta sempre valori minori. L'aumento di consumo a partire dalla fine dell'estate può probabilmente essere conseguenza dell'aggiornamento di Instagram del 6 agosto 2020: sono stati introdotti i *reels*, ossia dei brevi video simili a quelli presenti su TikTok.

A conferma di questa ipotesi, la pubblicazione "Italiani e Social Media" di Blogmeter – istituto di ricerca italiano per i servizi di Integrated Social Intelligence – riporta al terzo posto nella classifica dei social network più utilizzati in Italia, Instagram, secondo solo a Facebook e a Youtube. Per tale motivo è quindi probabile che un suo maggiore utilizzo abbia influenzato la categoria di traffico.

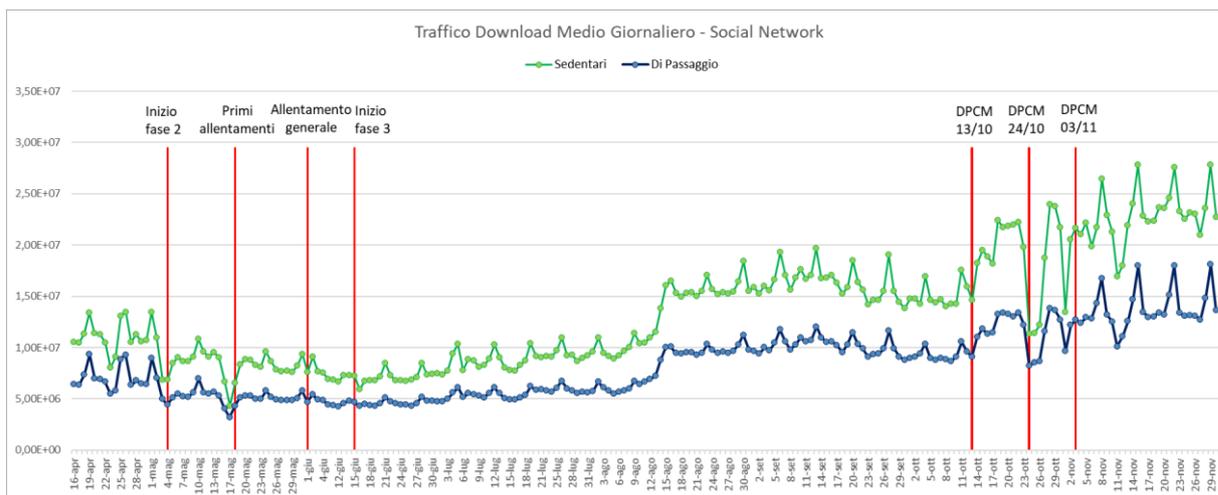


Grafico 13 – Social Network: traffico download medio giornaliero

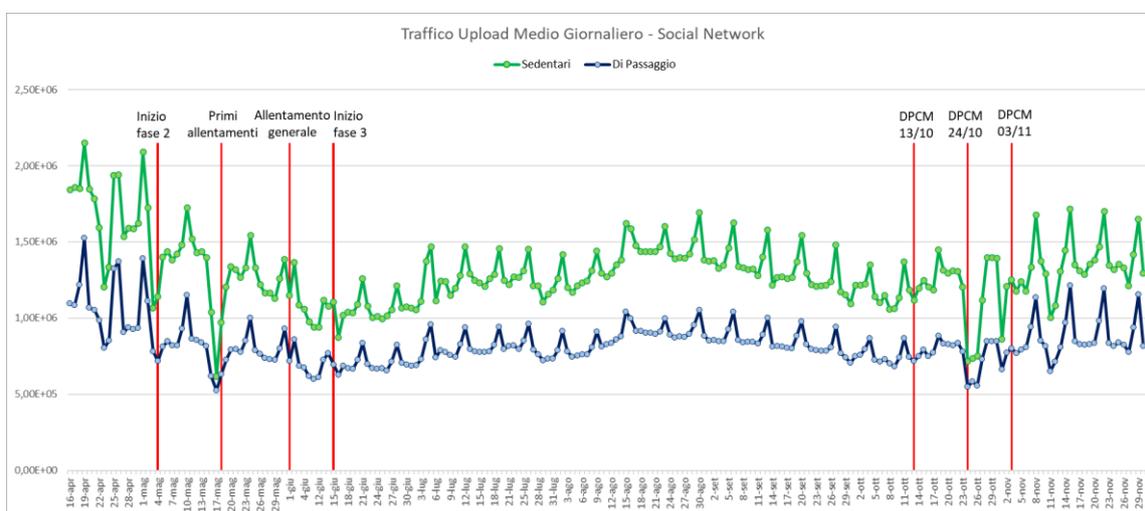


Grafico 14 – Social Network: traffico upload medio giornaliero

Coerentemente con quanto visto per il traffico fisso, anche in questo caso Gaming, Cloud, Miscellaneous e Software Update non hanno mostrato pattern legati alle varie fasi della pandemia in Italia.

Di seguito si riporta, a titolo di esempio, il grafico relativo alla categoria Gaming (Grafico 15). Come si può osservare, l'andamento sembra svincolato da quello della pandemia e legato piuttosto a eventi occasionali che costituiscono i picchi verso l'alto, tra i quali il lancio della terza stagione di Fortnite, noto web game, il 17 giugno, oppure la trasmissione in diretta streaming interattiva di Gamescon, la fiera internazionale dei videogiochi online, il 27 agosto.

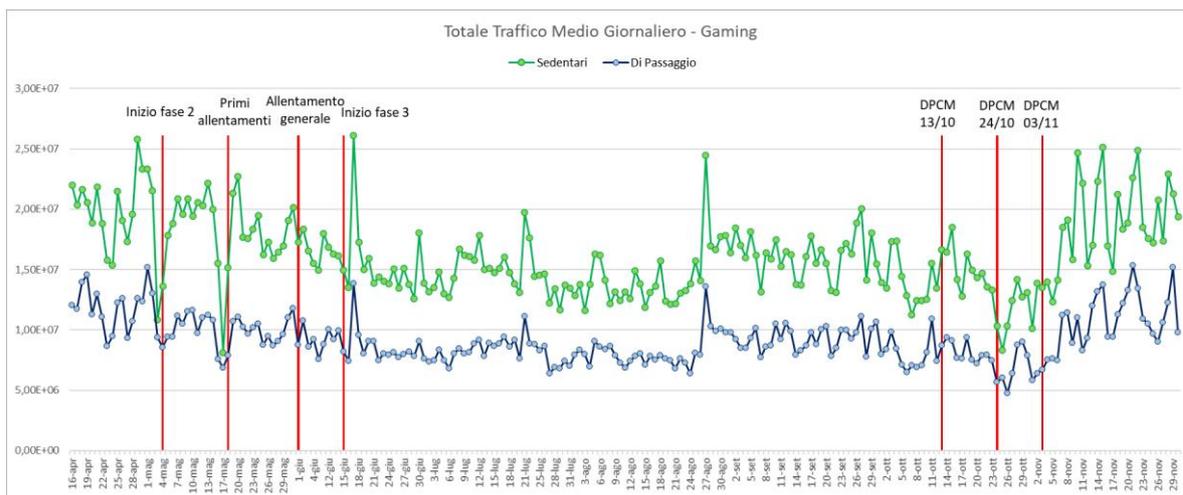


Grafico 15 – Gaming: totale traffico medio giornaliero

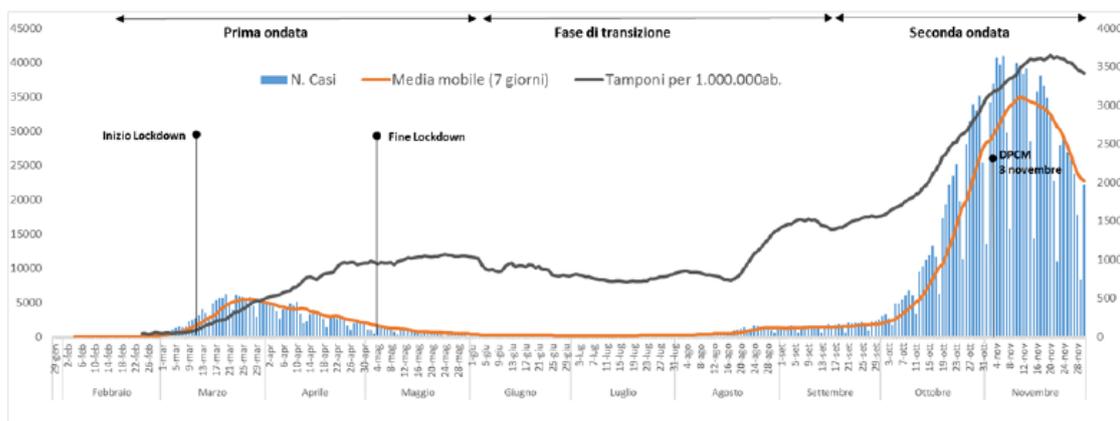
2.3 Province più e meno colpite

Nel Quarto Rapporto pubblicato dall'Istat in collaborazione con l'Istituto Superiore di Sanità, è stata presentata un'analisi di mortalità nel periodo compreso tra gennaio e novembre 2020, in cui si evince anche che in Italia sono stati registrati 1.651.229 casi positivi di Covid-19. Come visto in precedenza, l'orizzonte temporale in esame può essere suddiviso in una prima fase, da febbraio a maggio, concentrata prevalentemente in specifiche zone del Nord, in cui si è verificato una rapida diffusione sia dei contagi sia delle vittime; una seconda fase, cosiddetta di *transizione*, con un apparente arresto della diffusione del virus; e una terza fase, la *seconda ondata*, in cui la curva dei contagi e quella dei decessi hanno ripreso a crescere in modo esponenziale.

Febbraio 2020 si è posto in totale controtendenza rispetto ai mesi antecedenti, che a partire dal 2019 avevano registrato una diminuzione del tasso di mortalità rispetto agli anni dal 2015 al 2018. Solo durante la prima fase dell'epidemia sono stati registrati 211 mila decessi, ben 50 mila in più rispetto allo stesso periodo (marzo-maggio) degli anni precedenti, dato ancor più sconcertante se si pensa che 45 mila di essi sono relativi a residenti del Nord Italia, zona maggiormente colpita. Durante i mesi estivi, invece, tale crescita si è arrestata, riportando i valori di mortalità alla pari con quelli registrati dal 2015 al 2019, per poi riprendere nuovamente a crescere durante la seconda ondata.

Nell'arco temporale in esame, si è registrato anche un anomalo incremento nel numero di morti avvenute a causa di interventi tardivi o cure non tempestivamente somministrate: la saturazione dei posti in terapia intensiva è l'esempio più eclatante, a cui vanno aggiunti tutti i casi in cui pazienti con patologie non hanno ricevuto i farmaci necessari in tempo.

Come si può osservare nel *Grafico 16*, il numero di contagi è notevolmente aumentato in termini assoluti durante la seconda ondata. È necessario però precisare che a tale dato si deve affiancare l'incredibile miglioramento nella capacità diagnostica: in questo modo è stato possibile rilevare molti più soggetti asintomatici, operando su scala nazionale.



Fonte: ISS, Sistema di sorveglianza integrata Covid-19.

Grafico 16 - Numero di casi di Covid-19 per data di prelievo/diagnosi e numero di tamponi per milione di abitanti.

Per tali motivi, si è deciso di focalizzare la seconda parte dell'analisi, propedeutica all'ultima, sul calcolo delle province più e meno colpite dal Covid-19, a partire dal dataset presente sul sito della Protezione Civile contenente, per ogni giorno, il totale dei casi verificatisi nella provincia.

Per ognuna delle 107 unità territoriali presenti, è stata individuata la mediana del numero totale di contagi nel periodo analizzato ed è stata divisa per il totale della popolazione residente nella provincia, con i dati disponibili sul sito dell'Istat, che riportava la popolazione aggiornata al 1° gennaio 2020.

È stato calcolato il seguente indice:

$$\text{Contagiosità} = \frac{\text{mediana totale contagiati}}{\text{popolazione}} \quad (2.3)$$

Gli indici ottenuti sono poi stati ordinati e sono state selezionate le province più e meno colpite, utilizzando come riferimento primo e il nono decile della distribuzione.

Le province più colpite sono risultate essere Cremona, Piacenza, Lodi, Bergamo, Brescia, Pavia, Alessandria, Valle d'Aosta, Reggio nell'Emilia, Trento e Asti; mentre quelle meno colpite Vibo Valentia, Potenza, Taranto, Palermo, Oristano, Nuoro, Agrigento, Trapani, Ragusa e la provincia del Sud Sardegna (Tabella 1).

| Provincia | Abbreviazione | Popolazione | Mediana | Contagiosità |
|--------------------------------|---------------|-------------|----------|--------------|
| Cremona | CR | 358.347 | 6.678 | 1,8636% |
| Piacenza | PC | 287.236 | 4.487 | 1,5621% |
| Lodi | LO | 230.607 | 3.600 | 1,5611% |
| Bergamo | BG | 1.116.384 | 14.668,5 | 1,3139% |
| Brescia | BS | 1.268.455 | 15.762 | 1,2426% |
| Pavia | PV | 546.515 | 5.592,5 | 1,0233% |
| Alessandria | AL | 419.037 | 4.080 | 0,9737% |
| Valle d'Aosta / Vallée d'Aoste | AO | 125.501 | 1.196 | 0,9530% |
| Reggio nell'Emilia | RE | 531.751 | 4.970,5 | 0,9347% |
| Trento | TN | 542.739 | 4.880 | 0,8991% |
| Asti | AT | 213.216 | 1.880 | 0,8817% |
| Sondrio | SO | 180.941 | 1.595 | 0,8815% |
| Mantova | MN | 411.062 | 3.610,5 | 0,8783% |
| Lecco | LC | 337.087 | 2.855 | 0,8470% |
| Parma | PR | 453.930 | 3.703 | 0,8158% |
| Vercelli | VC | 170.296 | 1.344 | 0,7892% |
| Pesaro e Urbino | PU | 357.137 | 2.763 | 0,7737% |
| Novara | NO | 368.040 | 2.803 | 0,7616% |
| Milano | MI | 3.279.944 | 24.585 | 0,7496% |
| Verbano-Cusio-Ossola | VB | 157.455 | 1.146 | 0,7278% |
| Imperia | IM | 213.919 | 1.527 | 0,7138% |
| Torino | TO | 2.252.379 | 15.930,5 | 0,7073% |
| Como | CO | 603.828 | 4.125 | 0,6831% |
| Genova | GE | 835.829 | 5.675,5 | 0,6790% |
| Monza e della Brianza | MB | 878.267 | 5.811,5 | 0,6617% |
| Rimini | RN | 339.796 | 2.175,5 | 0,6402% |
| Biella | BI | 174.384 | 1.053 | 0,6038% |
| Trieste | TS | 233.276 | 1.404,5 | 0,6021% |
| Belluno | BL | 201.972 | 1.198,5 | 0,5934% |
| Savona | SV | 274.183 | 1.593 | 0,5810% |
| Modena | MO | 707.292 | 3.940,5 | 0,5571% |
| Verona | VR | 930.339 | 5.132,5 | 0,5517% |

| | | | | |
|---------------|----|-----------|---------|---------|
| Massa-Carrara | MS | 193.934 | 1.052 | 0,5425% |
| Bologna | BO | 1.017.806 | 5.368 | 0,5274% |
| Pescara | PE | 318.678 | 1.596 | 0,5008% |
| Bolzano | BZ | 532.080 | 2.659,5 | 0,4998% |
| Cuneo | CN | 586.568 | 2.892 | 0,4930% |
| Forlì-Cesena | FC | 394.833 | 1.757 | 0,4450% |
| Varese | VA | 892.532 | 3.933 | 0,4407% |
| Padova | PD | 939.672 | 3.980 | 0,4236% |
| La Spezia | SP | 219.196 | 895 | 0,4083% |
| Ancona | AN | 469.750 | 1.878,5 | 0,3999% |
| Macerata | MC | 312.146 | 1.157 | 0,3707% |
| Lucca | LU | 388.678 | 1.367 | 0,3517% |
| Firenze | FI | 1.004.298 | 3.374,5 | 0,3360% |
| Vicenza | VI | 862.363 | 2.876 | 0,3335% |
| Venezia | VE | 851.663 | 2.697 | 0,3167% |
| Ferrara | FE | 344.840 | 1.064,5 | 0,3087% |
| Treviso | TV | 888.309 | 2.694 | 0,3033% |
| Ravenna | RA | 389.634 | 1.066,5 | 0,2737% |
| Fermo | FM | 173.004 | 473 | 0,2734% |
| Enna | EN | 162.368 | 438 | 0,2698% |
| Rieti | RI | 154.232 | 413 | 0,2678% |
| Pistoia | PT | 293.059 | 749 | 0,2556% |
| Pordenone | PN | 312.619 | 710 | 0,2271% |
| Pisa | PI | 422.310 | 932 | 0,2207% |
| Prato | PO | 258.152 | 562,5 | 0,2179% |
| Chieti | CH | 383.189 | 828 | 0,2161% |
| Teramo | TE | 307.412 | 650 | 0,2114% |
| Arezzo | AR | 341.766 | 685 | 0,2004% |
| Grosseto | GR | 220.785 | 422 | 0,1911% |
| Rovigo | RO | 233.386 | 445 | 0,1907% |
| Udine | UD | 526.256 | 1.001 | 0,1902% |
| Foggia | FG | 616.310 | 1.170 | 0,1898% |
| Sassari | SS | 489.634 | 877 | 0,1791% |
| Terni | TR | 224.882 | 382 | 0,1699% |
| Brindisi | BR | 390.456 | 660 | 0,1690% |
| Campobasso | CB | 218.679 | 364 | 0,1665% |
| Siena | SI | 266.238 | 437,5 | 0,1643% |
| Livorno | LI | 333.509 | 529,5 | 0,1588% |
| Gorizia | GO | 139.206 | 218,5 | 0,1570% |
| Perugia | PG | 655.403 | 1.009 | 0,1540% |
| Viterbo | VT | 316.142 | 459 | 0,1452% |
| Ascoli Piceno | AP | 206.363 | 293 | 0,1420% |
| Roma | RM | 4.333.274 | 6.021 | 0,1389% |
| Avellino | AV | 413.926 | 571 | 0,1379% |
| Frosinone | FR | 485.241 | 668 | 0,1377% |
| Bari | BA | 1.249.246 | 1.494 | 0,1196% |

| | | | | |
|-----------------------|----|-----------|---------|---------|
| Latina | LT | 5.766.55 | 622 | 0,1079% |
| Matera | MT | 195.998 | 211 | 0,1077% |
| Barletta-Andria-Trani | BT | 388.390 | 382 | 0,0984% |
| Catania | CT | 1.104.974 | 1.007,5 | 0,0912% |
| Messina | ME | 620.721 | 550 | 0,0886% |
| Napoli | NA | 3.082.905 | 2.659,5 | 0,0863% |
| L'Aquila | AQ | 296.491 | 246 | 0,0830% |
| Siracusa | SR | 397.037 | 329 | 0,0829% |
| Benevento | BN | 274.080 | 209 | 0,0763% |
| Isernia | IS | 83.586 | 62 | 0,0742% |
| Caltanissetta | CL | 260.779 | 186 | 0,0713% |
| Crotone | KR | 170.718 | 120 | 0,0703% |
| Cosenza | CS | 700.385 | 469 | 0,0670% |
| Lecce | LE | 791.122 | 525 | 0,0664% |
| Salerno | SA | 1.092.779 | 698,5 | 0,0639% |
| Caserta | CE | 922.171 | 580 | 0,0629% |
| Catanzaro | CZ | 354.851 | 214 | 0,0603% |
| Cagliari | CA | 430.914 | 255 | 0,0592% |
| Reggio di Calabria | RC | 541.278 | 296 | 0,0547% |
| Vibo Valentia | VV | 157.469 | 86 | 0,0546% |
| Potenza | PZ | 360.936 | 191 | 0,0529% |
| Taranto | TA | 572.772 | 281 | 0,0491% |
| Palermo | PA | 1.243.328 | 547,5 | 0,0440% |
| Oristano | OR | 156.078 | 61 | 0,0391% |
| Nuoro | NU | 206.843 | 79 | 0,0382% |
| Agrigento | AG | 429.611 | 141 | 0,0328% |
| Trapani | TP | 428.377 | 139 | 0,0324% |
| Ragusa | RG | 321.215 | 98 | 0,0305% |
| Sud Sardegna | SU | 347.005 | 101 | 0,0291% |

Tabella 1- Indice di contagiosità per ogni provincia

2.4 Traffico nelle province più e meno colpite

La terza analisi aveva l'obiettivo di studiare i differenti consumi delle tipologie di traffico separatamente per province più e meno colpite.

