



UNIVERSITÀ POLITECNICA DELLE MARCHE
FACOLTÀ DI ECONOMIA “GIORGIO FUÀ”

Corso di Laurea triennale in Economia Aziendale

**UN MODELLO PANEL PER L'EPIDEMIA
COVID-19 IN ITALIA
A PANEL MODEL FOR THE COVID-19
EPIDEMIC IN ITALY**

Relatore:
Prof. Giulio Palomba

Rapporto Finale di:
Silvia D'Amario

Anno Accademico 2020/2021

Indice

CAPITOLO 1. INTRODUZIONE.....	2
CAPITOLO 2. MODELLI EPIDEMIOLOGICI PER L'ITALIA.....	4
2.1 DA CODOGNO AD OGGI: IL VIRUS IN ITALIA	5
2.2 STUDIARE L'EPIDEMIA MEDIANTE MODELLI MATEMATICI	7
2.3 UNO SGUARDO ALLA STORIA DEI MODELLI DEL PASSATO	9
2.4 MODELLI EPIDEMIOLOGICI PER IL COVID-19.....	10
2.4.1 Controllo dell'epidemia con il modello SIDARTHE	10
2.4.2 Modello che include le NPI	12
2.4.3 Un modello previsionale NETWORK-BASED	14
2.4.4 Modello SIR modificato dal DET.....	14
2.4.5 Due modelli che includono l'eterogeneità regionale	15
2.4.6 Un tentativo per arginare l'incertezza dei dati.....	17
2.4.7 L'efficacia di una strategia graduale.....	18
2.4.8 Un modello per l'importanza dell'isolamento	19
CAPITOLO 3. ANALISI EMPIRICA.....	21
3.1 I DATI E LE VARIABILI.....	21
3.2 MODELLO PER LE SERIE STORICHE POOLED	23
3.3 MODELLI PER I DATI LONGITUDINALI.....	27
3.3.1 Il modello ad effetti fissi	27
3.3.2 Il modello ad effetti casuali	29
3.3.3 Inferenza statistica	31
3.4 GLI OSPEDALIZZATI.....	31
3.5 IL TASSO DI POSITIVITÀ.....	32
3.6 REVISIONE DELL'INFORMAZIONE DISPONIBILE.....	33
CAPITOLO 4. CONCLUSIONI.....	35
Bibliografia e Sitografia.....	37
Ringraziamenti.....	39

CAPITOLO 1. INTRODUZIONE

Questo lavoro si colloca nell'ambito dei modelli epidemiologici che descrivono la diffusione di un virus da un punto di vista matematico. La tesi si concentra sullo studio della pandemia mondiale che si è diffusa dagli ultimi mesi del 2019, analizzando modelli già elaborati in letteratura.

COVID-19 è il nome della malattia da nuovo coronavirus: "CO" indica corona, "VI" virus, "D" significa *disease*, 19 fa riferimento all'anno in cui si è manifestata. È chiamato così per le punte a forma di corona che sono presenti sulla superficie del virus.

Con il termine COVID-19 si fa riferimento alla malattia infettiva associata al virus SARS-CoV2, un nuovo ceppo della famiglia dei coronavirus identificato per la prima volta nell'uomo a inizio dicembre 2019. La scoperta è avvenuta nella città di Wuhan, capoluogo con 11 milioni di abitanti della provincia di Hubei in Cina, dove sono apparsi celermente numeri significativi di casi di polmonite dall'origine sconosciuta; il 31 dicembre 2019 la Commissione Sanitaria Municipale di Wuhan ha segnalato tale *cluster* di casi di polmonite a eziologia ignota all'Organizzazione Mondiale della Sanità. Solamente il 3 gennaio 2020 il Chinese Center for Disease Control and Prevention ha dichiarato responsabile di tali polmoniti un nuovo tipo di coronavirus, denominato dalla World Health Organization (WHO) SARS-CoV-2, definendo la malattia associata ad esso "*Coronavirus disease 19*" (COVID-19).

Il 30 gennaio 2020 l'OMS ha dichiarato l'epidemia in Cina emergenza internazionale di salute pubblica, successivamente; l'11 marzo 2020 il direttore generale dell'OMS, ha dichiarato che la diffusione del Covid-19 era ormai diventata una pandemia diffusa in tutto il pianeta.

Fatta questa premessa, una discussione più dettagliata sulla diffusione del virus in Italia e l'utilità dei modelli matematici, sarà compresa nel secondo capitolo, il quale racchiude una panoramica dei modelli epidemiologici già presenti in letteratura. Dopo una breve illustrazione dei modelli per dati i formato panel, nel terzo capitolo è proposto un modello econometrico su dati epidemiologici messi a disposizione dalla Protezione Civile e un'analisi della stima ottenuta. A concludere l'elaborato sarà un conciso sunto sui risultati ottenuti dall'elaborazione a conferma dell'attuale situazione pandemica nel Paese.

CAPITOLO 2. MODELLI EPIDEMIOLOGICI PER L'ITALIA

È passato più di un anno da quando il Coronavirus ha fatto il suo ingresso in Italia: era la sera del 20 febbraio 2020 quando all'ospedale di Codogno si presentava un uomo di 38 anni, Mattia Maestri, con una polmonite grave, con sintomi riconducibili al Covid-19. Il giorno successivo veniva annunciato il paziente 1 in Italia per contagio da Sars-CoV2. L'uomo veniva poi dimesso il 25 marzo, dopo tre settimane di terapia e quattro di ricovero. Da quel momento, in tutto il Paese hanno iniziato ad esplodere i primi focolai che poi hanno portato alla situazione di emergenza ancora in atto ed al primo severo *lockdown*.

La principale modalità di trasmissione del virus è attraverso *droplets*, goccioline di respiro che possono passare da una persona all'altra tramite uno starnuto, un colpo di tosse e contatti personali diretti, ma anche attraverso le mani che, se non lavate possono contaminare e trasmettere il virus tramite il semplice contatto. Essendo troppo pesanti per rimanere nell'aria, le goccioline cadono rapidamente, poggiandosi su superfici e pavimenti. Tuttavia, ad oggi sappiamo che sicuramente il contagio da superfici non è fra le principali vie di trasmissione. I sintomi più diffusi sono febbre pari o superiore a 37,5°C e brividi, tosse, perdita (anosmia) o diminuzione (iposmia) improvvisa dell'olfatto, perdita (ageusia) o alterazione (disgeusia) del gusto, difficoltà respiratorie, mal di gola, mal di testa, naso che cola, affaticamento, debolezza e dolore muscolare. Nei casi più gravi si presenta

in maniera più seria e grave, ossia con polmonite, insufficienza respiratoria, sepsi e shock settico, che possono portare al decesso. Si può essere anche asintomatici, ovvero essere infetti e non mostrare mai sintomi; tali persone sono comunque contagiose e possono trasmettere il virus.