Rete fissa

Anche in questo caso si è incominciato analizzando la linea fissa e focalizzandosi solo sulle quattro tipologie di traffico che nell'analisi iniziale avevano mostrato legami con l'andamento epidemico (Streaming, Social Network, Instant Messaging e Browsing).

Sia nella query delle province maggiormente colpite, sia in quella delle province minormente colpite, il dataset è stato filtrato per tipologia di rete ADSL o FTTC e per le unità territoriali appartenenti alle due selezioni effettuate nel paragrafo 2.3.

La parte successiva è del tutto analoga a quanto visto in precedenza: è stato effettuato un raggruppamento per data e calcolato il consumo medio giornaliero espresso nella (2.2).

Si riporta l'estratto di codice relativo alle unità territoriali più colpite per la tipologia Streaming, ma lo stesso è stato utilizzato anche negli altri casi.

```
#"Filtrate righe" = Table.SelectRows([...], each [atype] = "ADSL" or [atype] = "FTTC"),

#"Filtrate righe1" = Table.SelectRows([...], each [Provincia] = "CR" or [Provincia] = "PC" or [Provincia] = "LO" or [Provincia] = "BG" or [Provincia] = "BS" or [Provincia] = "PV" or [Provincia] = "AL" or [Provincia] = "AO" or [Provincia] = "RE" or [Provincia] = "TN" or [Provincia] = "AT"),

#"Raggruppate righe" = Table.Group([...], {"date"}, {"Dw streaming", each List.Sum([rx_streaming]), type nullable number}, {"Up streaming", each List.Sum([tx_streaming]), type nullable number}, {"Ut streaming", each List.Sum([du_streaming]), type nullable number}, [...]}),

#"Aggiunta colonna personalizzata" = Table.AddColumn("#Raggruppate righe", "Streaming", each ([Dw streaming]+[Up streaming])/[Ut streaming])
```

Come già rilevato nell'analisi iniziale, solamente due sono le tipologie di fonti di traffico sul fisso che risultano avere chiari legami con l'andamento della pandemia: Social Network (in Grafico 17) e Instant Messaging (Grafico 18).

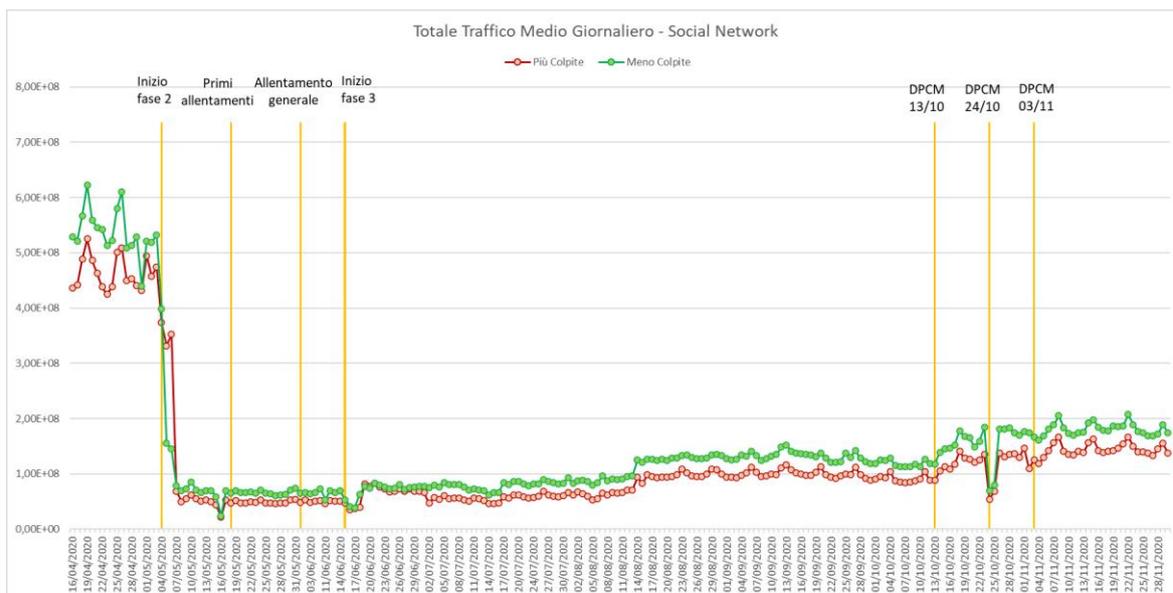


Grafico 17 – Social Network: traffico medio giornaliero di province più e meno colpite

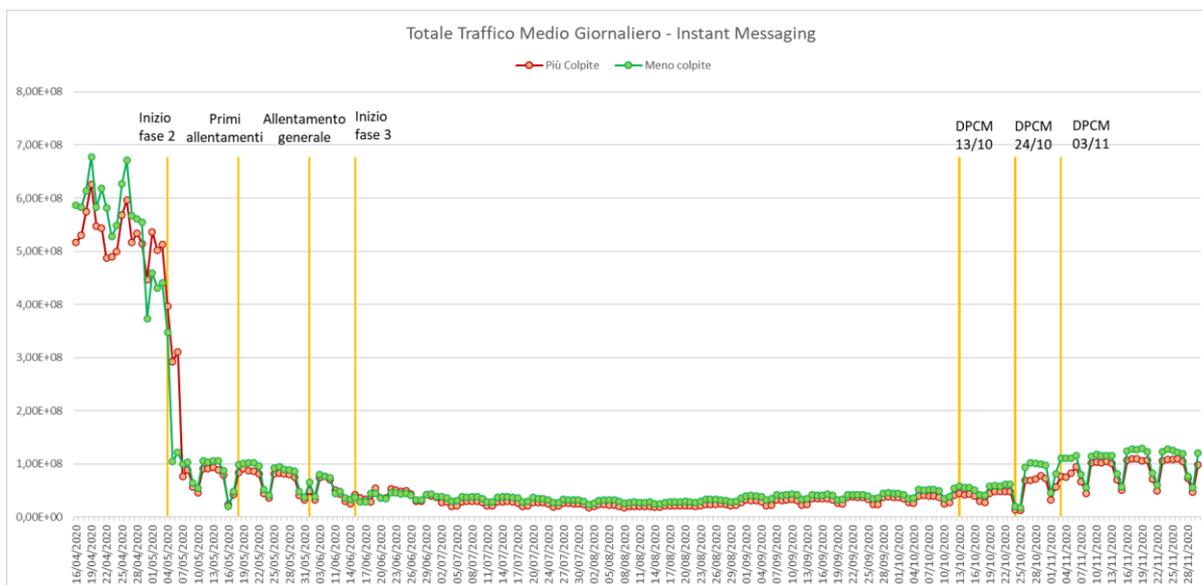


Grafico 18 – Instant Messaging: totale traffico medio giornaliero su rete fissa;

Contrariamente alle aspettative, le province meno colpite riportano un consumo leggermente maggiore rispetto alle più colpite, mantenendo trend simili lungo tutto l’asse temporale.

Già S. Castriota, M. Delmastro, M. Tonin [2020] hanno mostrato nella loro pubblicazione “National or Local? The Demand for News in Italy during Covid-19” un comportamento inatteso e per certi versi riconducibile a quanto ottenuto sopra: hanno analizzato la domanda di news a livello nazionale e locale, scoprendo che l’aumento dei telespettatori

dei telegiornali, in entrambi i casi, è fortemente correlato con il numero dei positivi delle altre regioni, mentre il numero di positivi della propria regione è quasi sempre negativo e mai statisticamente significativo.

Rete mobile

Il processo che riguarda la rete mobile è stato analogo a quello utilizzato per la rete fissa, con gli opportuni adattamenti già visti nel paragrafo 2.2: ADSL e FTTC sono stati sostituiti da 2G, 3G e 4G; il raggruppamento non è stato più solamente a livello di data, ma anche per tipologia di utente, così da ottenere una differenziazione tra i *sedentari* e gli utenti *di passaggio*.

Permane la differenza tra fisso e mobile in termini di categorie di fonti di traffico legate alla pandemia: qui si confermano Streaming, Social Network, Instant Messaging e Browsing.

Per quanto riguarda gli utenti *sedentari*, il risultato ottenuto è simile a quanto rilevato sulla linea fissa: le province meno colpite hanno mostrato un valore quasi sempre maggiore rispetto alle più colpite.

Per facilitare il confronto, si riportano i due grafici relativi a Social Network (Grafico 19) e a Instant Messaging (Grafico 20).

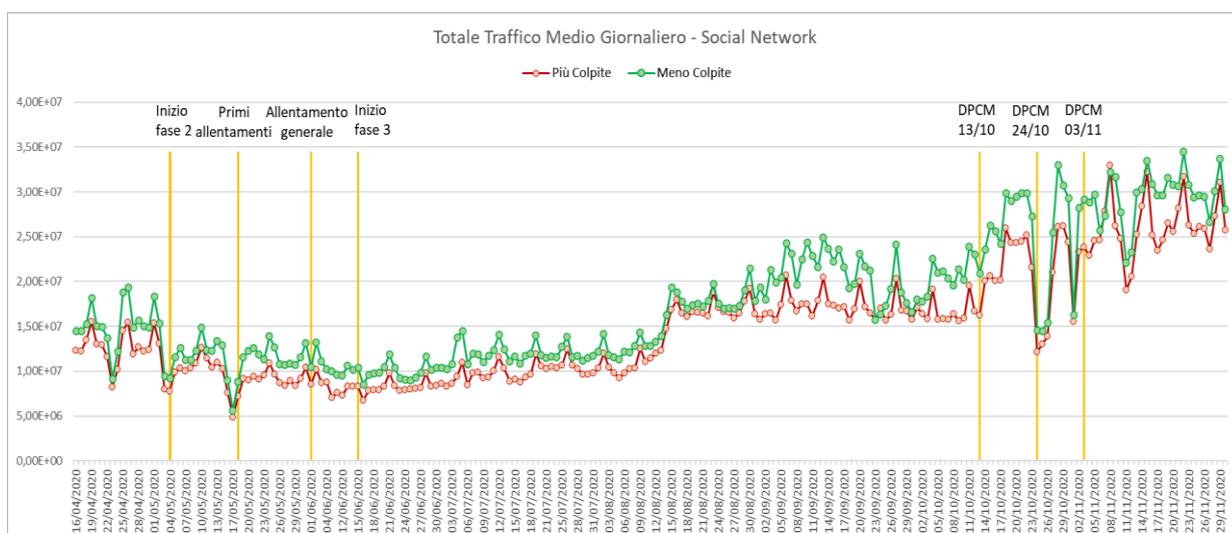


Grafico 19 - Social Network: totale traffico medio giornaliero su rete mobile – Sedentari

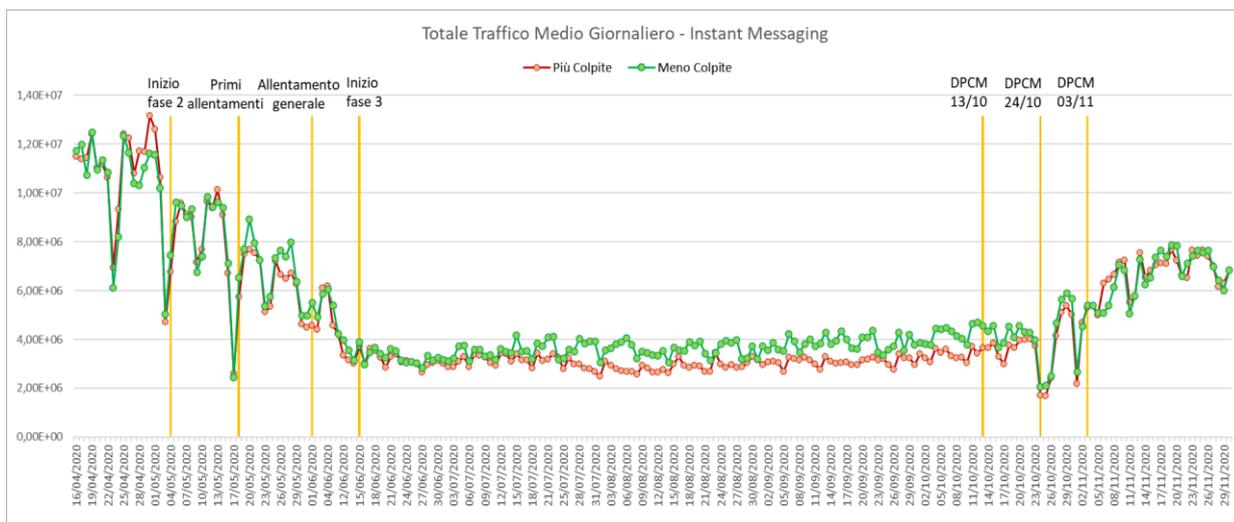


Grafico 20 - Instant Messaging: totale traffico medio giornaliero su rete mobile - Sedentari

Come si può notare, i due andamenti sono differenti su rete fissa e su mobile: nel primo caso il consumo risulta notevolmente maggiore fino alla fine del lockdown nazionale, per poi stabilizzarsi a valori più bassi e riprendere a crescere solo di poco nella seconda fase; diverso è quello che capita sul mobile, dove i valori ricominciano a salire in modo rilevante con l'aumento delle restrizioni, seppur sempre rimanendo lontano dai valori iniziali.

Anche per quanto riguarda gli utenti *di passaggio* (Grafico 21 e Grafico 22), il consumo medio delle province più colpite risulta sempre inferiore rispetto alle meno colpite, mentre i trend sono allineati a quanto verificato per i *sedentari* su rete mobile.

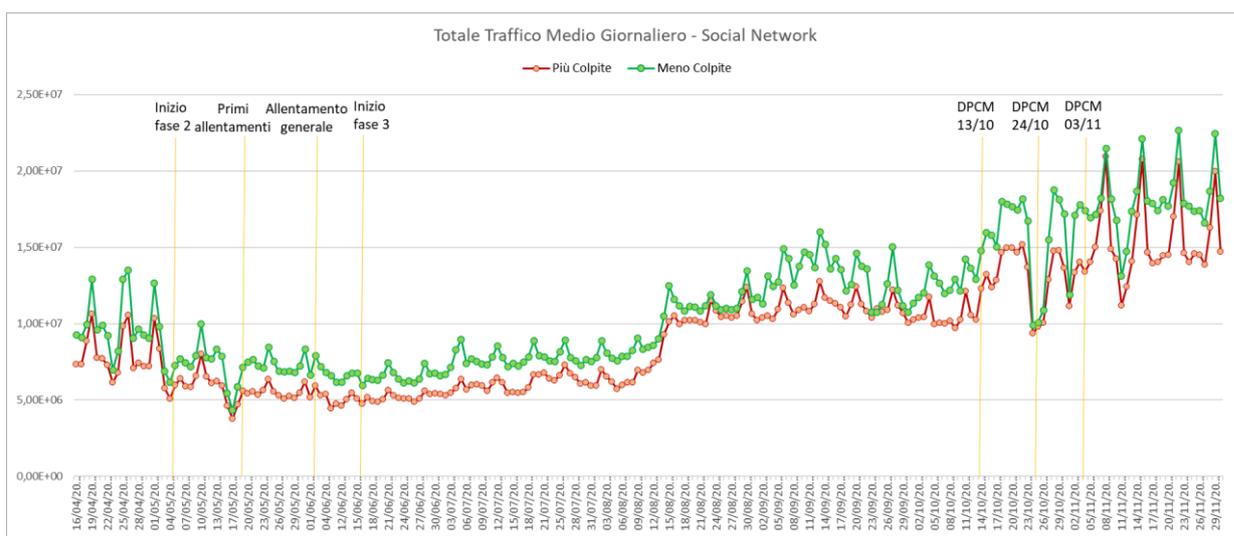


Grafico 21 - Social Network: totale traffico medio giornaliero - Utenti di passaggio

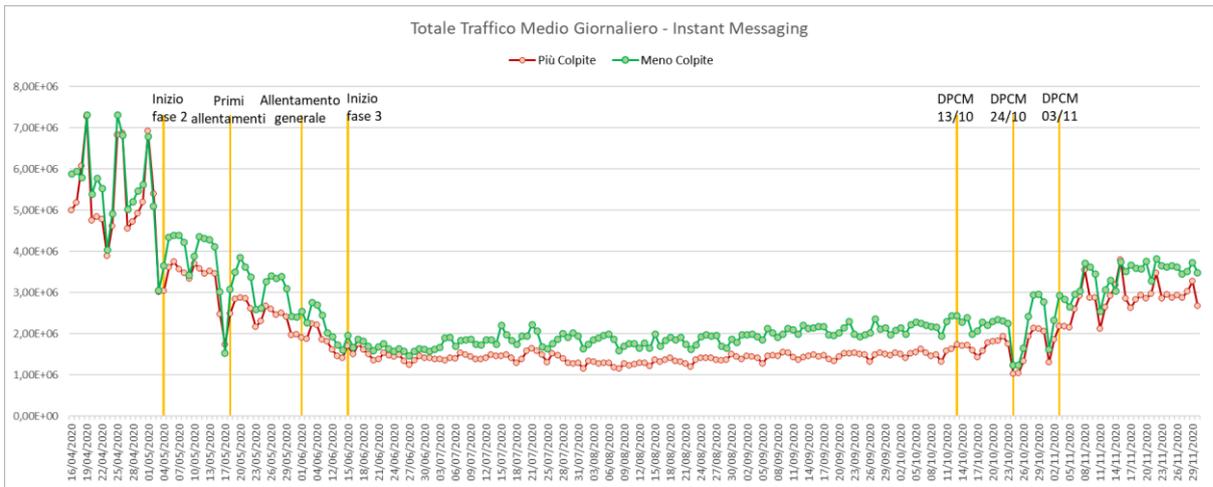


Grafico 22 - Instant Messaging: totale traffico medio giornaliero - Utenti di passaggio

L'ultima analisi ha avuto come scopo quello di indagare, distintamente per zone meno e più colpite (in Grafico 23 e Grafico 24), l'andamento del traffico medio consumato da utenti *sedentari* e *di passaggio* su rete mobile, per verificare se ci fosse un differente gap tra le due categorie, che però è risultato pressoché simile.

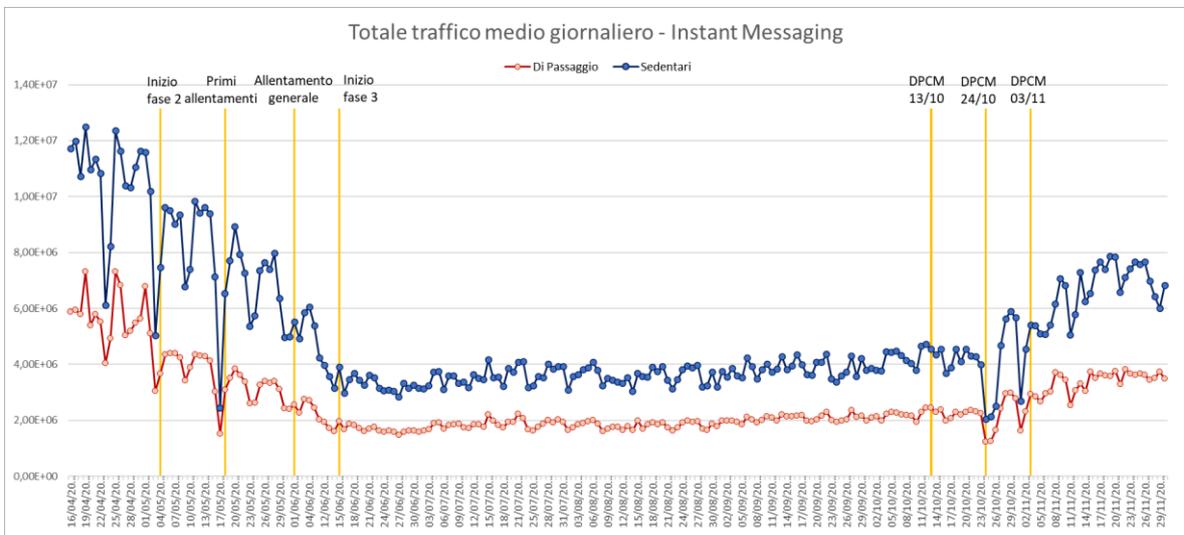


Grafico 23 - Instant Messaging: totale traffico medio giornaliero – Province meno colpite

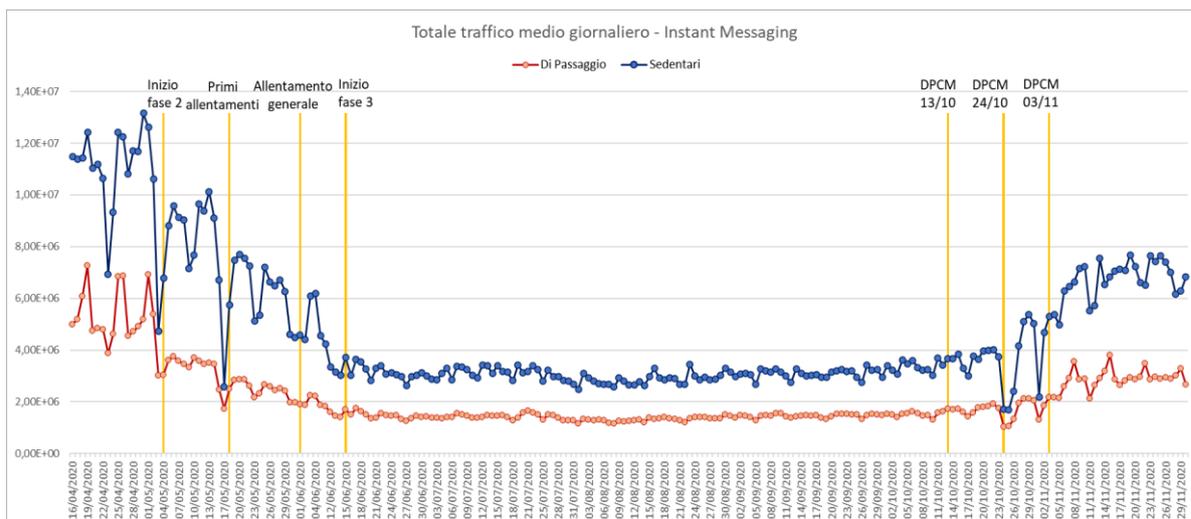


Grafico 24 - Instant Messaging: totale traffico medio giornaliero - Province più colpite

2.5 Suddivisione in zone

Con la suddivisione delle province eseguita in precedenza, è risultato evidente come le unità più colpite (Cremona, Piacenza, Lodi, Bergamo, Brescia, Pavia, Alessandria, Aosta, Reggio nell'Emilia, Trento e Asti) fossero tutte appartenenti al Nord Italia, mentre quelle con un minor indice di contagiosità appartenessero al Sud Italia e alle Isole (Vibo Valentia, Potenza, Taranto, Palermo, Oristano, Nuoro, Agrigento, Trapani, Ragusa, Sud Sardegna). Per tale motivo, si è deciso di effettuare un'analisi più accurata, tramite la suddivisione dell'Italia in tre zone di riferimento (Nord, Centro e Sud), utilizzando la ripartizione presente sul dataset dei comuni italiani dell'Istat già descritto al paragrafo 1.1.

I valori soglia per identificare le unità territoriali appartenenti alle due categorie in questo caso sono stati il quinto e il novantacinquesimo percentile della distribuzione.

Per agevolare il confronto, si riporteranno gli stessi grafici per le tre zone.

Nord Italia

Per quanto riguarda il Nord Italia, la suddivisione è stata la seguente (Tabella 2):

- Province più colpite: Cremona, Piacenza, Lodi;
- Province meno colpite: Rovigo, Udine, Gorizia.

| Provincia | Abbreviazione | Zona | Popolazione | Mediana | Contagiosità |
|-----------------------|---------------|------------|-------------|----------|--------------|
| Cremona | CR | Nord-ovest | 358.347 | 6.678 | 1,864% |
| Piacenza | PC | Nord-est | 287.236 | 4.487 | 1,562% |
| Lodi | LO | Nord-ovest | 230.607 | 3.600 | 1,561% |
| Bergamo | BG | Nord-ovest | 1.116.384 | 14.668,5 | 1,314% |
| Brescia | BS | Nord-ovest | 1.268.455 | 15.762 | 1,243% |
| Pavia | PV | Nord-ovest | 546.515 | 5.592,5 | 1,023% |
| Alessandria | AL | Nord-ovest | 419.037 | 4.080 | 0,974% |
| Valle d'Aosta | AO | Nord-ovest | 125.501 | 1.196 | 0,953% |
| Reggio nell'Emilia | RE | Nord-est | 531.751 | 4.970,5 | 0,935% |
| Trento | TN | Nord-est | 542.739 | 4.880 | 0,899% |
| Asti | AT | Nord-ovest | 213.216 | 1.880 | 0,882% |
| Sondrio | SO | Nord-ovest | 180.941 | 1.595 | 0,882% |
| Mantova | MN | Nord-ovest | 411.062 | 3.610,5 | 0,878% |
| Lecco | LC | Nord-ovest | 337.087 | 2.855 | 0,847% |
| Parma | PR | Nord-est | 453.930 | 3.703 | 0,816% |
| Vercelli | VC | Nord-ovest | 170.296 | 1.344 | 0,789% |
| Novara | NO | Nord-ovest | 368.040 | 2.803 | 0,762% |
| Milano | MI | Nord-ovest | 3.279.944 | 24.585 | 0,750% |
| Verbano-Cusio-Ossola | VB | Nord-ovest | 157.455 | 1.146 | 0,728% |
| Imperia | IM | Nord-ovest | 213.919 | 1.527 | 0,714% |
| Torino | TO | Nord-ovest | 2.252.379 | 15.930,5 | 0,707% |
| Como | CO | Nord-ovest | 603.828 | 4.125 | 0,683% |
| Genova | GE | Nord-ovest | 835.829 | 5.675,5 | 0,679% |
| Monza e della Brianza | MB | Nord-ovest | 878.267 | 5.811,5 | 0,662% |
| Rimini | RN | Nord-est | 339.796 | 2.175,5 | 0,640% |
| Biella | BI | Nord-ovest | 174.384 | 1.053 | 0,604% |
| Trieste | TS | Nord-est | 233.276 | 1.404,5 | 0,602% |
| Belluno | BL | Nord-est | 201.972 | 1.198,5 | 0,593% |
| Savona | SV | Nord-ovest | 274.183 | 1.593 | 0,581% |
| Modena | MO | Nord-est | 707.292 | 3.940,5 | 0,557% |
| Verona | VR | Nord-est | 930.339 | 5.132,5 | 0,552% |
| Bologna | BO | Nord-est | 1.017.806 | 5.368 | 0,527% |
| Bolzano | BZ | Nord-est | 532.080 | 2.659,5 | 0,500% |
| Cuneo | CN | Nord-ovest | 586.568 | 2.892 | 0,493% |
| Forlì-Cesena | FC | Nord-est | 394.833 | 1.757 | 0,445% |
| Varese | VA | Nord-ovest | 892.532 | 3.933 | 0,441% |
| Padova | PD | Nord-est | 939.672 | 3.980 | 0,424% |
| La Spezia | SP | Nord-ovest | 219.196 | 895 | 0,408% |
| Vicenza | VI | Nord-est | 862.363 | 2.876 | 0,334% |
| Venezia | VE | Nord-est | 851.663 | 2.697 | 0,317% |
| Ferrara | FE | Nord-est | 344.840 | 1.064,5 | 0,309% |
| Treviso | TV | Nord-est | 888.309 | 2.694 | 0,303% |

| | | | | | |
|-----------|----|----------|---------|---------|--------|
| Ravenna | RA | Nord-est | 389.634 | 1.066,5 | 0,274% |
| Pordenone | PN | Nord-est | 312.619 | 710 | 0,227% |
| Rovigo | RO | Nord-est | 233.386 | 445 | 0,191% |
| Udine | UD | Nord-est | 526.256 | 1.001 | 0,190% |
| Gorizia | GO | Nord-est | 139.206 | 218,5 | 0,157% |

Tabella 2 – Indice di contagiosità per ogni provincia del Nord Italia

Contrariamente a quanto rilevato nell'indagine su scala nazionale, analizzando il consumo medio su rete fissa, in tutte e quattro le categorie in esame le province maggiormente colpite presentavano valori mediamente superiori rispetto alle meno colpite.

Il legame con le fasi della pandemia, invece, ancora una volta è stato confermato solo per quanto riguarda Social Network (Grafico 25) e Instant Messaging.

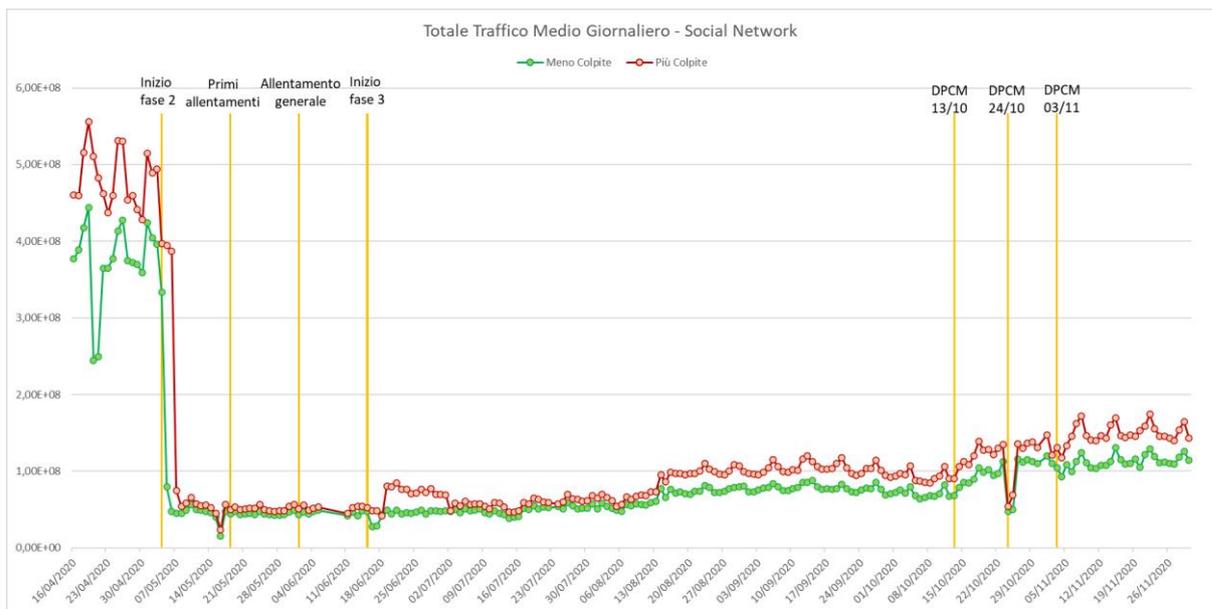


Grafico 25 – Social Network: totale traffico medio giornaliero su rete fissa

Ancor più evidente è stata la differenza rilevata tra le due tipologie di province quando si sono studiati gli andamenti su rete mobile, sia nel caso di utenti *sedentari* (Grafico 26), sia nel caso di utenti *di passaggio* (Grafico 27).

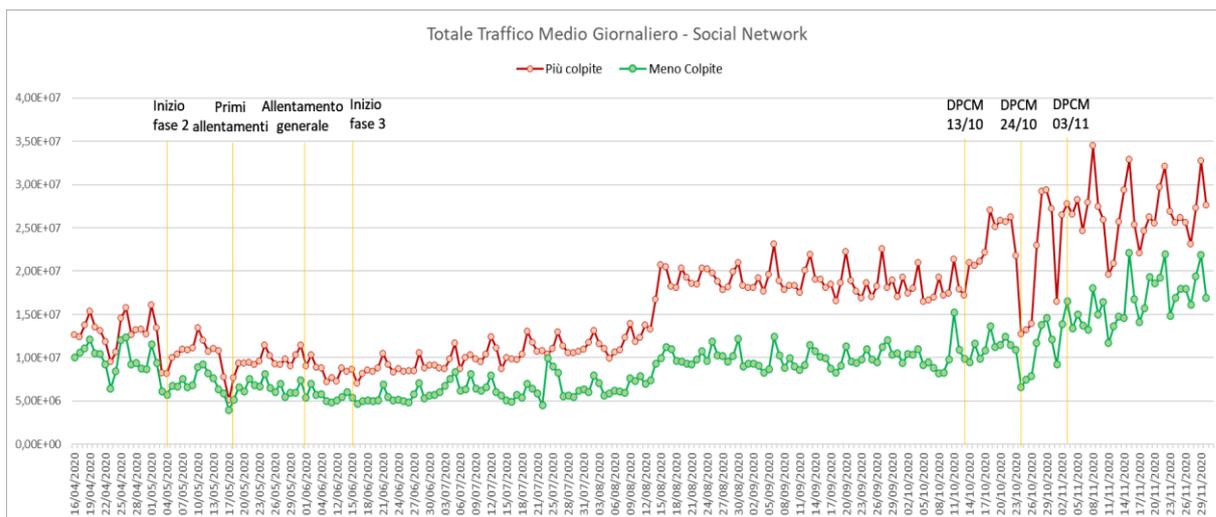


Grafico 26 – Social Network: totale traffico medio giornaliero su rete mobile per utenti sedentari

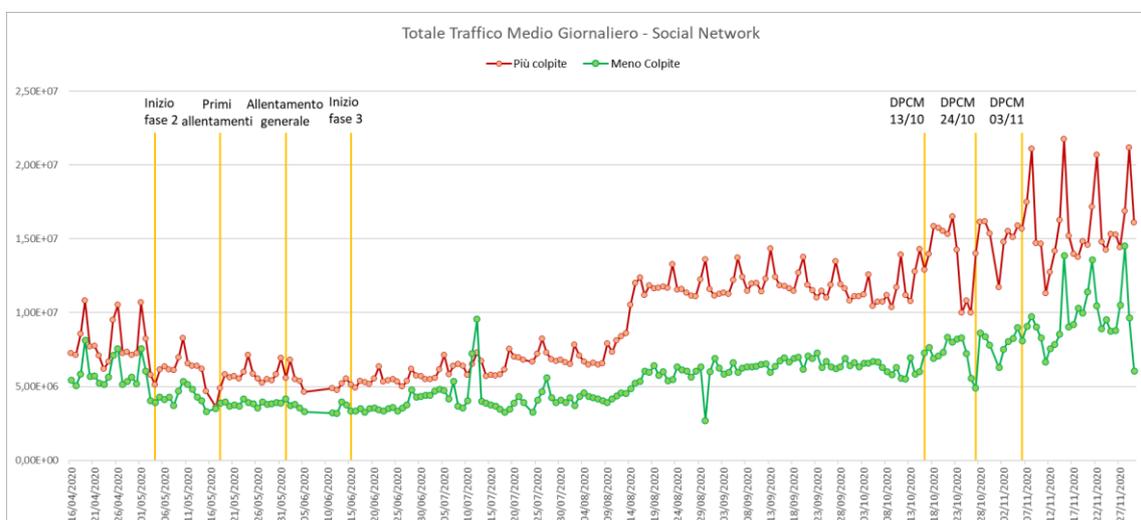


Grafico 27 – Social Network: totale traffico medio giornaliero per utenti di passaggio

Ponendo in relazione i due andamenti del traffico degli utenti *sedentari* e di quelli *di passaggio*, separando questa volta le unità meno colpite (Grafico 28) da quelle più colpite (Grafico 29), è stato osservato un maggior gap nelle seconde.