2.1 DA CODOGNO AD OGGI: IL VIRUS IN ITALIA

Dopo la Cina l'Italia è stata la prima a dover fare i conti con il nuovo Coronavirus: diventa il nuovo epicentro del virus che si diffonde in tutte le Regioni, bloccando il movimento di persone e l'economia. È doveroso ripercorrere i principali eventi che hanno caratterizzato il nostro paese a più di un anno dall'inizio dell'epidemia.

Il 30 gennaio 2020 vengono sospesi tutti i voli per la Cina e il giorno seguente l'allora presidente del Consiglio, Giuseppe Conte, conferma i primi due casi di contagio riscontrati in Italia di due turisti cinesi, dichiarando già l'emergenza sanitaria nazionale. Il 21 febbraio 2020 un trentottenne ricoverato all'ospedale di Codogno risultato positivo al Coronavirus: è il paziente 1. In poche ore si riscontra la positività di altre 14 persone e si registra la prima vittima a Vo' Euganeo, in provincia di Padova. Il 23 febbraio 2020 scattano le prime "zone rosse" in 11 comuni tra Lombardia e Veneto, con divieto di accesso o di allontanamento dal territorio comunale, sospensione di eventi e manifestazioni di ogni forma in luogo pubblico o privato. Con il decreto del 9 marzo,

soprannominato *#iorestoacasa*, l'intero Paese finisce in *lockdown*. Il nostro è il primo tra gli stati Occidentali a adottare misure a tal punto restrittive. Il 4 maggio 2020 l'Italia entra nella "Fase 2", con un piano nazionale per le riaperture che prevede un allentamento graduale delle precedenti misure di contenimento, essendo la curva epidemica in fase di discesa. Il nuovo decreto prevede mascherine e distanziamento sociale fino a quando non sarà disponibile il vaccino. Il 18 maggio termina ufficialmente il lockdown, il Paese comincia a respirare. L'11 giugno 2020 apre la "Fase 3", di convivenza con il virus, che allenta ulteriormente i provvedimenti; il DPCM lascia alle regioni la libertà di allentare o restringere ulteriormente le decisioni prese a livello nazionale, o di posticiparle, in base alla situazione epidemiologica dei diversi territori. I mesi estivi rappresentano per milioni di italiani un ritorno alla normalità o quasi: si può andare in vacanza, ci si può spostare tra regioni. Ottobre registra una importante seconda ondata di contagi: i numeri tornano a crescere e gli ospedali a riempirsi. Infatti, il 19 ottobre arriva un nuovo DPCM con misure di sicurezza più stringenti. A novembre l'Italia viene divisa in tre zone di rischio: fascia gialla, arancione e rossa con scenari normativi differenziati. Il 27 dicembre 2020 parte ufficialmente in Italia il "Vaccine-day" europeo, la campagna di vaccinazione contro il Coronavirus. A gennaio 2021 parte la campagna vaccinale. I punti fondamentali del piano prevedono: la vaccinazione gratuita e garantita a tutti; l'identificazione di categorie da vaccinare con priorità nella fascia iniziale a limitata disponibilità

dei vaccini. Si assiste anche a una crisi di governo: il 26 gennaio Giuseppe Conte sale al Quirinale per rassegnare le dimissioni. Nel mese di febbraio si continua ad osservare un'importante accelerazione dell'aumento dell'incidenza a livello nazionale e scende l'età media degli individui contagiati. Il 13 febbraio entra in carica il nuovo governo Draghi: il Neopresidente del Consiglio Mario Draghi nel suo secondo intervento pubblico ha ribadito la necessità di un potenziamento della campagna di vaccinazione. Continua la suddivisione dell'Italia, con l'aggiunta della zona bianca. Si registra in questo periodo la terza ondata del virus, con i ricoveri in terapia intensiva in aumento, anche se in minor numero rispetto all'ondata precedente. La speranza per un alleviamento dei contagi è riposta, oltre che nella spinta climatica primaverile come è avvenuto nell'anno passato, anche in un aumento dell'efficacia del vaccino, che aumenterebbe la platea degli immuni. Il nuovo DPCM del 17 maggio 2021 prevede grandi riaperture e un allentamento graduale delle restrizioni anti-contagio. Intanto la campagna vaccinale prosegue in tutte le Province Italiane.

2.2 STUDIARE L'EPIDEMIA MEDIANTE MODELLI MATEMATICI

La diffusione di un'epidemia può essere studiata mediante opportuni modelli matematici. Mai come nel corso di questa pandemia c'è stato un boom dei modelli matematici che hanno elaborato previsioni sulla possibile diffusione del virus, differenziando le stime sul numero di posti occupati nelle terapie intensive e di

pazienti deceduti sulla base della maggiore o minore rigidità delle misure di distanziamento sociale. Epidemiologia è un termine che deriva dal greco “*epi*” (=attorno), “*démos*” (=popolazione) e “*logos*” (=discussione, studio) ed è, quindi, lo studio dei fattori che determinano la frequenza e la distribuzione delle malattie e dei problemi sanitari in una popolazione. Tale disciplina ha principalmente per oggetto lo studio delle caratteristiche di manifestazione di un evento in una determinata popolazione e l’individuazione dei fattori che ne sono all’origine, delle loro modalità di intervento e delle condizioni che ne favoriscono od ostacolano l’azione.

L’incertezza e la problematicità fondamentale per i modelli di previsione del coronavirus derivano dall’affidabilità dei dati disponibili, che costituiscono l’*“input”* del modello. Un esempio è quello relativo ai casi totali registrati per mezzo di tamponi in una giornata, il cui numero è influenzato dalla decisione di eseguire tamponi solo ai sintomatici, per cui il numero di persone realmente positive al virus non è possibile registrarlo. Quindi il conteggio non fornisce una vera immagine delle persone positive alla malattia ma dipende dall’efficacia del metodo utilizzato per monitorarla. È necessario dunque esaminare con particolare attenzione la fonte dei dati e le metodologie di analisi su cui si basano le attuali previsioni sull’evoluzione futura dell’epidemia; altrettanto importante è prestare grande prudenza nell’analisi della qualità dei dati, nella loro elaborazione e nella comunicazione dei risultati.

2.3 UNO SGUARDO ALLA STORIA DEI MODELLI DEL PASSATO

Da più di cento anni, modelli matematici sono utilizzati per descrivere la diffusione di epidemie. I modelli attualmente utilizzati derivano in gran parte da quello proposto da Kermack e Mckendrick nel 1927. Il modello prende il nome di SIR, acronimo dei 3 compartimenti in cui viene divisa la popolazione: S che sta per “*Susceptible*”, o “Suscettibili”, ossia gli individui suscettibili di infezione, cioè tutti quei soggetti che sono sani ma possono ammalarsi e diventare infetti; I sta per “*Infectious*” o “Infettivi”, indica gli individui che hanno contratto la malattia e che possono trasmetterla; infine, R sta per “*Recovered*” o “Rimossi”, rappresenta il numero di individui che sono diventati immuni dopo aver superato la malattia o che sono deceduti a causa di essa.

Per arginare alla semplicità del modello appena presentato, numerose sono le integrazioni utilizzate in letteratura. Ad esempio, il modello SIRS considera anche l’immunità alla malattia decrescente nel tempo. Il modello SIR presuppone che le persone mantengano l’immunità per tutta la vita a una malattia al momento del recupero; questo è il caso di una varietà di malattie. Per un’altra serie di malattie trasportate nell’aria, come per l’influenza stagionale, l’immunità di un individuo può diminuire nel tempo. C’è il modello SIRD che considera anche il compartimento dei morti (D). Altro modello più realistico è il SEIR che prende in considerazione anche il caso in cui un individuo infetto non diventa subito infettivo: il patogeno ha bisogno di tempo per replicarsi e stabilirsi nel nuovo

ospite. Questo tempo, in cui l'individuo è infetto, ma non infettivo, prende il nome di "periodo latente". Altro concetto è il "periodo di incubazione" che corrisponde all'intervallo di tempo tra l'infezione e il manifestarsi dei sintomi; non è detto che i due coincidano. Per introdurre il periodo latente nel modello SIR, si considera una nuova classe indicata con la lettera E , quella degli "Exposed", o "Esposti", che generalmente segue quella dei suscettibili.

2.4 MODELLI EPIDEMIOLOGICI PER IL COVID-19

I modelli compartimentali classici devono essere modificati e studiati per descrivere correttamente gli effetti delle azioni esterne forzate per ridurre l'impatto della malattia. Spesso anche i dati sono incompleti ed eterogenei, per cui si deve tener conto anche dell'alto grado di incertezza.

2.4.1 Controllo dell'epidemia con il modello SIDARTHE

SIDARTHE è un modello proposto da G. Giordano, F. Blanchini, R. Bruno, P. Colanieri, A. Di Filippo, A. Di Matteo, M. Colanieri, in "*Modelling the COVID-19 epidemic and implementation of population-wide interventions in Italy*", per studiare il caso italiano. Si tratta di un modello di tipo compartimentale che estende i modelli epidemiologici SIR/SEIR. Come è possibile osservare nella Figura 1, la popolazione è suddivisa in 8 compartimenti in relazione al contatto con il virus. In questo modo si tenta di considerare le numerose peculiarità della malattia.

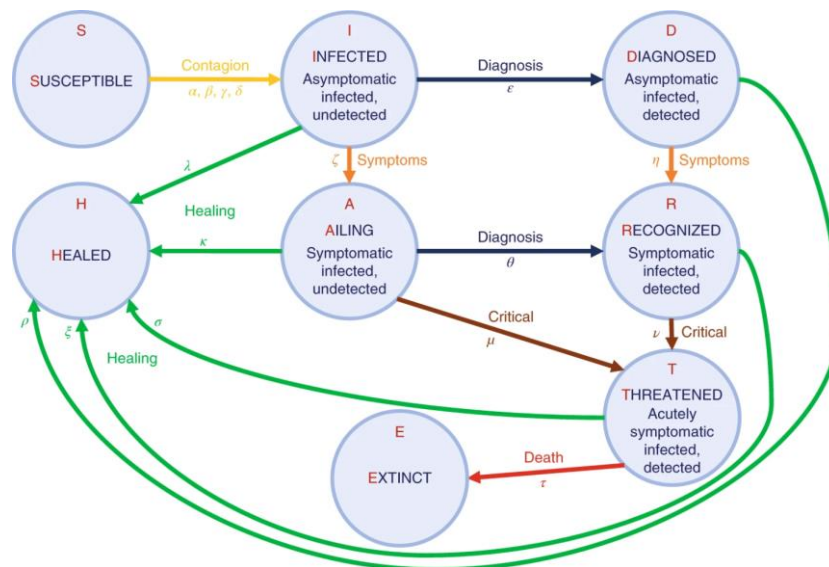


Figura 1. Schema grafico che rappresenta le interazioni tra le diverse fasi dell'infezione nel modello, indicati con le lettere maiuscole, e tutti i parametri considerati, denotati da lettere greche. Fonte: www.nature.com

L'obiettivo principale di questo lavoro è la modellazione e ricostruzione del "sommerso", cioè l'identificazione del numero di contagiati non diagnosticati, a partire dai dati messi a disposizione della Protezione Civile. Le conclusioni sulla letalità e sull'evoluzione del contagio mettono in evidenza l'importanza dei test a tappeto per il tracciamento del contagio e delle manovre di *lockdown* e distanziamento sociale. Confrontando diversi scenari, i risultati del modello prevedono che un'implementazione parziale delle misure di blocco comporta un ritardo nel picco di individui infetti e pazienti ricoverati in terapia intensiva, in contrasto con un'unica diminuzione moderata del numero totale di individui infetti e ammissioni in terapia intensiva. Al contrario, l'attuazione di strategie di

distanziamento sociale molto forti si tradurrebbe in un picco inferiore previsto di individui infetti e pazienti ricoverati in terapia intensiva, con una mancata diminuzione del numero totale di individui infetti e ricoverati in terapia intensiva a causa della malattia. Infine, una massiva campagna di test su tutta la popolazione e un rafforzamento del tracciamento dei contatti porterebbero a raggiungere il picco epidemico prima.

2.4.2 Modello che include le NPI

“Modelling and predicting the effect of social distancing and travel restrictions on COVID-19 spreading”, di F. Parino, L. Zino, M. Porfiri, A. Rizzo , è un modello elaborato su base provinciale e include le caratteristiche salienti del comportamento umano, insieme a una rappresentazione realistica della demografia della popolazione e degli spostamenti, sia dei flussi pendolari, sia di quelli più a lunga distanza e durata. La pandemia ha costretto la maggior parte dei paesi ad attuare interventi non farmaceutici senza precedenti (NPIs, dall’ inglese *“non-pharmaceutical interventions”*). Gli interventi più comuni sono stati di due tipi: la riduzione dell’attività degli individui e l’applicazione delle restrizioni alla mobilità. Il notevole costo economico e sociale derivante da tali scelte rende fondamentale valutarne l’efficacia. Il modello incapsula il periodo di latenza tra contagio e infettività e il ritardo nelle segnalazioni ufficiale di infezioni e decessi. Lo schema è illustrato nella Figura 2: dopo il contagio, gli individui infetti

diventano esposti (E) prima di spostarsi tra gli infettivi (I); terminato il periodo infettivo, gli individui diventano non infettivi (N), prima di recuperare o morire.