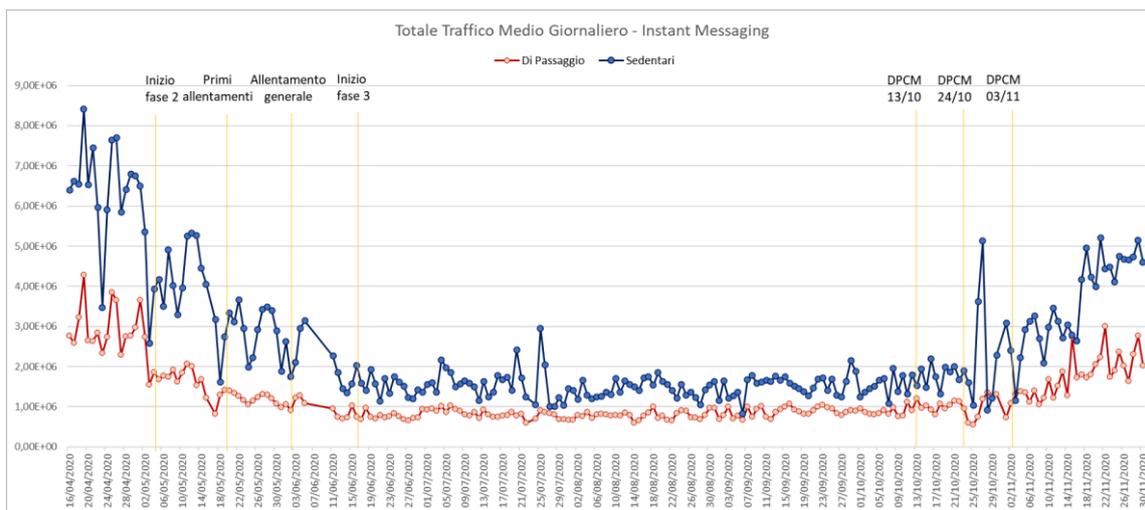


Grafico 28- Instant Messaging: totale traffico medio giornaliero nelle province meno colpite

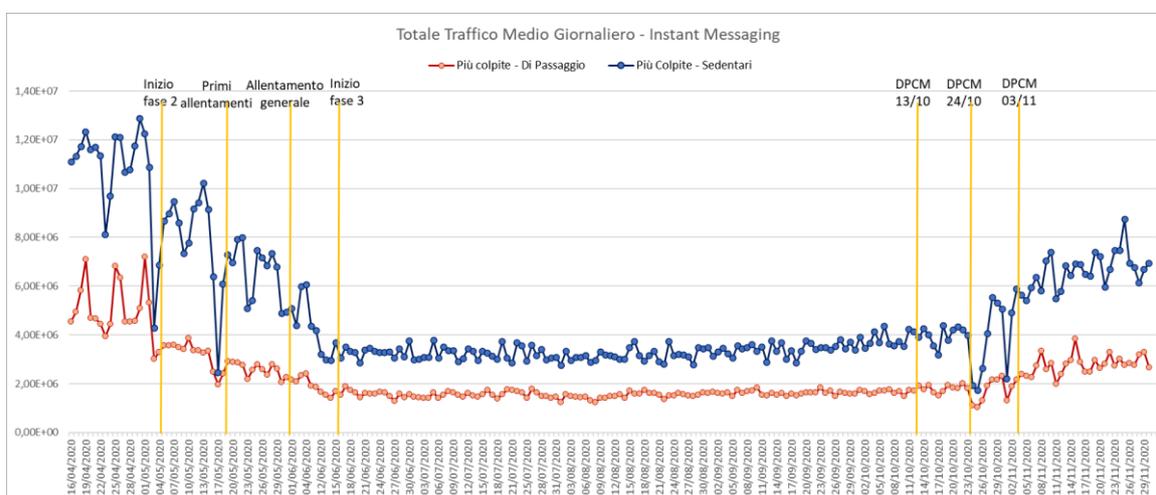


Grafico 29 - Instant Messaging: totale traffico medio giornaliero nelle province più colpite

Centro Italia

Le 23 province del Centro Italia, sono state così suddivise in base alla soglia del top e bottom 5% (Tabella 3):

- Province più colpite: Pesaro e Urbino, Massa-Carrara;
- Province meno colpite: Frosinone, Latina.

| Provincia | Abbreviazione | Zona | Popolazione | Mediana | Contagiosità |
|-----------------|---------------|--------|-------------|---------|--------------|
| Pesaro e Urbino | PU | Centro | 357.137 | 2763 | 0,774% |
| Massa-Carrara | MS | Centro | 193.934 | 1052 | 0,542% |

| | | | | | |
|---------------|----|--------|-----------|--------|--------|
| Ancona | AN | Centro | 469.750 | 1878,5 | 0,400% |
| Macerata | MC | Centro | 312.146 | 1157 | 0,371% |
| Lucca | LU | Centro | 388.678 | 1367 | 0,352% |
| Firenze | FI | Centro | 1.004.298 | 3374,5 | 0,336% |
| Fermo | FM | Centro | 173.004 | 473 | 0,273% |
| Rieti | RI | Centro | 154.232 | 413 | 0,268% |
| Pistoia | PT | Centro | 293.059 | 749 | 0,256% |
| Pisa | PI | Centro | 422.310 | 932 | 0,221% |
| Prato | PO | Centro | 258.152 | 562,5 | 0,218% |
| Arezzo | AR | Centro | 341.766 | 685 | 0,200% |
| Grosseto | GR | Centro | 220.785 | 422 | 0,191% |
| Terni | TR | Centro | 224.882 | 382 | 0,170% |
| Siena | SI | Centro | 266.238 | 437,5 | 0,164% |
| Livorno | LI | Centro | 333.509 | 529,5 | 0,159% |
| Perugia | PG | Centro | 655.403 | 1009 | 0,154% |
| Viterbo | VT | Centro | 316.142 | 459 | 0,145% |
| Ascoli Piceno | AP | Centro | 206.363 | 293 | 0,142% |
| Roma | RM | Centro | 4.333.274 | 6021 | 0,139% |
| Frosinone | FR | Centro | 485.241 | 668 | 0,138% |
| Latina | LT | Centro | 576.655 | 622 | 0,108% |

Tabella 3 - Indice di contagiosità per ogni provincia del Centro Italia

Come sempre, l'analisi è iniziata dalla rete fissa e subito si è potuto notare un comportamento differente rispetto al Nord: nelle tipologie di traffico in esame non sono emerse differenze rilevanti tra le province più e meno colpite, fatta eccezione, in entrambi i casi di Social Network e Instant Messaging, per il periodo compreso tra il 16 aprile e il 4 maggio, dove è prevalente, seppur di poco, il consumo medio nelle province più colpite (Grafico 30 e Grafico 31).

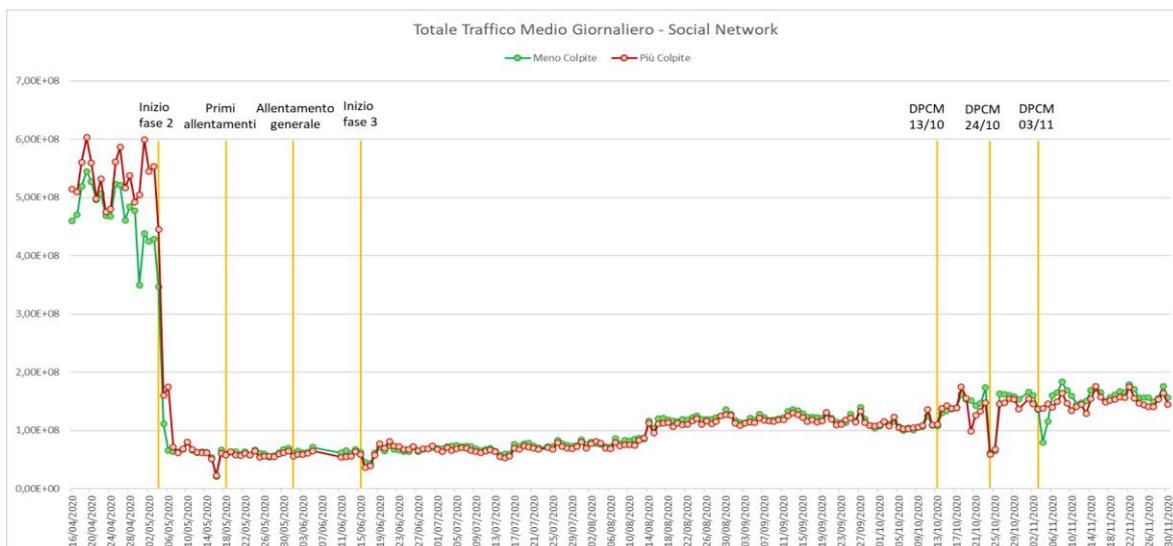


Grafico 30 – Social Network: totale traffico medio giornaliero su linea fissa

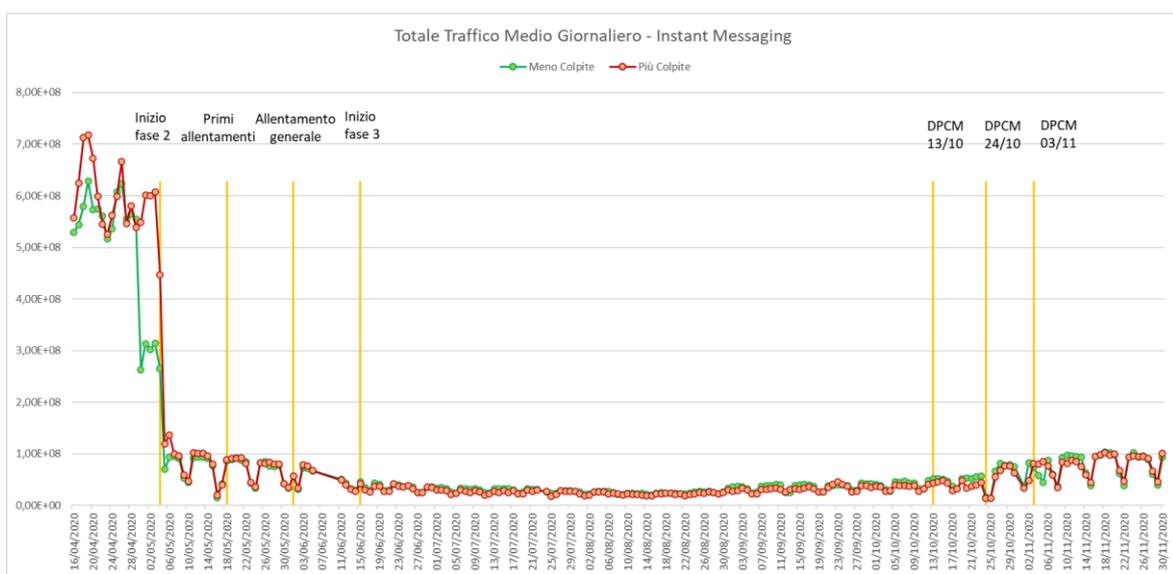


Grafico 31 - Instant Messaging: totale traffico medio giornaliero su rete fissa

Nell'analisi relativa al traffico su mobile per gli utenti *sedentari* (Grafico 32) il consumo medio è prevalente nelle province meno colpite in tutte e quattro le categorie, contrariamente a quanto avveniva nel caso delle province del Nord Italia.

Diverso è il discorso per gli utenti di passaggio, che invece mostrano caratteristiche simili a quelle rilevate nell'analisi su rete fissa: non emergono particolari discrepanze tra il consumo rilevato nelle zone maggiormente o minormente colpite (Grafico 33).

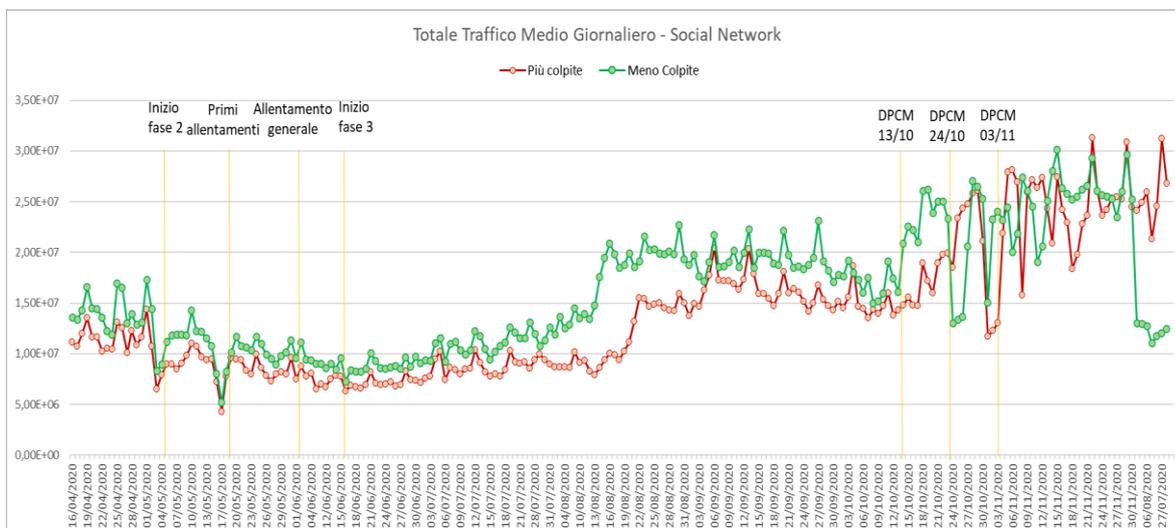


Grafico 32 - Social Network: totale traffico medio giornaliero su mobile per utenti sedentari

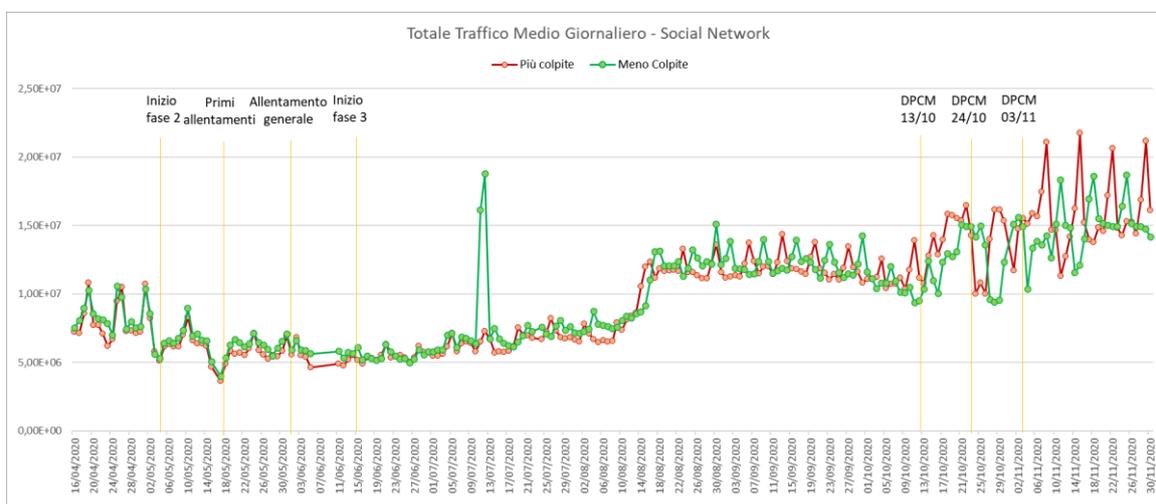


Grafico 33 – Social Network: totale traffico medio giornaliero per utenti di passaggio

Analizzando i gap tra utenti *sedentari* e *di passaggio* distintamente per le zone meno (Grafico 34) e più colpite (Grafico 35), appaiono simili nella fase iniziale, ma successivamente si ha un più notevole distacco nelle province meno colpite.

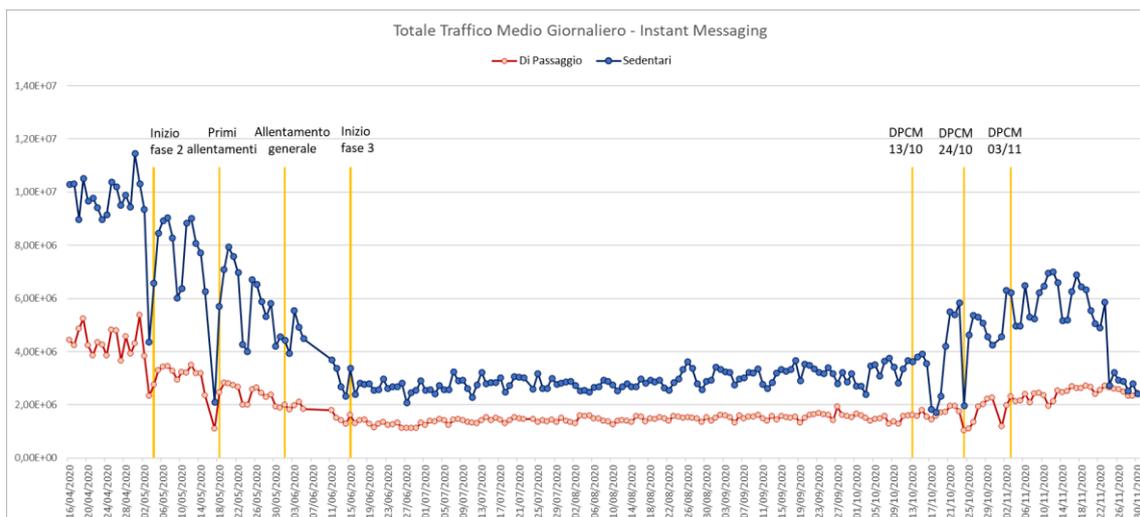


Grafico 34 - Instant Messaging: totale traffico medio giornaliero per le province meno colpite

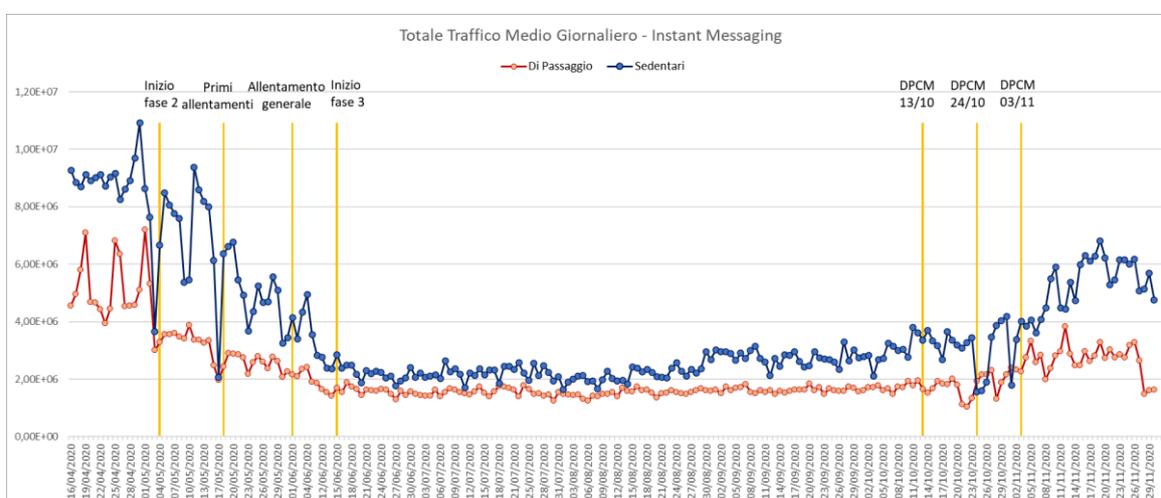


Grafico 35 - Instant Messaging: totale traffico medio giornaliero per le province più colpite

Sud Italia e Isole

L'ultima analisi si è focalizzata sul Sud Italia e sulle Isole (Tabella 4):

- Le province più colpite sono risultate essere Pescara ed Enna;
- Le province meno colpite sono state Ragusa e il Sud Sardegna.

| Provincia | Abbreviazione | Zona | Popolazione | Mediana | Contagiosità |
|-----------|---------------|-------|-------------|---------|--------------|
| Pescara | PE | Sud | 318.678 | 1.596 | 0,5008% |
| Enna | EN | Isole | 162.368 | 438 | 0,2698% |
| Chieti | CH | Sud | 383.189 | 828 | 0,2161% |
| Teramo | TE | Sud | 307.412 | 650 | 0,2114% |

| | | | | | |
|-----------------------|----|-------|-----------|---------|---------|
| Foggia | FG | Sud | 616.310 | 1.170 | 0,1898% |
| Sassari | SS | Isole | 489.634 | 877 | 0,1791% |
| Brindisi | BR | Sud | 390.456 | 660 | 0,1690% |
| Campobasso | CB | Sud | 218.679 | 364 | 0,1665% |
| Avellino | AV | Sud | 413.926 | 571 | 0,1379% |
| Bari | BA | Sud | 1.249.246 | 1.494 | 0,1196% |
| Matera | MT | Sud | 195.998 | 211 | 0,1077% |
| Barletta-Andria-Trani | BT | Sud | 388.390 | 382 | 0,0984% |
| Catania | CT | Isole | 1.104.974 | 1.007,5 | 0,0912% |
| Messina | ME | Isole | 620.721 | 550 | 0,0886% |
| Napoli | NA | Sud | 3.082.905 | 2.659,5 | 0,0863% |
| L'Aquila | AQ | Sud | 296.491 | 246 | 0,0830% |
| Siracusa | SR | Isole | 397.037 | 329 | 0,0829% |
| Benevento | BN | Sud | 274.080 | 209 | 0,0763% |
| Isernia | IS | Sud | 83.586 | 62 | 0,0742% |
| Caltanissetta | CL | Isole | 260.779 | 186 | 0,0713% |
| Crotone | KR | Sud | 170.718 | 120 | 0,0703% |
| Cosenza | CS | Sud | 700.385 | 469 | 0,0670% |
| Lecce | LE | Sud | 791.122 | 525 | 0,0664% |
| Salerno | SA | Sud | 1.092.779 | 698,5 | 0,0639% |
| Caserta | CE | Sud | 922.171 | 580 | 0,0629% |
| Catanzaro | CZ | Sud | 354.851 | 214 | 0,0603% |
| Cagliari | CA | Isole | 430.914 | 255 | 0,0592% |
| Reggio di Calabria | RC | Sud | 541.278 | 296 | 0,0547% |
| Vibo Valentia | VV | Sud | 157.469 | 86 | 0,0546% |
| Potenza | PZ | Sud | 360.936 | 191 | 0,0529% |
| Taranto | TA | Sud | 572.772 | 281 | 0,0491% |
| Palermo | PA | Isole | 1.243.328 | 547,5 | 0,0440% |
| Oristano | OR | Isole | 156.078 | 61 | 0,0391% |
| Nuoro | NU | Isole | 206.843 | 79 | 0,0382% |
| Agrigento | AG | Isole | 429.611 | 141 | 0,0328% |
| Trapani | TP | Isole | 428.377 | 139 | 0,0324% |
| Ragusa | RG | Isole | 321.215 | 98 | 0,0305% |
| Sud Sardegna | SU | Isole | 347.005 | 101 | 0,0291% |

Tabella 4 - Indice di contagiosità per ogni provincia del Centro Italia

Per quanto riguarda il fisso, sia su Social Network sia su Instant Messaging, gli andamenti dei due tipi di province sono pressoché sovrapposti (Grafico 36), ancor più di quanto già succedeva nel Centro Italia.

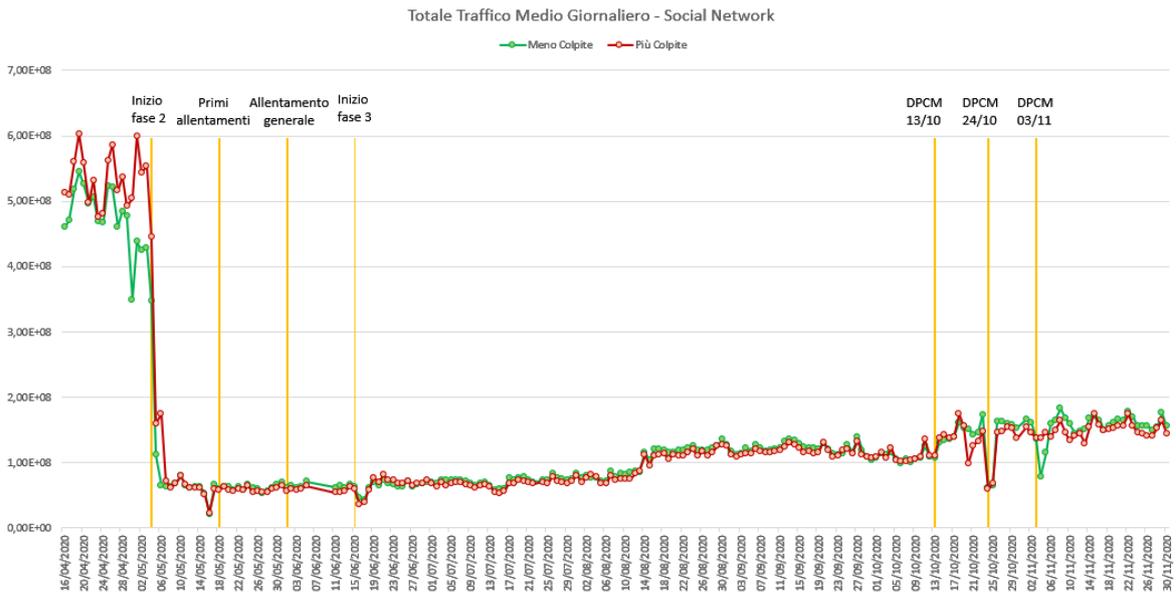


Grafico 36 – Social Network: totale traffico medio giornaliero su rete fissa

Anche sul mobile non si evidenziano particolari discrepanze, né per gli utenti *sedentari* (Grafico 37), né per quelli *di passaggio* (Grafico 38).

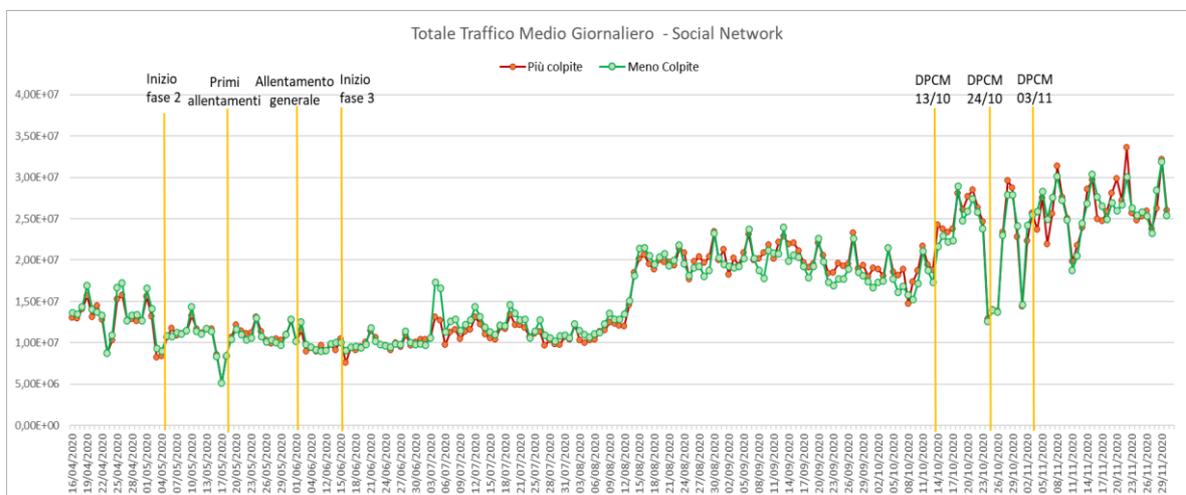


Grafico 37 – Sociale Network: totale traffico medio giornaliero su rete mobile per utenti sedentari

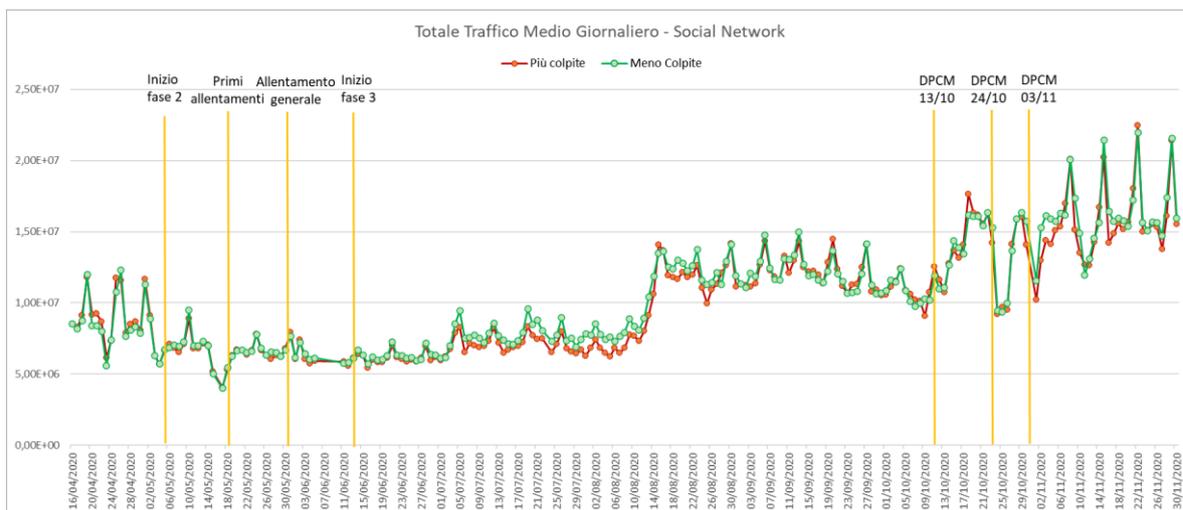


Grafico 38 - Sociale Network: totale traffico medio giornaliero su rete mobile per utenti di passaggio

Per quanto riguarda i gap tra i consumi medi degli utenti *sedentari* e di quelli *di passaggio*, separatamente per i due gruppi di province meno e più colpite (rispettivamente Grafico 39 e Grafico 40), non sembrano esserci rilevanti differenze.

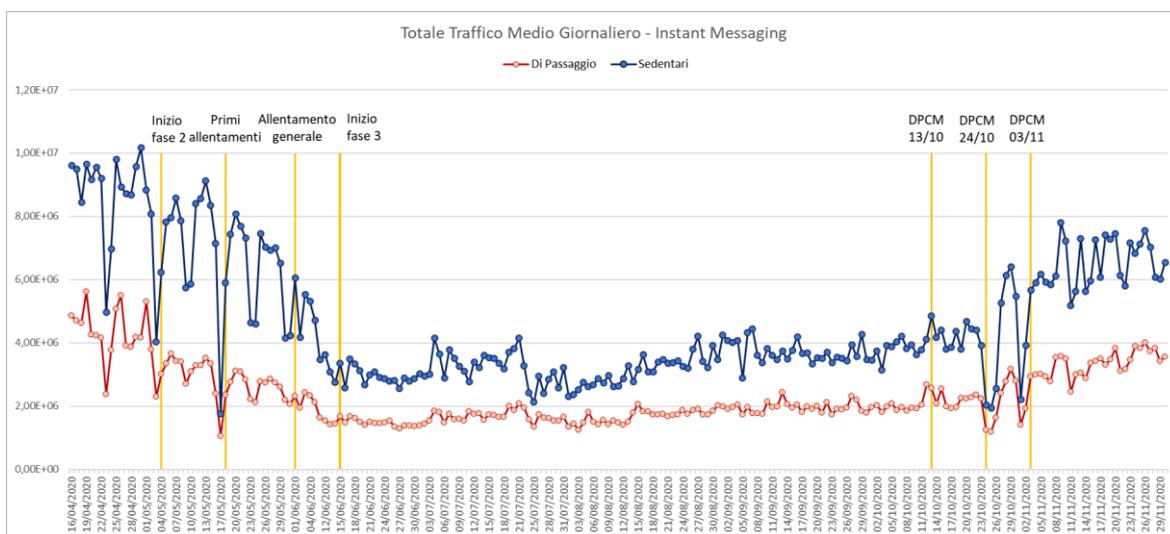


Grafico 39 - Instant Messaging: totale traffico medio giornaliero su rete mobile per province meno colpite

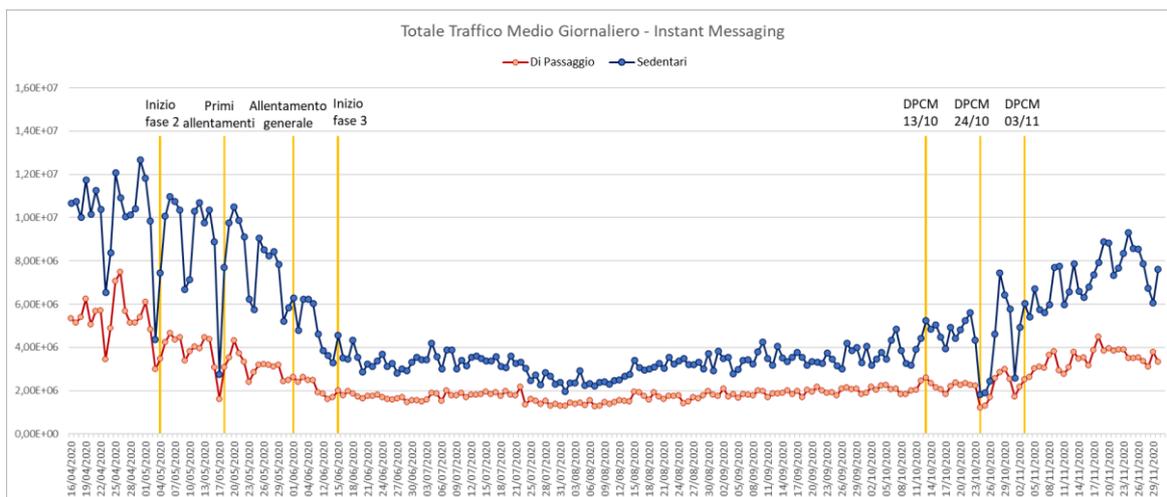


Gráfico 40 - Instant Messaging: totale traffico medio giornaliero su rete mobile per province più colpite

2.6 Considerazioni di fine capitolo

A valle dell'esecuzione di queste analisi preliminari, sono state tratte in particolare due conclusioni:

- Non tutte le tipologie di traffico mostrano andamenti legati alla pandemia. Sono state rilevate, infatti, solamente quattro categorie (Streaming, Social Network, Instant Messaging e Browsing) per la rete mobile e di queste solamente due (Social Network e Instant Messaging) sono state confermate anche sul fisso;
- La suddivisione dell'Italia in tre zone e i confronti tra province più e meno colpite al loro interno produce risultati interessanti e nettamente differenti rispetto a quelli ottenuti con la distinzione su scala nazionale.