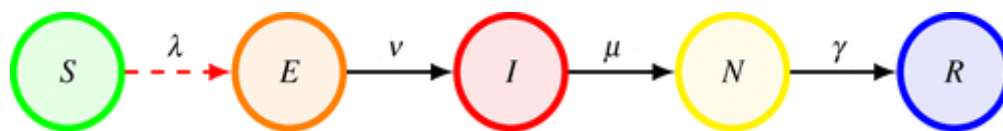


Figura 2. Schema della progressione epidemica. Fonte: royalsocietypublishing.org

Nello studio viene esaminata l'interazione tra riduzione dell'attività sociale, che agisce sulla rete di contatti, e restrizioni alla mobilità, che modificano i modelli di mobilità tra le comunità. I risultati rivelano che essenziale è la tempistica per la loro effettiva attuazione. Sebbene la riduzione delle attività sia una strategia decisiva, l'impatto della riduzione della mobilità ha un effetto più sfumato. In primo luogo, esse producono benefici solo se applicate nella fase iniziale dell'epidemia e solo se abbinate a adeguate politiche di riduzione delle attività. In secondo luogo, l'effetto di tali restrizioni è influenzato fortemente dallo spazio.

Un secondo ambito di analisi è stato sulla possibile implementazione di NPI mirate, con restrizioni più forti solo per individui di fasce di età più a rischio, attuando al contempo lievi restrizioni per i giovani. Sebbene tali misure sarebbero auspicabili in termini di ragioni economiche, esse avrebbero valore limitato in termini di salute pubblica.

2.4.3 Un modello previsionale NETWORK-BASED

C. Pizzuti, A. Socievole, B. Prasse, P. V. Mieghem, nell'elaborato "*Network-based Prediction of COVID-19 Epidemic Spreading in Italy*", propongono un modello NIPA-LD (*Network Inference-based Prediction Algorithm with LockDown*) in cui l'approccio utilizzato è *network-based*: le popolazioni regionali italiane sono rappresentate da nodi di rete e i collegamenti di rete rappresentano le interazioni tra le regioni. Il classico modello SIR, che prevede che i tassi di infezione rimangano costanti nel tempo, viene modificato, introducendo l'applicazione dei "*transmission modifiers*" per adeguare i tassi di trasmissione del virus in base alle misure di chiusura imposte dal Governo. I risultati mostrano che la percentuale di errore nelle previsioni, in media su tutte le regioni italiane, è nettamente minore nel modello NIPA-LD rispetto al modello NIPA classico, in cui vengono omesse le misure di contenimento. Ciò significa che tale modello può essere utilizzato per eseguire previsioni affidabili e permetterebbe di valutare quali strategie di blocco sono più appropriate per la gestione dell'epidemia, sia a breve che a lungo termine, anche se generalmente queste ultime sono difficili.

2.4.4 Modello SIR modificato dal DET

Il modello contenuto in "*A Modified SIR Model for the COVID-19 Contagion in Italy*" di G. C. Calafiore, C. Novara e C. Possieri, tenta di superare due limiti del modello standard, introducendo nuovi parametri. In primo luogo, i parametri non

sono considerati costanti, in quanto in un ambiente epidemico reale possono verificarsi dei cambiamenti di parametro, ad esempio dovuti a misure di contenimento adottate. Il secondo problema riguarda il numero degli individui infetti, nella realtà vengono considerati solo quelli che sono risultati “positivi” al test del tampone. I risultati del modello riguardano l’efficacia delle misure di controllo imposte dalle autorità nel diminuire la velocità di trasmissione della malattia, per lo più nel breve periodo; il tasso di recupero tende ad aumentare con il tempo, e il tasso di mortalità tende a diminuire: tali andamenti sembrano però legati a un miglioramento dell’efficacia della risposta del sistema sanitario nazionale, e una possibile modifica nella natura, virulenza e letalità del virus.

2.4.5 Due modelli che includono l’eterogeneità regionale

“A network model of Italy shows that intermittent regional strategies can alleviate the COVID-19 epidemic”, a opera di F. Della Rossa, D. Salzano, A. Di Meglio, F. De Lellis, M. Coraggio, C. Calabrese, A. Guarino, R. Cardona-Rivera, P. De Lellis, D. Liuzza, F. Lo Iudice, G. Russo, M. Di Bernardo, è un modello che modello considera l’Italia come una rete di regioni, ognuna modellata con parametri diversi. Infatti, ogni regione è indipendente sulla Salute e può prendere decisioni riguardo la propria quota del Servizio Sanitario Nazionale, rafforzando o indebolendo le norme nazionali di contenimento mediante l’azione dei presidenti regionali. L’obiettivo è capire se e quando le misure prese dal Governo hanno

avuto un effetto sia a livello nazionale a livello regionale. Il primo risultato consiste nel successo delle misure nazionali di *lockdown*, prima in due regioni (Veneto e Lombardia), poi a livello nazionale. Si riscontra un certo grado di omogeneità delle varie regioni in risposta alle rigide norme di distanziamento sociale. Altro risultato dell'analisi riguarda il tasso di mortalità che varia in funzione del livello di occupazione dei letti ospedalieri in ogni regione. Inoltre, l'efficacia di misure differenziate è confermata dal modello quanto lo sono le misure nazionali.

“Modelling and predicting the spatio-temporal spread of COVID-19 in Italy, BMC Infectious Diseases”, di D. Giuliani, M. M. Dickson, G. Espa, F. Santi, è un altro modello che considera l'eterogeneità regionale italiana, evidenziando la disuguaglianza tra le regioni meridionali e quelle settentrionali, che rappresentano il centro industriale, economico e finanziario. Il territorio del Nord è caratterizzato da un grande flusso di persone provenienti da altre aree a causa del diffuso pendolarismo del lavoro e di persone in cerca di un sistema sanitario più efficiente. Tali elementi si ripercuotono sulla diffusione della malattia. A partire da tali regioni sono state applicate misure forti, poi estese all'intera nazione. È fondamentale includere una componente che spieghi la dipendenza spaziale tra le varie aree di diffusione del virus. Le misure di *lockdown* prese hanno seguito un'articolazione crescente e divisa in tre fasi: all'inizio solo alcuni comuni di Lombardia e Veneto sono stati obbligati alla quarantena, poi l'intera regione e

altre province settentrionali, infine tutti i territori italiani. Infatti, nelle province centro-meridionali la malattia si è diffusa con ritardo, ed è stata anche influenzata dall'arrivo di persone fuggite dalle regioni del nord poi sottoposte al blocco.

2.4.6 Un tentativo per arginare l'incertezza dei dati

De Natale, V. Ricciardi, G. De Luca, D. De Natale, G. Di Meglio, A. Ferragamo, V. Marchitelli, A. Piccolo, A. Scala, R. Somma, E. Spina, C. Troise, in “*The COVID-19 Infection in Italy: A Statistical Study of an Abnormally Severe Disease*” hanno focalizzato lo studio sulle peculiarità che distinguono l'evoluzione dell'epidemia in Italia rispetto ad altri paesi, che appare particolarmente aggressiva sia in termini di diffusione che di letalità. Il QRC (rapporto tra numero di deceduti e numero di infetti registrati) ha raggiunto circa l'11% nel paese complessivo. Date le procedure di laboratorio complesse per diagnosticare il virus, stimare un valore affidabile e realistico dell'IFR (numero di decessi su numero di infezioni totale) è molto complicato. Sono stati seguiti tre diversi approcci, stimando un numero reale di infezioni che oscilla tra circa 0,3 e 8 milioni di persone al 30 marzo 2020, che, sebbene sembri un numero elevato, non è irragionevole per diversi motivi. Attraverso il ricorso a studi indipendenti di utilizzo dei mezzi pubblici, è stato stimato il numero delle possibili infezioni mediante i contatti tra persona e persona: circa 30 milioni di persone viaggiano in Italia ogni giorno per lavoro o studio. Includendo questo dato si ottiene una stima

minima di circa 11 milioni di persone a stretto contatto tra loro ogni giorno. Altri fattori che potrebbero giustificare la notevole letalità in Italia sono l'elevato inquinamento atmosferico dell'area circostante la Pianura Padana, l'età media elevata, il numero di fumatori, l'elevata resistenza agli antibiotici; tali elementi però non sembrano svolgere un ruolo fondamentale. Da considerare anche l'inadeguatezza del sistema sanitario. Il modello si pone anche l'obiettivo di stimare un possibile rallentamento del tasso di infezione e di verificare le previsioni in relazione alle misure di contenimento adottate dalle autorità. Anche questo modello dimostra il successo delle misure di blocco nazionale, in particolare rileva l'elevata efficacia della chiusura delle scuole.

2.4.7 L'efficacia di una strategia graduale

“Control with uncertain data of socially structured compartmental epidemic models”, elaborato da G. Albi, L. Pareschi, M. Zannella, è un modello basato sull'ipotesi che applicare misure di contenimento molto forti comporta costi sulla popolazione ed economici significativi, di conseguenza il Governo potrebbero non essere in grado di mantenerle per più di un breve periodo di tempo. Dover pianificare interventi a lungo termine comporta la necessità di analizzare strategie di controllo sostenibili, basate sulla struttura sociale del sistema, che possono essere efficaci quanto la strategia ottimale più costosa. L'importanza dei tempi e dell'intensità degli interventi sono elementi cruciali da considerare nel caso di

parametri incerti sul numero effettivo di persone infette. I risultati dimostrano che una strategia graduale è efficace, consente di contenere il picco dei contagi e avere un risultato simile al caso in cui si applicano misure di controllo complete, ma ad un costo socioeconomico di gran lunga minore.

2.4.8 Un modello per l'importanza dell'isolamento

E. Bertuzzo, L. Mari, D. Pasetto, S. Miccoli, R. Casagrandi, M. Gatto & A. Rinaldo, in “*The geography of COVID-19 spread in Italy and implications for the relaxation of confinement measures*”, studiando il possibile effetto delle misure di allentamento graduale, hanno creato un modello che ha riscontrato nell'isolamento dei contatti la più grande mossa da attuare. Il modello considerato è il SEIPA, che ingloba l'interazione tra le comunità, le restrizioni alla mobilità, i tempi di infezione e il distanziamento sociale. Ne deriva che le infezioni si possono verificare all'interno delle comunità e possono essere importate o esportate tramite i collegamenti. In un'ottica di allentamento delle misure adottate, bisogna considerare anche il grado di connettività dalle comunità dopo la parziale riapertura. L'isolamento dei casi come possibile risposta a un nuovo aumento del tasso di trasmissione in seguito all'allentamento delle misure restrittive è una delle strategie implementabili, per evitare nuove applicazioni di misure di blocco forzate. Tali misure di isolamento sarebbero efficaci se applicate ai compartimenti degli esposti e dei soggetti al picco di infettività. Per ogni regione italiana, il

modello calcola il numero giornaliero di tali categorie che dovrebbe essere isolato per controbilanciare l'aumento del tasso di trasmissione derivante dalla riapertura, mantenendo così l'andamento decrescente della curva epidemica che si è raggiunto con il lockdown. Se l'emergenza viene adeguatamente controllata e il numero di infetti continua a diminuire, tale sforzo di isolamento diminuisce proporzionalmente nel tempo, ma in termini percentuali rimane costante. La risposta delle varie regioni rimane differenziata. Tracciare e isolare tutte le infezioni primarie è comunque un compito impegnativo, per cui il modello suggerisce che ogni regione dovrebbe valutare le proprie strategie di tracciamento. In caso di fallimento di tali misure, si dovrebbe ritardare ulteriormente l'allentamento; per tenere continuamente sotto controllo la situazione vengono proposte misure *stop-and-go*.

CAPITOLO 3. ANALISI EMPIRICA

3.1 I DATI E LE VARIABILI

In Italia, il monitoraggio dell'epidemia viene effettuato attraverso due flussi di dati: il primo è inviato dalle Regioni e coordinato dal Ministero della Salute, con il supporto della Protezione Civile e l'Istituto Superiore di Sanità, per raccogliere informazioni sul numero di test positivi, decessi, ricoveri in ospedale e terapia intensiva in ogni Provincia d'Italia; il secondo è il flusso dei dati individuali inviati dalle Regioni all'Istituto Superiore di Sanità, che comprende anche i dati demografici, la comorbidità, lo stato clinico, per un'analisi più accurata. Ogni giorno tra le 17:00 e le 18:00 la Protezione Civile aggiorna il database che mostra tutti i dati pubblici dell'epidemia.