Tali considerazioni saranno utilizzate come punto di partenza per sviluppare il modello econometrico presentato nel prossimo capitolo.

Capitolo 3: Analisi di regressione

In seguito alle considerazioni tratte dall'analisi descrittiva del capitolo precedente, è stato effettuato uno studio di regressione al fine di indagare se e in che misura la diffusione dei contagi da COVID-19 influenzasse il consumo delle diverse tipologie di Internet già selezionate: Streaming, Social Network, Instant Messaging e Browsing.

Coerentemente con quanto fatto in precedenza, si è ritenuto necessario continuare a differenziare il traffico fisso da quello mobile: per tale motivo sono state effettuate analisi distinte per le due tipologie, anche al fine di catturare un eventuale effetto sostituzione tra di esse.

Sono stati separati i quattro principali periodi presenti nell'orizzonte temporale in esame e, sempre in proseguimento con quanto già descritto, si è deciso di inserire una distinzione geografica, dapprima suddividendo per le tre zone d'Italia e in un secondo momento, per aumentare il livello di dettaglio, inserendo un focus sulle province più colpite.

Il modello utilizzato è l'OLS, *Ordinary Least Squares* e tutte le analisi sono state effettuate con il software Stata12.

3.1 Preparazione del dataset

Dal momento che il numero dei contagi presente sul dataset utilizzato aveva una granularità a livello provinciale, è stato necessario ottenere i consumi medi giornalieri provinciali per le quattro categorie di traffico identificate. A tal fine, è stato utilizzato il seguente codice su Power Query (qui riportato per il caso dello Streaming su mobile, ma analogo metodo è stato adottato anche per le altre tipologie e sulla rete fissa):

```

let [...]

#"Filtrate righe" = Table.SelectRows("#Rimosse colonne", each [atype] = "2G" or
[atype] = "3G" [atype] = "4G")

#"Raggruppate righe" = Table.Group("#Rinominate colonne", {"date",
"Provincia"}, {"Dw Streaming", each List.Sum([rx_streaming]), type nullable
number}, {"Up Streaming", each List.Sum([tx_streaming]), type nullable number},
{"Ut Streaming", each List.Sum([du_streaming]), type nullable number}, [...] }),

#"Aggiunta colonna personalizzata" = Table.AddColumn("#Raggruppate righe",
"Streaming medio", each ([Dw Streaming]+[Up Streaming])/[Ut Streaming])

```

Nel dataset ottenuto è stata aggiunta una variabile temporale indicante quattro periodi selezionati:

- Il primo, dal 16 aprile 2020 al 3 maggio 2020, ossia giorni appartenenti al lockdown nazionale;
- La seconda, dal 4 maggio 2020 al 14 giugno 2020, la cosiddetta *fase due*;
- La terza, dal 15 giugno 2020 al 12 ottobre 2020, la fase estiva con minori restrizioni;
- La quarta, dal 13 ottobre 2020 al 30 novembre 2020, l'inizio della *seconda ondata*.

Sono stati definiti, inoltre, altri due attributi: uno che permettesse l'associazione della provincia con la rispettiva zona di appartenenza (Nord, Centro e Sud Italia) e un altro al fine di collocarla o meno nell'insieme delle più colpite (sulla base della classificazione già fatta nel Capitolo 2: Cremona, Enna, Lodi, Massa-Carrara, Pesaro e Urbino, Pescara, Piacenza).

3.2 Definizione del modello

Dal momento che l'obiettivo era quello di indagare come il Covid-19 avesse cambiato i consumi nell'universo ICT in Italia, sono state effettuate analisi differenti per ogni tipologia di fonte di traffico (Streaming, Social Network, Instant Messaging e Browsing). Il loro valore medio giornaliero ha costituito la variabile dipendente, mentre come variabile indipendente è stata scelta l'unica che potesse rappresentare la diffusione del Covid-19, ossia il numero dei contagi.

È stata definita la seguente retta di regressione, per ognuna delle quattro fonti di traffico:

$$y_{t,p} = \beta_0 + \beta_1 C_{t,p} + \delta_t + \delta_p + \varepsilon_{t,p} \quad (3.1)$$

La variabile dipendente, $y_{t,p}$, indica, per la singola tipologia di traffico, il consumo medio totale sulla rete (fissa o mobile) al giorno t , nella provincia p .

La variabile di interesse è $C_{t,p}$, e rappresenta il numero di casi registratisi in totale nella provincia p al giorno t .

Le due dummy δ_t e δ_p sono state inserite per controllare eventuali effetti fissi sulla data o sulla provincia, mentre $\varepsilon_{t,p}$ rappresenta l'errore. Inoltre, in tutti i casi è stato introdotto il comando `cl(provincia_code)` in Stata, al fine di effettuare una clusterizzazione a livello provinciale degli errori standard eteroschedastici.

Il coefficiente β_0 è l'intercetta della retta di regressione, mentre β_1 ne indica il coefficiente angolare:

$$\frac{dy_{t,p}}{dC_{t,p}} = \beta_1$$

La sua interpretazione è la seguente: l'aumento di un'unità nella variabile che rappresenta l'effetto Covid-19, produrrà un aumento o una diminuzione di β_1 unità di consumo medio giornaliero per la tipologia di traffico in esame.

Il metodo OLS ha come obiettivo quello di fornire la stima migliore possibile di β_1 , identificando quindi una curva di ottimo che più si avvicini a quella reale, costituita dalla distribuzione dei punti della popolazione. Per fare ciò, la funzione individuata deve essere quella che minimizza la somma degli errori quadrati tra la curva definita e quella dei punti osservati. Nella retta (3.1) gli scarti sono rappresentati dal termine $\varepsilon_{t,p}$, che altro non è che la differenza tra i valori osservati della variabile dipendente e quelli predetti dal modello.

Analisi effettuate

Al fine di decidere quale fosse la miglior rappresentazione per il caso in esame, sono state eseguite diverse analisi.

Gli effetti fissi sono sempre stati introdotti gradualmente: in ogni caso si è partiti prima da un'analisi che non ne conteneva alcuno, per poi aggiungere quello sulla data e successivamente anche quello sulla provincia. Quest'ultima è stata la scelta utilizzata poi nel modello finale, anche poiché garantiva output maggiormente controllati.

Per ogni analisi sono inoltre state considerate tre specifiche nel definire la variabile indicante il numero di casi:

- Essa era costituita dal numero di casi verificatisi in totale in una provincia in un determinato giorno, così da identificare una possibile correlazione tra il numero giornaliero dei contagi e l'utilizzo delle diverse categorie di traffico;
- Era rappresentata dal numero di casi verificatisi nel giorno $t-1$ e di conseguenza permetteva di osservare come il numero di contagi del giorno precedente avesse o meno impattato sull'utilizzo delle varie fonti di traffico;
- Era formata dalla differenza tra i casi registratisi in una provincia fino al giorno t e quelli al giorno $t-1$. La scelta di includere tale specifica è stata conseguenza del modo in cui durante tutto l'orizzonte pandemico una delle informazioni su cui spesso è stata posta l'attenzione è stato l'incremento o il decremento marginale dei contagi da un giorno all'altro.

I risultati ottenuti dalle prime due specifiche erano, però, pressoché identici e per tale motivo si è scelto di proseguire considerando il totale dei casi giornalieri, omettendo quindi la seconda specifica. La terza, invece, ha quasi sempre permesso di ottenere risultati differenti, per tale motivo saranno presentati anche essi nell'analisi che sarà introdotta più avanti.

Oltre alla funzione lineare già descritta (3.1), ogni analisi è stata anche ripetuta in forma di funzione quadratica (3.2), al fine di osservare se una curva di secondo grado potesse costituire una migliore rappresentazione e fossero quindi presenti delle non linearità:

$$y_{t,p} = \beta_0 + \beta_1 C_{t,p} + \beta_2 C_{t,p}^2 + \delta_t + \delta_p + \varepsilon_{t,p} \quad (3.2)$$

Il modello che, però, si è adattato meglio di tutti al caso in esame è stato quello in cui si è effettuata una trasformazione logaritmica e, nello specifico, di tipo log-log:

$$\log y_{t,p} = \beta_0 + \beta_1 \log C_{t,p} + \delta_t + \delta_p + \varepsilon_{t,p} \quad (3.3)$$

In questo modo è stato possibile verificare come la variazione percentuale del numero dei contagi (o della loro differenza giornaliera) abbia provocato una variazione di un valore β_1 percentuale nel consumo medio giornaliero di traffico Internet.

Per completezza si riportano in Appendice alcuni output di esempio delle varie analisi svolte.

3.3 Interazioni con i periodi

Per continuità con quanto fatto nel Capitolo 2, si riportano di seguito gli output provenienti dall'analisi sulla rete mobile e il confronto con quelli ottenuti per la rete fissa, al fine di indagare un possibile effetto sostituzione tra le due, o comportamenti del tutto differenti.

Sono presentate le due specifiche descritte in precedenza:

- La prima, in cui la variabile di interesse è costituita dal totale dei casi registratisi fino al giorno t nella provincia p ;
- La seconda, in cui la variabile di interesse è ottenuta dalla differenza tra i casi di un giorno in una provincia e quelli del giorno precedente.

Prima specifica: numero totale di casi

La prima analisi di regressione è stata effettuata per indagare se e in che misura il numero di contagi nei quattro periodi identificati e le conseguenti restrizioni, avessero avuto impatti diversi sui consumi delle diverse fonti di traffico.

Per fare ciò, è stata interagita la variabile del numero dei casi con una dummy temporale ($D_{periodo}$), che indicava l'appartenenza a uno dei periodi sopra elencati:

$$\log y_{t,p} = \beta_0 + \beta_1 \log C_{t,p} \times D_{periodo} + \delta_t + \delta_p + \varepsilon_{t,p} \quad (3.4)$$

La significatività, che con il software Stata utilizzato può assumere valori pari all'1%, al 5% o al 10%, è definita tramite l'utilizzo del test T di Student:

$$t = \frac{\widehat{\beta}_1 - \beta_{1,0}}{SE(\widehat{\beta}_1)} \quad (3.5)$$

Così facendo è effettuato un test d'ipotesi per verificare l'eguaglianza tra lo stimatore $\widehat{\beta}_1$ e il valore ipotizzato $\beta_{1,0}$. La differenza tra i due β è divisa per la radice quadrata della varianza dello stimatore, ossia per il suo errore standard.

Per quanto riguarda la linea mobile (Tabella 5), è possibile osservare un coefficiente statisticamente significativo in corrispondenza dell'interazione tra numero di casi e primo, secondo e terzo periodo per quanto riguarda lo Streaming, con segno negativo; nel traffico su Instant Messaging, invece, emerge una correlazione positiva, ma solamente nel primo periodo.

I risultati ottenuti indicano che, a fronte di un aumento del 10% nel numero dei casi nel primo periodo (quello corrispondente alla fase di lockdown nazionale), il consumo medio giornaliero ha avuto una diminuzione dell'1,88% sullo Streaming, aumentando invece del 2,21% sull'Instant Messaging.

Su linea fissa (Tabella 6), invece, non sono presenti variabili statisticamente significative, indice che probabilmente i consumi sono stati guidati più dalle singole restrizioni che dal numero di casi.

| VARIABLES | (1) Log_Streaming | (2) Log_Social | (3) Log_IM | (4) Log_Browsing |
|--------------|----------------------|-------------------|---------------------|---------------------|
| log_casi1 | -0.188** (0.079) | -0.107 (0.100) | 0.221*** (0.071) | -0.126 (0.087) |
| log_casi2 | -0.113* (0.062) | 0.041 (0.082) | 0.097 (0.065) | -0.058 (0.086) |
| log_casi3 | -0.090* (0.054) | -0.031 (0.088) | 0.094 (0.068) | -0.125 (0.089) |
| log_casi4 | -0.096 (0.068) | -0.055 (0.111) | 0.150 (0.090) | -0.070 (0.099) |
| Observations | 23,687 | 23,687 | 23,688 | 23,688 |
| R-squared | 0.114 | 0.087 | 0.075 | 0.076 |

Robust standard errors
in parentheses

*** p<0.01, ** p<0.05,

* p<0.1

Tabella 5 - Traffico mobile: interazione tra numero di casi e periodo di appartenenza

| VARIABLES | (1) Log_Streaming | (2) Log_Social | (3) Log_IM | (4) Log_Browsing |
|--------------|----------------------|-------------------|------------------|---------------------|
| log_casi1 | -0.054 (0.043) | 0.027 (0.032) | 0.083 (0.059) | -0.001 (0.037) |
| log_casi2 | -0.028 (0.043) | -0.022 (0.029) | 0.078 (0.049) | -0.008 (0.033) |
| log_casi3 | -0.014 (0.037) | -0.018 (0.026) | 0.031 (0.038) | -0.003 (0.034) |
| log_casi4 | -0.031 (0.046) | -0.023 (0.030) | 0.020 (0.055) | -0.024 (0.032) |
| Observations | 21,812 | 21,811 | 21,857 | 21,861 |
| R-squared | 0.486 | 0.855 | 0.830 | 0.513 |

Robust standard errors
in parentheses

*** p<0.01, ** p<0.05,

* p<0.1

Tabella 6 - Traffico fisso: interazione tra numero di casi e periodo di appartenenza

Seconda specifica: differenza nel numero di casi

La stessa analisi è stata ripetuta per la seconda specifica, che sostituiva al totale assoluto dei casi, la differenza tra quelli di un giorno e quelli del giorno precedente.

Il traffico in Streaming ha un'evidente correlazione negativa con la variabile d'interesse, sia sulla rete mobile (in Tabella 7, dove ha una significatività dell'1%), sia su rete fissa (in Tabella 8, dove è sempre significativo, ma al 10%).

Il traffico su Social Network ha una correlazione negativa con la differenza di casi nel primo periodo e positiva nel secondo, su rete mobile, mentre su rete fissa non si evidenziano correlazioni.

La tipologia Instant Messaging comprende non solo i servizi di messaggistica istantanea come Whatsapp o Telegram, ma anche tutti i principali canali di VoIP. I *Voice over IP* sono applicazioni come Skype, Zoom o Teams, che permettono di effettuare tramite rete Internet una conversazione normalmente eseguita via telefono. L'utilizzo di tali piattaforme è aumentato in modo esponenziale durante i mesi della pandemia e a conferma di ciò si può osservare come sì, le correlazioni con le variabili di interesse nel secondo e nel quarto periodo siano negative su mobile, ma sul fisso esse siano positive e altamente statisticamente significative in ognuno dei quattro periodi: è evidente quindi un effetto sostituzione tra le due tipologie di connessione. Prendendo a titolo di esempio il secondo periodo, a fronte di un incremento del 10% nella differenza dei casi tra due giorni, si ha una diminuzione dello 0,9% nel traffico su mobile, ma un aumento dell'1,89% su rete fissa.

Sul Browsing, invece, solamente nel caso del terzo periodo la variabile indipendente ha un p-value inferiore al 10%, presentando una correlazione negativa.

| VARIABLES | (1) Log_Streaming | (2) Log_Social | (3) Log_IM | (4) Log_Browsing |
|--------------|----------------------|----------------------|----------------------|---------------------|
| log_casi1 | -0.092*** (0.035) | -0.089*** (0.028) | 0.029 (0.020) | -0.042 (0.035) |
| log_casi2 | -0.041 (0.028) | 0.050* (0.028) | -0.092*** (0.034) | 0.010 (0.032) |
| log_casi3 | 0.004 (0.019) | 0.004 (0.020) | 0.000 (0.015) | -0.033* (0.017) |
| log_casi4 | 0.010 (0.020) | -0.024 (0.039) | -0.053*** (0.017) | 0.011 (0.033) |
| Observations | 23,687 | 23,687 | 23,688 | 23,688 |
| R-squared | 0.114 | 0.087 | 0.056 | 0.075 |

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Tabella 7 - Traffico mobile: interazione tra la differenza del numero di casi e periodo di appartenenza

| VARIABLES | (1) Log_Streaming | (2) Log_Social | (3) Log_IM | (4) Log_Browsing |
|--------------|----------------------|-------------------|---------------------|---------------------|
| log_casi1 | -0.031* (0.018) | 0.021 (0.018) | 0.679*** (0.030) | 0.001 (0.015) |
| log_casi2 | -0.021 (0.023) | -0.024 (0.017) | 0.189*** (0.027) | -0.010 (0.018) |
| log_casi3 | 0.003 (0.007) | -0.005 (0.009) | 0.123*** (0.017) | -0.001 (0.006) |
| log_casi4 | -0.006 (0.020) | -0.005 (0.014) | 0.339*** (0.016) | -0.018 (0.019) |
| Observations | 21,812 | 21,811 | 21,857 | 21,861 |
| R-squared | 0.486 | 0.855 | 0.616 | 0.514 |

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Tabella 8 - Traffico fisso: interazione tra la differenza del numero di casi e periodo di appartenenza

3.4 Interazioni con le zone d'Italia

In seguito ai risultati ottenuti dalle statistiche descrittive riportate nel Capitolo 2, si è ritenuto necessario effettuare anche un'analisi su base geografica, che tenesse conto della diversa appartenenza delle singole province alle tre zone di Nord, Centro o Sud Italia.