Si tratta di dati in formato panel, ossia dati che combinano informazioni relative alle caratteristiche di un numero N di province/regioni nello stesso istante temporale con quelle rilevate per le stesse in T diversi giorni. Nel caso in esame, gli individui sono rappresentati dalla popolazione italiana, suddivisa in 21 Regioni, considerando separatamente le Province Autonome di Trento e Bolzano. L'arco di tempo preso in considerazione per le stime presenti nell'elaborato comprende i giorni tra il 24 febbraio 2020 e l'11 giugno 2021.

Nel modello viene considerato come variabile dipendente il tasso di crescita delle terapie intensive. Il dato "*Intensive Care*" (IC) identifica la quantità di posti letto

occupati in rianimazione. IC non indica i nuovi ingressi nelle precedenti 24 ore, allo stesso tempo non specifica quanti siano morti o siano stati spostati in altri reparti. Tale numero assoluto è considerato però l'aggiornamento più preciso e stabile per capire cosa sta succedendo e per monitorare l'epidemia, a differenza di dati come il numero giornaliero di nuovi contagi, condizionati da altri dati "sporchi", così come vengono definiti dagli esperti, ossia la variazione giornaliera del numero di tamponi eseguiti, i diversi criteri di accesso ai test stabiliti dalle regioni, i ritardi di comunicazione degli esiti, il sistema di contact tracing.

Le variabili indipendenti (o regressori) scelte per il modello sono il numero del totale dei positivi (tpos), dei nuovi attualmente positivi (npos), dei tamponi effettuati (tamp), dei soggetti in isolamento domiciliare (home), del totale degli ospedalizzati (hosp), dei dimessi guariti (guar) e dei deceduti (dec).

Utilizzare le variabili in logaritmo, per analizzare l'elasticità tra le grandezze, non è una strategia perseguibile perché le serie non sono stazionarie, di conseguenza si giungerebbe a un'inferenza fuorviante, non consistente.

Utilizzo quindi i tassi di crescita, dove le variabili sono definite come segue:

$$y_t = (\ln(x_t) - \ln(x_{t-1})) \cdot 100$$

dove y_t indica il tasso di crescita della variabile indipendente al tempo t ; x_t e x_{t-1} indicano il valore che la variabile dipendente assume rispettivamente al tempo t e al tempo $(t-1)$; tra parentesi è contenuta la differenza prima tra i logaritmi e infine moltiplico per 100 per ottenere risultati in termini percentuali.

Inoltre, sono state aggiunte cinque variabili di tipo dummy con valore pari a uno in corrispondenza dei principali DPCM emanati dal Governo nel periodo in esame. In particolare, i DPCM in corrispondenza di decisioni riguardanti le riaperture maggiori del 2020 e del 2021 (rispettivamente DPCM_mar10 e DPCM_mag17) e gli altri principali del 2020 (DPCM_apr26, DPCM_giu11, DPCM_ott24).

3.2 MODELLO PER LE SERIE STORICHE POOLED

La diponibilità di dati in formato panel ha reso possibile l'applicazione di tutti gli approcci utilizzabili. Il primo preso in analisi è il modello per le serie storiche pooled, in quanto ho una serie temporale nettamente superiore a quella cross-section; probabilmente sarà anche il più appropriato.

L'equazione del modello è:

$$\begin{aligned} y_{it} = & \beta_0 + \beta_1 tpos + \beta_2 npos + \beta_3 tamp + \beta_4 home + \beta_5 hosp + \beta_6 guar \\ & + \beta_7 dec + \beta_8 DPCM_mar10 + \beta_9 DPCM_apr26 + \beta_{10} DPCM_giu11 \\ & + \beta_{11} DPCM_ott24 + \beta_{12} DPCM_mag17 + \varepsilon_{it} \end{aligned}$$

La seguente è la stima elaborata con i dati epidemiologici:

Tabella 1: POOLED ESTIMATION.

Modello 1: Pooled OLS, usando 8184 osservazioni
 Include 21 unità cross section
 Lunghezza serie storiche: minimo 270, massimo 470
 Variabile dipendente: y

	coefficiente	errore std.	rapporto t	p-value	
const	-0,0882397	0,175833	-0,5018	0,6158	
tpos	0,0186372	0,0527680	0,3532	0,7240	
npos	-0,000982888	0,00255898	-0,3841	0,7009	
tamp	-0,0672302	0,0635202	-1,058	0,2899	
home	0,0648908	0,0350822	1,850	0,0644	*
hosp	0,620436	0,0233627	26,56	3,99e-149	***
guar	0,0180762	0,0199858	0,9044	0,3658	
dec	0,0449415	0,0295971	1,518	0,1289	
DPCM_mar10	47,4228	4,99700	9,490	2,96e-021	***
DPCM_apr26	-2,78277	3,10781	-0,8954	0,3706	
DPCM_giu11	-3,25417	5,24855	-0,6200	0,5353	
DPCM_ott24	4,81355	3,03933	1,584	0,1133	
DPCM_mag17	-9,68556	3,04265	-3,183	0,0015	***
Media var. dipendente	0,266696	SQM var. dipendente	14,90394		
Somma quadr. residui	1572627	E.S. della regressione	13,87315		
R-quadro	0,134811	R-quadro corretto	0,133541		
F(12, 8171)	106,0985	P-value(F)	5,0e-246		
Log-verosimiglianza	-33129,64	Criterio di Akaike	66285,29		
Criterio di Schwarz	66376,42	Hannan-Quinn	66316,45		
rho	-0,134331	Durbin-Watson	2,198343		

Il primo elemento che salta all'occhio è la significatività della stima dei vari coefficienti, per cui si dovrebbero predisporre delle revisioni, eccezion fatta per la variabile degli ospedalizzati e dei DPCM delle riaperture.

La bontà di adattamento (*R-quadro*) è molto bassa. Tale effetto potrebbe essere dovuto al fatto che il modello necessita di altri dati di cui non disponiamo, ragion per cui sono state inserite anche variabili dummy.

Un apposito test è stato svolto per verificare che i residui non siano costituiti da una serie di realizzazioni incorrelate. L'ipotesi nulla prevede l'assenza di autocorrelazione nei residui del modello.

Tipo di test	Distribuzione	Statistica test	p-value	Decisione
Test di Wooldridge	t_{20}	-3,20931	0,0044	RIFIUTO

È confermata la presenza di autocorrelazione di primo ordine. Inoltre, il test Durbin-Watson, con statistica test pari a 2,198343 ($d_L=1,9144$ e $d_U=1,9385$), rifiuta l'ipotesi nulla di assenza di autocorrelazione di primo ordine; c'è ulteriore prova di autocorrelazione della serie dei residui con il passato più vicino.