Prima specifica: numero totale di casi

Analogamente a quanto fatto nei paragrafi precedenti, è stata creata una dummy per ogni zona d'Italia e tutte e tre sono state interagite con il log del totale dei casi:

$$\log y_{t,p} = \beta_0 + \beta_1 \log C_{t,p} \times D_{zona} + \delta_t + \delta_p + \varepsilon_{t,p} \quad (3.6)$$

Su rete mobile (Tabella 9), gli unici risultati significativi al 5% sono stati ottenuti in corrispondenza della variabile interagita tra il numero di casi e le zone del Nord e Sud Italia: entrambe le volte è stata ottenuta una correlazione positiva per il traffico medio su Instant Messaging e negativa su Browsing.

| VARIABLES | (1) Log_Streaming | (2) Log_Social | (3) Log_IM | (4) Log_Browsing |
|----------------|----------------------|-------------------|--------------------|---------------------|
| log_casiNord | 0.057 (0.079) | 0.035 (0.117) | 0.207** (0.100) | -0.274** (0.121) |
| log_casiCentro | -0.074 (0.062) | -0,137 (0.125) | 0.130 (0.095) | -0.089 (0.103) |
| log_casiSud | -0.076 (0.057) | 0,029 (0.079) | 0.120** (0.059) | -0.162* (0.088) |
| Observations | 23,687 | 23,688 | 23,688 | 23,688 |
| R-squared | 0.114 | 0.999 | 0.073 | 0.077 |

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Tabella 9 - Traffico mobile: interazione tra numero di casi e zona italiana di appartenenza

Su rete fissa (Tabella 10), invece, solo considerando la zona del Sud Italia si ottiene un risultato statisticamente significativo al 5% nel traffico su Instant Messaging.

| VARIABLES | (1) Log_Streaming | (2) Log_Social | (3) Log_IM | (4) Log_Browsing |
|----------------|----------------------|-------------------|--------------------|---------------------|
| log_casiNord | -0.011 (0.060) | -0.032 (0.034) | 0.015 (0.064) | -0.021 (0.039) |
| log_casiCentro | -0.077 (0.133) | -0.066 (0.078) | -0.037 (0.158) | 0.008 (0.109) |
| log_casiSud | -0.006 (0.022) | -0.005 (0.019) | 0.069** (0.029) | -0.006 (0.023) |
| Observations | 21,812 | 21,811 | 21,857 | 21,861 |
| R-squared | 0.487 | 0.855 | 0.830 | 0.513 |

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Tabella 10 - Traffico fisso: interazione tra numero di casi e la zona italiana di appartenenza

Seconda specifica: differenza nel numero di casi

Anche in questo caso l'analisi è stata ripetuta per la seconda specifica in esame.

Solo nel traffico in Streaming su mobile (Tabella 11), è stato possibile osservare una correlazione con la differenza dei casi nel Centro e Sud Italia, con una significatività, rispettivamente, del 5% e 10%.

Sulla linea fissa (Tabella 12), invece, è stata rilevata una correlazione con la variabile di interesse solo nel traffico di Instant Messaging nella zona del Sud Italia.

| VARIABLES | (1) Log_Streaming | (2) Log_Social | (3) Log_IM | (4) Log_Browsing |
|----------------|----------------------|-------------------|------------------|---------------------|
| log_casiNord | 0.011 (0.012) | -0.000 (0.019) | 0.030 (0.022) | -0.024 (0.020) |
| log_casiCentro | -0.053** (0.021) | -0.049 (0.036) | 0.015 (0.020) | 0.016 (0.028) |
| log_casiSud | -0.033* (0.017) | 0.004 (0.027) | 0.015 (0.018) | -0.027 (0.022) |
| Observations | 23,687 | 23,688 | 23,688 | 23,688 |
| R-squared | 0.114 | 0.836 | 0.073 | 0.075 |

Robust

standard errors

in parentheses

*** p<0.01, **

p<0.05, * p<0.1

Tabella 11 - Traffico mobile: interazione tra la differenza nel numero di casi giornalieri e la zona italiana di appartenenza

| VARIABLES | (3) Log_Streaming | (3) Log_Social | (3) Log_IM | (3) Log_Browsing |
|----------------|----------------------|-------------------|--------------------|---------------------|
| log_casiNord | -0.005 (0.012) | -0.001 (0.011) | -0.003 (0.019) | -0.008 (0.012) |
| log_casiCentro | -0.023 (0.037) | -0.019 (0.020) | -0.018 (0.046) | -0.001 (0.023) |
| log_casiSud | -0.005 (0.008) | 0.001 (0.008) | 0.024** (0.010) | -0.006 (0.010) |
| Observations | 21,812 | 21,867 | 21,857 | 21,861 |
| R-squared | 0.486 | 0.840 | 0.830 | 0.513 |

Robust standard errors in parentheses

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

Tabella 12 - Traffico fisso: interazione tra la differenza nel numero di casi giornalieri e la zona italiana di appartenenza

3.5 Interazione con le province più colpite

Si è deciso di effettuare un'ulteriore distinzione tra le province più colpite e tutte le altre, al fine di focalizzare maggiormente l'attenzione sulla possibile differenza in termini di consumo tra le due tipologie.

È stata definita una dummy indicante l'appartenenza della provincia al gruppo delle più colpite, creato a partire dalle suddivisioni effettuate nel Capitolo 2: sono state selezionate le province appartenenti al top 5% nella distribuzione dell'indice di contagiosità (2.3) per ognuna delle tre zone. Tale dummy è poi stata interagita con la variabile log_casi:

$$\log y_{t,p} = \beta_0 + \beta_1 \log C_{t,p} \times D_{colpita} + \delta_t + \delta_p + \varepsilon_{t,p} \quad (3.6)$$

Prima specifica: numero totale di casi

I risultati ottenuti da questa analisi finale sono senza dubbio i più importanti.

Sono emerse, infatti, su rete mobile, evidenti correlazioni tra tutte le quattro categorie di traffico in esame e la variabile interagita con l'insieme delle province più colpite (Tabella 13):

- Nel caso del traffico in Streaming, che è l'unico in cui è presente una correlazione anche per tutte le altre province, la variabile interagita mostra un effetto del -0,567% sulla variabile dipendente. In valore assoluto, l'effetto catturato nelle province più colpite, è maggiore delle altre: a fronte di un aumento del 10% dei casi nelle prime, infatti, si è ottenuta una diminuzione del 5,67% di traffico medio, e solo dell'1,43% nelle altre;
- In corrispondenza del traffico su Social Network è presente un valore di 0,397, con una significatività del 5%;
- Nel caso del traffico su Instant Messaging, si è ottenuto un coefficiente negativo, ma anche qui è emerso un effetto sostituzione con il traffico fisso (Tabella 14), dove si è ottenuta una variazione dello 0,214 %;
- Il traffico in Browsing è positivamente correlato.

| VARIABLES | (1) Log_Streaming | (2) Log_Social | (3) Log_IM | (4) Log_Browsing |
|----------------------|----------------------|--------------------|--------------------|---------------------|
| log_casi_piu_colpite | -0.567*** (0.120) | 0.397** (0.194) | -0.372* (0.195) | 0.388** (0.173) |
| log_casi_altre | -0.143*** (0.053) | 0.020 (0.089) | 0.066 (0.060) | -0.080 (0.087) |
| Observations | 23,687 | 23,687 | 23,688 | 23,688 |
| R-squared | 0.114 | 0.086 | 0.074 | 0.076 |

Robust standard
errors in parentheses
*** p<0.01, **
p<0.05, * p<0.1

Tabella 13 - Traffico mobile: interazione tra il numero di casi e l'appartenenza all'insieme delle province più colpite

| VARIABLES | (1) Log_Streaming | (2) Log_Social | (3) Log_IM | (4) Log_Browsing |
|----------------------|----------------------|-------------------|--------------------|---------------------|
| log_casi_piu_colpita | 0.052 (0.071) | 0.021 (0.053) | 0.214** (0.090) | 0.079 (0.070) |
| log_casi_altre | -0.015 (0.040) | -0.009 (0.027) | 0.070 (0.045) | 0.007 (0.037) |
| Observations | 21,812 | 21,811 | 21,857 | 21,861 |
| R-squared | 0.485 | 0.855 | 0.830 | 0.513 |

Robust standard errors in parentheses

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

Tabella 14 - Traffico fisso: interazione tra il numero di casi e l'appartenenza all'insieme delle province più colpite

Seconda specifica: differenza nel numero di casi

Anche nel caso della seconda specifica sono emersi risultati interessanti per quanto riguarda le province più colpite: la variabile interagita presentava infatti, su mobile, una significatività dell'1% con correlazione negativa nei casi di traffico in Streaming e Instant Messaging, mentre nel caso di Browsing la correlazione è risultata positiva (Tabella 15).

Sulla rete fissa, invece, non sono emersi risultati statisticamente significativi (Tabella 16).

| VARIABLES | (1) Log_Streaming | (2) Log_Social | (3) Log_IM | (4) Log_Browsing |
|----------------------|----------------------|-------------------|----------------------|---------------------|
| log_casi_piu_colpita | -0.121*** (0.039) | 0.033 (0.022) | -0.065*** (0.022) | 0.133*** (0.025) |
| log_casi_altre | -0.013 (0.012) | -0.011 (0.016) | 0.025 (0.016) | -0.023 (0.014) |
| Observations | 23,687 | 23,687 | 23,688 | 23,688 |
| R-squared | 0.114 | 0.085 | 0.074 | 0.078 |

Robust standard errors in parentheses

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

Tabella 15 - Traffico mobile: interazione tra la differenza giornaliera dei casi e l'appartenenza all'insieme delle province più colpite

| VARIABLES | (1) Log_Streaming | (2) Log_Social | (3) Log_IM | (4) Log_Browsing |
|----------------------|----------------------|-------------------|------------------|---------------------|
| log_casi_piu_colpita | -0.005 (0.009) | 0.003 (0.009) | 0.010 (0.011) | 0.003 (0.014) |
| log_casi_altre | -0.009 (0.009) | -0.004 (0.008) | 0.002 (0.014) | -0.006 (0.007) |
| Observations | 21,812 | 21,811 | 21,857 | 21,861 |
| R-squared | 0.485 | 0.855 | 0.830 | 0.513 |

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Tabella 16 - Traffico fisso: interazione tra la differenza giornaliera dei casi e l'appartenenza all'insieme delle province più colpite

3.6 Considerazioni di fine capitolo

Confrontando tutte le analisi presentate si può osservare che le due categorie di traffico che più spesso hanno mostrato una correlazione con le variabili indipendenti sono lo Streaming (sempre con segno negativo) e l'Instant Messaging (con segno positivo in quanto, come si è visto, anche nei casi in cui esso risultava negativo era compensato da un effetto sostituzione su rete fissa).

Sicuramente gran parte dell'utilizzo delle applicazioni di Messagistica Istantanea è dovuto allo *smartworking* e alla didattica a distanza, però è pur vero che si sono riscontrati risultati significativi anche nei periodi estivi.

Sembra quindi emergere che, indistintamente dall'appartenenza alle province più o meno colpite e dalle restrizioni vigenti in un determinato periodo, il Covid-19 abbia portato a una crescita del loro consumo, mentre abbia sfavorito l'utilizzo delle piattaforme di Streaming.

Tutto ciò può essere facilmente spiegato: in un periodo ricco di paura e incertezza gli utenti hanno preferito applicazioni che permettessero loro di connettersi con altri, piuttosto che isolarsi ulteriormente vedendo film e serie tv in Streaming.

Conclusioni

Lo scopo di questo lavoro di tesi era quello di indagare l'impatto del Covid-19 sui consumi di Internet in Italia e, al fine di raggiungere l'obiettivo, sono stati inizialmente ricercati studi precedenti inerenti allo stesso ambito, così da poter ottenere una solida base da cui partire e approfondire le tante sfaccettature del mondo ICT nel periodo esaminato.

Il dataset utilizzato conteneva un numero significativo di record e variabili, perciò sarebbe stato possibile compiere studi differenti da quelli effettuati: numerose sarebbero state le possibili combinazioni tra le fonti di traffico e l'andamento della pandemia. Per esempio, una delle ipotesi iniziali era quella di utilizzare il tempo medio trascorso dagli utenti sulle diverse tipologie di traffico, anziché considerare i consumi in termini di bps su download e upload oppure considerare un'altra variabile anziché il numero di casi di una determinata provincia, focalizzandosi per esempio sul numero di ricoverati o di guariti.

Molte erano le dimensioni in cui tale dataset avrebbe potuto essere letto, ma in questo elaborato si è deciso di porre una maggiore attenzione su quelli che sono stati due dei capisaldi del periodo Covid-19 in Italia: le restrizioni e gli allentamenti nei diversi periodi e l'aspetto geografico, che ha visto una netta divisione tra nord e sud in termini di contagi.

L'analisi di regressione finale, a conclusione dell'elaborato, ha dimostrato come sia il singolo numero dei contagi sia il suo incremento giornaliero abbiano avuto impatto sul consumo delle diverse fonti ICT, su mobile e fisso.

I risultati ottenuti dalle analisi in cui è stato considerato il numero di casi in un determinato periodo mostrano risultati interessanti solo per quanto riguarda il traffico in Instant Messaging e Streaming (con relazione rispettivamente positiva e negativa) sulla rete mobile, mentre sul fisso non sono emersi dati statisticamente significativi: sembra quindi che il consumo di Internet sia guidato, più che dal numero di casi, dalle diverse restrizioni introdotte nei vari periodi.

Il risultato davvero importante è stato ottenuto quando si è considerato l'aspetto geografico: se inizialmente con la suddivisione tra Nord, Centro e Sud Italia non erano stati riscontrati risultati degni di nota, altrettanto non si può dire con l'introduzione della distinzione tra le province più colpite dal Covid-19 e le altre. In tutte le tipologie di traffico,

infatti, sono emersi risultati altamente statisticamente significativi nel caso in cui le province appartenessero all'insieme delle più colpite, fatto che invece non è accaduto per tutte le altre. In particolare, sembra lecito asserire che il consumo di Streaming sia negativamente correlato con il numero di casi in tali province, probabilmente per l'*effetto paura* oppure perché in momenti di particolare criticità le persone preferiscono cercare modi per connettersi con gli altri piuttosto che isolarsi ulteriormente.

A favore di questa ipotesi vi è il fatto che gli utenti sembrano spostare i loro consumi sui Social Network o sulle applicazioni di Instant Messaging. Anche la correlazione positiva con il Browsing non deve stupire: un aumento di contagi provoca un incremento nella domanda di informazioni, che possono essere ottenute in maniera mirata sui motori di ricerca.

Appendice

Sono di seguito riportati alcuni degli output ottenuti con le analisi di regressione inizialmente effettuate, che poi sono state scartate in fase di definizione del modello.

Regressione

$$y_{t,p} = \beta_0 + \beta_1 C_{t,p} + \beta_2 C_{t,p}^2 + \delta_t + \delta_p + \varepsilon_{t,p}$$

I numeri (1), (2), (3) sono relativi ad analisi in cui la variabile dipendente è il numero di casi. In (1) non è presente alcun controllo, in (2) è aggiunta la dummy temporale e in (3) è aggiunta anche la dummy delle province.

I numeri (4), (5) e (6) sono invece relativi ad analisi in cui è stata anche introdotta la variabile del numero di casi al quadrato. Come sopra, sono state introdotte gradatamente le variabili di controllo: in (4) non c'è nessun controllo, in (5) si introduce quello sulle date e in (6) anche quello sulle province).

| VARIABLES | (1) Streaming | (2) Streaming | (3) Streaming | (4) Streaming | (5) Streaming | (6) Streaming |
|--------------------|---------------------------|-----------------------------|--------------------------|-----------------------------|-----------------------------|----------------------------|
| TotaleCasi | -8,127.253 (6,169.453) | -11,744.554* (6,806.570) | 3,329.089 (5,047.180) | -17,102.835 (12,680.459) | -22,761.895 (14,557.722) | 15,736.214 (16,503.994) |
| TotaleCasiQuadrato | | | | 0.128 (0.094) | 0.146 (0.104) | -0.108 (0.107) |
| Observations | 23,687 | 23,687 | 23,687 | 23,687 | 23,687 | 23,687 |
| R-squared | 0.001 | 0.071 | 0.201 | 0.001 | 0.072 | 0.201 |

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Tabella 17 – Streaming

| VARIABLES | (1) Social Network | (2) Social Network | (3) Social Network | (4) Social Network | (5) Social Network | (6) Social Network |
|--------------------|--------------------------|------------------------------|----------------------------|----------------------------|------------------------------|-----------------------------|
| TotaleCasi | 9,762.678 (7,697.438) | -16,045.357** (6,850.245) | -7,362.647 (18,866.106) | 16,009.667 (19,692.550) | -34,186.838* (17,679.484) | -27,918.652 (43,113.668) |
| TotaleCasiQuadrato | | | | -0.089 (0.194) | 0.240 (0.166) | 0.178 (0.255) |
| Observations | 23,687 | 23,687 | 23,687 | 23,687 | 23,687 | 23,687 |
| R-squared | 0.001 | 0.064 | 0.141 | 0.001 | 0.066 | 0.141 |

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Tabella 18 - Social Network

| VARIABLES | (1) IM | (2) IM | (3) IM | (4) IM | (5) IM | (6) IM |
|--------------------|------------------------------|---------------------------|-----------------------------|-----------------------------|----------------------------|----------------------------|
| TotaleCasi | -14,491.205** (6,042.774) | -2,013.082 (4,853.296) | -12,437.156 (13,578.301) | -13,323.064 (15,758.935) | 11,252.219 (14,535.316) | 13,242.900 (35,689.354) |
| TotaleCasiQuadrato | | | | -0.017 (0.141) | -0.175 (0.128) | -0.223 (0.209) |
| Observations | 23,688 | 23,688 | 23,688 | 23,688 | 23,688 | 23,688 |
| R-squared | 0.002 | 0.044 | 0.124 | 0.002 | 0.044 | 0.124 |

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Tabella 19 - Instant Messaging

| VARIABLES | (1) Browsing | (2) Browsing | (3) Browsing | (4) Browsing | (5) Browsing | (6) Browsing |
|--------------------|--------------------------|-----------------------------|----------------------------|----------------------------|---------------------------|-----------------------------|
| TotaleCasi | 7,743.302 (7,316.294) | 14,665.541** (6,823.739) | 13,169.547 (10,547.907) | -6,846.590 (14,839.020) | 3,114.337 (14,868.551) | -18,521.140 (32,906.078) |
| TotaleCasiQuadrato | | | | 0.208 (0.126) | 0.153 (0.117) | 0.275 (0.211) |
| Observations | 23,688 | 23,688 | 23,688 | 23,688 | 23,688 | 23,688 |
| R-squared | 0.000 | 0.016 | 0.121 | 0.001 | 0.017 | 0.122 |

Robust standard
errors in parentheses

*** p<0.01, **

p<0.05, * p<0.1

Tabella 20 - Browsing

Regressione

$$y_{t,p} = \beta_0 + \beta_1 C_{t,p} * D_{periodo} + \beta_2 C_{t,p}^2 * D_{periodo} + \delta_t + \delta_p + \varepsilon_{t,p}$$

Interazione delle due variabili (casi giornalieri e il loro quadrato) con i 4 periodi identificati.