Sono stati eseguiti i test di eteroschedasticità nel modello.

Tipo di test	Distribuzione	Statistica test	p-value	Decisione
Test di White	χ^2_{74}	1347,428750	0,0000	RIFIUTO
Test di Breusch-Pagan	χ^2_{12}	2347,274175	0,0000	RIFIUTO

Rifiutando l'ipotesi nulla di assenza di omoschedasticità, essi segnalano la possibilità di un contesto eteroschedastico.

Valuto inoltre la forma funzionale tramite il test RESET di Ramsey, che sotto l'ipotesi nulla prevede che il modello ha problemi di specificazione.

Tipo di test	Distribuzione	Statistica test	p-value	Decisione
Test RESET	F _{1,8170}	29,389219	0,0000	RIFIUTO

L'ipotesi nulla di adeguata specificazione è rigettata.

Vado ora ad analizzare i coefficienti ottenuti, valutando gli impatti marginali tra le grandezze considerate. Innanzitutto, emerge la relazione negativa delle terapie intensive con il numero dei tamponi giornalieri con esito positivo e il numero dei tamponi totali, con valori rispettivamente di circa 0,00098 e 0,06723. Relazione positiva si ha invece con il totale dei positivi (circa 0,01864), con i casi in isolamento domiciliare (circa 0,06489), con i ricoverati in ospedale (circa 0,6204), con i guariti (circa 0,01808) e con i deceduti (circa 0,04494).

Per quanto concerne i provvedimenti del Governo, dal modello emerge una bassa significatività per i DPCM. In particolare, si nota come abbiano avuto un qualche impatto sui dati i provvedimenti atti ad emanare nuove misure di riapertura, ossia quello del 10 marzo 2020 e l'ultimo risalente al 17 maggio 2021; entrambi hanno contribuito con un impatto marginale negativo rispetto alle terapie intensive. I restanti DPCM considerati, che probabilmente integravano in qualche modo decisioni passate, non hanno avuto alcun impatto significativo.

Un test è stato eseguito per valutare la presenza di un break strutturale durante la seconda ondata del virus. In particolare, ipotizzando il 1° ottobre 2020 come data

di riferimento per distinguere tra la prima e la seconda ondata, il test di Chow prevede sotto l'ipotesi nulla l'assenza di break strutturali.

Tipo di test	Distribuzione	Statistica test	p-value	Decisione
Test di Chow	F _{8,8163}	1,74823	0,0823	NON RIFIUTO

L'ipotesi nulla del test non è rifiutata, quindi non si riscontra in termini di tasso di crescita percentuale la presenza di un break strutturale.

3.3 MODELLI PER I DATI LONGITUDINALI

Quando si lavora con dati longitudinali si ha a che fare con matrici di dati caratterizzate generalmente da una elevata numerosità di individui e da una relativamente contenuta dimensione temporale. È possibile quindi stimare due differenti modelli, il modello ad effetti fissi e il modello ad effetti casuali, per poi individuare quello che più si adatta al panel e all'informazione disponibile.

3.3.1 Il modello ad effetti fissi

Il primo è il modello ad effetti fissi, in cui l'effetto fisso corrisponde alla costante dell'*i*-esimo individuo; tale costante cambia per ogni individuo.

Considerando l'*i*-esimo individuo, la configurazione del modello è la seguente:

$$y_i = \alpha_i + x_i'\beta + \varepsilon_i$$

dove β è il vettore contenente i *k* parametri da stimare e α_i è la costante che misura l'effetto individuale.

Applicando tale stima al modello epidemico e alle variabili individuate precedentemente si ottiene la seguente stima:

Tabella 2: FIXED EFFECTS MODEL.

Modello 2: Effetti fissi, usando 8184 osservazioni
 Includere 21 unità cross section
 Lunghezza serie storiche: minimo 270, massimo 470
 Variabile dipendente: y

	coefficiente	errore std.	rapporto t	p-value	
const	-0,0807325	0,176080	-0,4585	0,6466	
tpos	0,0201459	0,0528819	0,3810	0,7032	
npos	-0,00100030	0,00256160	-0,3905	0,6962	
tamp	-0,0735787	0,0637507	-1,154	0,2485	
home	0,0648200	0,0351344	1,845	0,0651	*
hosp	0,621133	0,0233986	26,55	5,41e-149	***
guar	0,0189530	0,0200311	0,9462	0,3441	
dec	0,0446703	0,0296647	1,506	0,1321	
DPCM_mar10	47,4112	5,00476	9,473	3,49e-021	***
DPCM_apr26	-2,77846	3,11104	-0,8931	0,3718	
DPCM_giu11	-3,37801	5,25711	-0,6426	0,5205	
DPCM_ott24	4,80503	3,04240	1,579	0,1143	
DPCM_mag17	-9,67988	3,04568	-3,178	0,0015	***
Media var. dipendente	0,266696	SQM var. dipendente	14,90394		
Somma quadr. residui	1571801	E.S. della regressione	13,88652		
R-quadro LSDV	0,135266	R-quadro intra-gruppi	0,134616		
LSDV F(32, 8151)	39,84427	P-value(F)	6,2e-229		
Log-verosimiglianza	-33127,50	Criterio di Akaike	66320,99		
Criterio di Schwarz	66552,32	Hannan-Quinn	66400,08		
rho	-0,134899	Durbin-Watson	2,199517		

Attraverso un opportuno test è possibile valutare l'effettiva presenza di effetti regionali nel campione. L'ipotesi nulla del test prevede l'uguaglianza degli effetti regionali per tutte le regioni/province, per cui se rifiutata ci sarebbe conferma della diversità del tasso di crescita delle varie variabili del virus per ogni località esaminata.

Tipo di test	Distribuzione	Statistica test	p-value	Decisione
Test per la differenza delle intercette di gruppo	F _{20,8351}	0,214169	0,9999	NON RIFIUTO

L'ipotesi nulla non è rifiutata, quindi non sono individuabili effetti differenti per le regioni in termini di tassi di crescita.

3.3.2 Il modello ad effetti casuali

Nel modello ad effetti casuali tutta l'eterogeneità si scarica nel termine di errore. L'errore contiene quindi una componente di eterogeneità propria di ciascun individuo e una componente omogenea per tutta la popolazione e per tempo. Considerando l'*i*-esimo individuo, la formula analitica di questo modello è:

$$y_i = \alpha + x_i'\beta + \mu_i + \varepsilon_i$$

dove α è costante per ogni individuo e la componente dell'errore μ_i varia per ogni soggetto.