Di seguito è riportato l'esempio del traffico in Streaming con le medesime numerazioni del caso precedente:

| VARIABLES | (1) Streaming | (2) Streaming | (3) Streaming | (4) Streaming | (5) Streaming | (6) Streaming |
|---------------------|--------------------------------|--------------------------------|-----------------------------|---------------------------------|---------------------------------|---------------------------------|
| TotaleCasi1 | 126,728.902** (54,075.685) | -80,006.745 (75,904.962) | -78,852.347 (74,780.538) | 232,814.431* (119,343.513) | -446,375.899** (176,030.704) | -488,749.750** (186,587.987) |
| TotaleCasi2 | 1,418.707 (39,782.365) | -59,531.116 (51,771.098) | -58,855.523 (54,539.709) | -25,782.047 (62,686.773) | -256,361.230*** (93,275.825) | -282,797.175*** (99,252.316) |
| TotaleCasi3 | -68,958.391*** (16,346.556) | -35,002.362*** (12,671.786) | -33,970.989 (21,300.728) | -131,791.396*** (42,782.326) | -28,977.570 (41,747.470) | -57,464.527 (61,277.884) |
| TotaleCasi4 | -814.940 (4,923.462) | -2,826.751 (4,634.826) | -5,069.085 (5,657.558) | -433.289 (10,875.756) | 4,349.792 (11,055.248) | -6,285.850 (16,853.743) |
| TotaleCasiQuadrato1 | | | | -10.483 (7.246) | 28.005** (11.466) | 31.810** (12.309) |
| TotaleCasiQuadrato2 | | | | 1.555 (3.085) | 12.053** (4.797) | 14.097*** (5.087) |
| TotaleCasiQuadrato3 | | | | 3.634* (1.874) | -0.315 (1.690) | 1.526 (2.017) |
| TotaleCasiQuadrato4 | | | | -0.025 (0.085) | -0.078 (0.082) | 0.042 (0.109) |
| Observations | 23,687 | 23,687 | 23,687 | 23,687 | 23,687 | 23,687 |
| R-squared | 0.013 | 0.074 | 0.202 | 0.016 | 0.081 | 0.207 |

Robust standard errors in parentheses
 *** p<0.01, ** p<0.05, * p<0.1

Tabella 21 - Streaming; Interazione con i periodi

Regressione

$$y_{t,p} = \beta_0 + \beta_1 C_{t,p} + \beta_2 C_{t,p}^2 + \delta_t + \delta_r + \varepsilon_{t,p}$$

In questo caso il controllo sulla provincia è stato sostituito da una dummy a livello regionale. In entrambe le province è stata considerata direttamente il modello con più restrizioni.

| VARIABLES | (1) Streaming | (2) Streaming |
|--------------------|--------------------------|----------------------------|
| TotaleCasi | 4,769.115 (7,349.594) | 14,727.670 (15,581.865) |
| TotaleCasiQuadrato | | -0.116 (0.111) |
| Observations | 23,687 | 23,687 |
| R-squared | 0.127 | 0.127 |

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Tabella 22 - Streaming; controllo sulla regione.

Regressione

Gli output di seguito riportati sono stati ottenuti dalla stessa analisi vista in precedenza, ma anziché considerare il totale dei casi è stato considerato il totale dei casi del giorno precedente.

| VARIABLES | (1) Streaming | (2) Streaming | (3) Streaming | (4) Streaming | (5) Streaming | (6) Streaming |
|------------------|---------------------------|-----------------------------|--------------------------|-----------------------------|-----------------------------|----------------------------|
| Lag_casi | -8,654.178 (6,344.577) | -12,548.341* (6,997.493) | 3,840.422 (5,026.898) | -18,265.446 (12,970.041) | -24,376.821 (14,835.009) | 16,185.802 (16,694.887) |
| Lag_casiQuadrato | | | | 0.139 (0.096) | 0.159 (0.107) | -0.108 (0.109) |
| Observations | 23,255 | 23,255 | 23,255 | 23,255 | 23,255 | 23,255 |
| R-squared | 0.001 | 0.070 | 0.202 | 0.001 | 0.070 | 0.202 |

Robust standard errors in parentheses

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

Tabella 23 - Streaming

| VARIABLES | (1) Social Network | (2) Social Network | (3) Social Network | (4) Social Network | (5) Social Network | (6) Social Network |
|------------------|--------------------------|------------------------------|----------------------------|----------------------------|------------------------------|-----------------------------|
| Lag_casi | 9,622.395 (7,825.968) | -16,368.566** (7,047.837) | -7,895.833 (18,975.740) | 15,470.018 (20,005.975) | -34,503.577* (18,025.375) | -29,813.288 (43,438.444) |
| Lag_casiQuadrato | | | | -0.085 (0.198) | 0.244 (0.171) | 0.192 (0.260) |
| Observations | 23,255 | 23,255 | 23,255 | 23,255 | 23,255 | 23,255 |
| R-squared | 0.001 | 0.064 | 0.142 | 0.001 | 0.065 | 0.143 |

Robust standard

errors in

parentheses

*** p<0.01, **

p<0.05, * p<0.1

Tabella 24 - Social Network

| VARIABLES | (1) IM | (2) IM | (3) IM | (1) IM | (2) IM | (3) IM |
|------------------|------------------------------|---------------------------|-----------------------------|-----------------------------|----------------------------|----------------------------|
| Lag_casi | -14,491.205** (6,042.774) | -2,013.082 (4,853.296) | -12,437.156 (13,578.301) | -13,323.064 (15,758.935) | 11,252.219 (14,535.316) | 13,242.900 (35,689.354) |
| Lag_casiQuadrato | | | | -0.017 (0.141) | -0.175 (0.128) | -0.223 (0.209) |
| Observations | 23,688 | 23,688 | 23,688 | 23,688 | 23,688 | 23,688 |
| R-squared | 0.002 | 0.044 | 0.124 | 0.002 | 0.044 | 0.124 |

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Tabella 25 - Instant Messaging

| VARIABLES | (1) Browsing | (2) Browsing | (3) Browsing | (1) Browsing | (2) Browsing | (3) Browsing |
|------------------|--------------------------|-----------------------------|----------------------------|----------------------------|---------------------------|-----------------------------|
| Lag_casi | 7,743.302 (7,316.294) | 14,665.541** (6,823.739) | 13,169.547 (10,547.907) | -6,846.590 (14,839.020) | 3,114.337 (14,868.551) | -18,521.140 (32,906.078) |
| Lag_casiQuadrato | | | | 0.208 (0.126) | 0.153 (0.117) | 0.275 (0.211) |
| Observations | 23,688 | 23,688 | 23,688 | 23,688 | 23,688 | 23,688 |
| R-squared | 0.000 | 0.016 | 0.121 | 0.001 | 0.017 | 0.122 |

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Tabella 26 - Browsing

Indice delle tabelle

| | |
|--|-----------|
| <i>Tabella 1- Indice di contagiosità per ogni provincia.....</i> | <i>39</i> |
| <i>Tabella 2 – Indice di contagiosità per ogni provincia del Nord Italia</i> | <i>47</i> |
| <i>Tabella 3 - Indice di contagiosità per ogni provincia del Centro Italia</i> | <i>50</i> |
| <i>Tabella 4 - Indice di contagiosità per ogni provincia del Centro Italia</i> | <i>54</i> |
| <i>Tabella 5 - Traffico mobile: interazione tra numero di casi e periodo di appartenenza</i> | <i>64</i> |
| <i>Tabella 6 - Traffico fisso: interazione tra numero di casi e periodo di appartenenza</i> | <i>64</i> |
| <i>Tabella 7 - Traffico mobile: interazione tra la differenza del numero di casi e periodo di appartenenza.....</i> | <i>66</i> |
| <i>Tabella 8 - Traffico fisso: interazione tra la differenza del numero di casi e periodo di appartenenza.....</i> | <i>66</i> |
| <i>Tabella 9 - Traffico mobile: interazione tra numero di casi e zona italiana di appartenenza</i> | <i>68</i> |
| <i>Tabella 10 - Traffico fisso: interazione tra numero di casi e la zona italiana di appartenenza.....</i> | <i>68</i> |
| <i>Tabella 11 - Traffico mobile: interazione tra la differenza nel numero di casi giornalieri e la zona italiana di appartenenza</i> | <i>69</i> |
| <i>Tabella 12 - Traffico fisso: interazione tra la differenza nel numero di casi giornalieri e la zona italiana di appartenenza</i> | <i>70</i> |
| <i>Tabella 13 - Traffico mobile: interazione tra il numero di casi e l'appartenenza all'insieme delle province più colpite</i> | <i>71</i> |
| <i>Tabella 14 - Traffico fisso: interazione tra il numero di casi e l'appartenenza all'insieme delle province più colpite</i> | <i>72</i> |
| <i>Tabella 15 - Traffico mobile: interazione tra la differenza giornaliera dei casi e l'appartenenza all'insieme delle province più colpite.....</i> | <i>72</i> |
| <i>Tabella 16 - Traffico fisso: interazione tra la differenza giornaliera dei casi e l'appartenenza all'insieme delle province più colpite.....</i> | <i>73</i> |
| <i>Tabella 17 – Streaming.....</i> | <i>76</i> |
| <i>Tabella 18 - Social Network</i> | <i>77</i> |
| <i>Tabella 19 - Instant Messaging</i> | <i>77</i> |
| <i>Tabella 20 - Browsing</i> | <i>78</i> |

| | |
|---|-----------|
| <i>Tabella 21 - Streaming; Interazione con i periodi.....</i> | <i>79</i> |
| <i>Tabella 22 - Streaming; controllo sulla regione.....</i> | <i>80</i> |
| <i>Tabella 23 - Streaming.....</i> | <i>81</i> |
| <i>Tabella 24 - Social Network</i> | <i>81</i> |
| <i>Tabella 25 - Instant Messaging</i> | <i>82</i> |
| <i>Tabella 26 - Browsing</i> | <i>82</i> |

Bibliografia

- Abrahams, S., AlFakhri, M., Hotz, V. J., Schoeni, R. F., Seltzer, J. A., & Wiemers, E. E. (2020). *Disparities in vulnerability to severe complications from COVID-19 in the United States*. National Bureau of Economic Research.
- AGCOM (2020). *Le comunicazioni nel 2020*. www.agcom.it
- Askitas, N., Tatsiramos, K., & Verheyden, B. (2020). *Lockdown Strategies, Mobility Patterns and COVID-19*. CESifo Working Papers.
- Barrios, J. M., Benmelech, E., Hochberg, Y. V., Sapienza, P., & Zingales, L. (2020). *Civic Capital and Social Distancing during the COVID-19 pandemic*. National Bureau of Economic Research.
- Bartscher, A. K., Seitz, S., Siegloch, S., Slotwinski, M., & Wehrhöfer, N. (2020). *Social Capital and the Spread of COVID-19: Insights from European Countries*. CESifo Working Papers.
- Biswas, S., Dubè, J. H., Sacher, S. K., & Simonov, A. (2020). *The persuasive effect of Fox News: non-compliance with social distancing during the COVID-19 pandemic*. National Bureau of Economic Research.
- Blumenshine, P., Reingold, A., Ergert, S., Mockenhaupt, R., Braveman, P., & Marks, J. (2008). Pandemic Influenza Planning in the United States from a Health Disparities Perspective. *Emerging Infectious Diseases* 14(5), 709-15.
- Bonezzi, A., Brendl, C.M., & De Angelis, M. (2011). *Stuck in the middle: The psychophysics of goal pursuit*. *Psychological science*, 22(5), pp.607-612.
- Briscese, G., Lacetera, N., Macis, M., & Tonin, M. (2020). *Compliance with COVID-19 social-distancing measures in Italy: the role of expectations and duration*. National Bureau of Economic Research.
- Campos-Urbina, A. M., Carranza Esteban, R. F., Catay-Medina, J. B., Garayar-Peceros, H., Ignacio-Quinte, C., Mejia, C. R., Porta-Quinto, T., Rodriguez-Alarcon, J.F., Ruiz Mamani, P. G., Ticona, D., & Tovani-Palone, M. R. (2020). *The Media and their*

Informative Role in the Face of the Coronavirus Disease 2019 (COVID-19): Validation of Fear Perception and Magnitude of the Issue (MED-COVID-19). Electronic Journal of General Medicine.

Castriota, S., Delmastro, M., Tonin, M. (2020). *National or Local? The Demand for News in Italy during Covid-19*. CESifo Working Papers.

Chakravarthy, K., Chaturvedi, R., & Williams, C. M. (2020). *Cybersecurity Risks in a Pandemic*. JOURNAL OF MEDICAL INTERNET RESEARCH.

Chiou, L., & Tucker, C. (2020). *Social Distancing, internet access and inequality*. National Bureau of Economic Research.

Couch, K. A., Fairlie, R. W., & Xu, H. (2020). *The Impacts of COVID-19 on Minority Unemployment: First Evidence from April 2020 CPS Microdata*. CESifo Working Papers.

Durante, R., Guiso, L., & Gulino, G. (2020). *Asocial Capital: Civic Culture and Social Distancing during COVID-19*.

Ecarma, C. (2020). *Fox News Is Preparing to Be Sued Over Coronavirus Misinformation*. Vanity Fair.

Istat (2020). *Impatto dell'epidemia COVID-19 sulla mortalità totale della popolazione residente - periodo gennaio-novembre 2020*.

Knack, S., & Keefer, P. (1997). *Does Social Capital Have an Economic Payoff? A Cross-Country Investigation*. Quarterly Journal of Economics 112.4, pp. 1251-1288.

Kuchler, T., Russel, D., & Stroebel, J. (2020). *The geographic spread of COVID-19 correlates with structure of social networks as measured by Facebook*. National Bureau of Economic Research.

Mancini, E., & Scocchi, N. (2020). *E con il Coronavirus anche un'epidemia di fake news*. lavoce.info.

Miller, A. R., & Tucker, C. E. (2011). *Can health care information technology save babies?* Journal of Political Economy 119 (2), 289-324.

- Murugesan, S., & Weil, T. (2020). *IT Risk and Resilience-Cybersecurity Response to COVID-19*. Institute of Electrical and Electronics Engineers (IEEE 2020).
- Putnam, Robert D. (1993). *Making Democracy Work: Civic Traditions in Modern Italy*. Princeton University Press, Princeton, NJ.
- Schmitt-Grohé, S., Teoh, K., & Uribe, M. (2020). *COVID-19: testing inequality in New York city*. National Bureau of Economic Research.
- Stacey, K. (2020). *Coronavirus exposes America's broadband problem*. Financial Times.
- Tabellini, Guido (2010). *Culture and Institutions: Economic Development in the Regions of Europe*. In: *Journal of the European Economic Association* 8.4, pp. 677–716.
- Tucker, C. E., & Yu, S. (2019). *Does IT Lead to More Equal Treatment? An Empirical Study of the Effect of Smartphone Use on Customer Complaint Resolution*. Mimeo, MIT.
- Wirth, A. (2020). *Electronic Journal of General Medicine*. The Association for the Advancement of Medical Instrumentation (AAMI 2020).

Sitografia

<https://www.corrierecomunicazioni.it/media/fake-news/le-notizie-al-tempo-del-covid-19-5-su-100-sono-fake-news/>

<https://www.pwc.co.uk/issues/crisis-and-resilience/covid-19/how-fake-news-has-exploited-covid19-cyber.html>

<https://www.pandasecurity.com/mediacenter/news/covid-cybersecurity-statistics/>

<https://www.capgemini.com/it-it/news/capgemini-report-the-great-digital-divide/>

<https://theconversation.com/technology-is-improving-why-is-rural-broadband-access-still-a-problem-60423>

<https://www.cybersecurity360.it/nuove-minacce/cyber-attacchi-al-tempo-della-covid-19-i-target-e-le-azioni-di-contrasto/>

https://www.ilsole24ore.com/art/pil-2020-rimbalzo-terzo-trimestre-non-evitera-flessione-1011percento-ADuB0Di?refresh_ce=1

<https://www.ilsole24ore.com/art/la-tv-on-demand-fa-pieno-abbonati-ADMawvs>

https://www.repubblica.it/tecnologia/social-network/2019/01/28/news/fake_news_sui_social_le_condividono_in_pochi_e_soprattutto_over_60_la_conferma_su_twitter-217666294/

<https://www.weforum.org/agenda/2020/04/coronavirus-education-global-covid19-online-digital-learning/>

<https://www.weforum.org/agenda/2020/04/coronavirus-covid-19-pandemic-digital-divide-internet-data-broadband-mobbile/>

<https://www.nytimes.com/interactive/2020/04/07/technology/coronavirus-internet-use.html>

https://www.repubblica.it/tecnologia/sicurezza/2020/04/21/news/_allarme_cyber-attacchi_anche_in_italia_il_covid-19_e_la_nuova_esca_-254604755/

https://www.repubblica.it/tecnologia/sicurezza/2020/04/27/news/coronavirus_l_allarme_dell_oms_aumentati_di_5_volte_i_cyber_attacchi_contro_di_noi_-254994848/

https://www.repubblica.it/tecnologia/social-network/2020/04/30/news/lockdown_covid19_boom_di_smartphone_e_social_e_twitter_offre_i_suoi_dati_alla_ricerca-255286215/

<https://www.lastampa.it/economia/2020/05/22/news/internet-dopo-il-boom-del-lockdown-non-si-sgonfia-il-traffico-della-rete-1.38876713>

https://www.repubblica.it/tecnologia/2020/06/11/news/internet_durante_il_lockdown_30_del_traffico_ma_al_top_anche_attacchi_informatici_e_pirateria-258964450/

<https://www.nytimes.com/2020/06/16/upshot/coronavirus-test-cost-varies-widely.html>

https://www.repubblica.it/tecnologia/social-network/2020/08/12/news/coronavirus_migliaia_di_ricoveri_per_danni_da_fake_news-264473726/

<https://www.nytimes.com/2020/09/10/books/review/brian-stelter-hoax.html>

https://www.repubblica.it/economia/2020/10/02/news/giu_reddito_e_potere_d_acquisto_delle_famiglie_su_il_deficit_l_italia_ai_tempi_del_lockdown-269183930/