La stima ottenuta applicando il precedente modello ai dati epidemici è quella che segue:

Tabella 3: RANDOM EFFECTS MODEL.

Modello 3: Effetti casuali (GLS), usando 8184 osservazioni
 Include 21 unità cross section
 Lunghezza serie storiche: minimo 270, massimo 470
 Variabile dipendente: y

	coefficiente	errore std.	z	p-value	
const	-0,0882397	0,175833	-0,5018	0,6158	
tpos	0,0186372	0,0527680	0,3532	0,7239	
npos	-0,000982888	0,00255898	-0,3841	0,7009	
tamp	-0,0672302	0,0635202	-1,058	0,2899	
home	0,0648908	0,0350822	1,850	0,0644	*
hosp	0,620436	0,0233627	26,56	2,15e-155	***
guar	0,0180762	0,0199858	0,9044	0,3658	
dec	0,0449415	0,0295971	1,518	0,1289	
DPCM_mar10	47,4228	4,99700	9,490	2,30e-021	***
DPCM_apr26	-2,78277	3,10781	-0,8954	0,3706	
DPCM_giu11	-3,25417	5,24855	-0,6200	0,5352	
DPCM_ott24	4,81355	3,03933	1,584	0,1132	
DPCM_mag17	-9,68556	3,04265	-3,183	0,0015	***
Media var. dipendente	0,266696	SQM var. dipendente	14,90394		
Somma quadr. residui	1572627	E.S. della regressione	13,87231		
Log-verosimiglianza	-33129,64	Criterio di Akaike	66285,29		
Criterio di Schwarz	66376,42	Hannan-Quinn	66316,45		
rho	-0,134899	Durbin-Watson	2,199517		

Dalla tabella è possibile notare come le stime dei parametri siano identiche per il modello pooled OLS e per il modello ad effetti casuali. Ciò è insolito, ma si potrebbe spiegare con i dati relativi alla varianza forniti dal software:

Varianza 'between' = 0
 Varianza 'within' = 192,835
 theta medio = 0

In particolare, la varianza “*between*” è nulla, indicando una situazione in cui sembra non esserci varianza tra le medie dei gruppi. La conseguenza di questo evento si riversa nel Theta usato per la trasformazione del modello ad effetti casuali che assume valore nullo, annullando gli effetti della trasformazione stessa.

3.3.3 Inferenza statistica

Eseguo quindi il test di Hausman che prevede sotto l’ipotesi nulla che lo stimatore ad effetti casuali sia consistente ed efficiente, mentre quello ad effetti fissi consistente ma non efficiente. Sotto l’ipotesi alternativa invece il primo è inconsistente e il secondo consistente.

Tipo di test	Distribuzione	Statistica test	p-value	Decisione
Test di Hausman	χ^2_{11}	3,26825	0,986685	NON RIFIUTO

L’ipotesi nulla del test quindi non è rifiutata, quindi che lo stimatore del modello ad effetti fissi è preferibile allo stimatore nel modello ad effetti casuali, l’eventuale presenza di effetti regionali si scarica tutta sulla costante.

3.4 GLI OSPEDALIZZATI

In tutte le stime effettuate, i coefficienti associati alle variabili sembrano mantenere all’incirca lo stesso valore in tutti i modelli.

Analizzando invece i *p-value*, esso è minimo per la variabile degli ospedalizzati in tutti i modelli. Ciò sottolinea la maggiore significatività del tasso di crescita di

questo dato in relazione alle terapie intensive. La dipendenza tra queste due grandezze è positiva e con impatto marginale circa dello 0,065. Infatti, come è prevedibile aspettarsi, se aumentano i casi di individui che hanno bisogno di assistenza in ospedale, sicuramente tra costoro ci saranno soggetti che arrivano in rianimazione a causa di insufficienza respiratoria acuta e che hanno bisogno di supporto respiratorio (ventilazione meccanica, cioè l'intubazione o ventilazione non invasiva, come i 'caschi' o maschera Venturi).

3.5 IL TASSO DI POSITIVITÀ

Un'altra interessante analisi è stata fatta riguardo il tasso di positività dei tamponi, ossia il rapporto tra i tamponi risultati positivi e il totale dei tamponi effettuati, per valutare se assume significatività all'interno dei modelli. Potrebbe risultare un valore più utile, ai fini dell'andamento dell'epidemia, rispetto al numero dei tamponi assoluto, perché questo deve essere sempre rapportato alla popolazione del territorio che si sta considerando; più utile anche rispetto al numero di contagi scoperti, perché una regione che scopre pochi casi perché fa pochi test non sta necessariamente meglio di un'altra che ne scopre di più perché fa più test, rilevando anche i casi asintomatici.

Tale indicatore è sempre più utilizzato come esemplare dell'andamento dell'epidemia e dell'estensione dei test in una certa regione.

Quest'ultimo è stato inserito nei modelli, a sostituzione delle due variabili che formano il rapporto. Il coefficiente relativo a tale variabile (share) risulta per ogni modello come in tabella:

	Modello 1: Pooled OLS	Modello 2: Effetti fissi	Modello 3: Effetti casuali (GLS)
coefficiente	-0,000975730	-0,000991955	-0,000975730
p-value	0,7030	0,6986	0,7030
rapporto t	-0,3813	-0,3872	-0,3813

Per valutare la significatività del tasso di positività nei modelli, osservo i risultati del test t che prevede come ipotesi nulla l'uguaglianza del coefficiente a zero. Se l'ipotesi nulla non è rifiutata la variabile non impatta nella regressione, invece se è rifiutata il coefficiente è significativo del punto di vista statistico, quindi impatta.

Com'è possibile osservare, effettuando il test *t*, l'ipotesi nulla non è rifiutata.

La variabile terapie intensive sembra non essere funzione diretta del loro rapporto tra le variabili esaminate.

3.6 REVISIONE DELL'INFORMAZIONE DISPONIBILE

Dall'analisi delle stime riportate si possono trarre alcune conclusioni preliminari, che possono portare a degli interventi correttivi effettuabili sull'informazione disponibile nell'ottica di miglioramento dei risultati. I modelli presentati sono, infatti, da considerarsi *misspecified*, e il basso *p-value* del test RESET ne fornisce

la conferma. Per risolvere tale problema è opportuno effettuare una selezione più accurata delle variabili da inserire.

Inoltre, la disponibilità di un'informazione completa riguardo alle osservazioni delle variabili considerate nel modello e la possibilità di integrarle con altre, dovrebbe dare luogo a stime probabilmente più significative.

In ogni caso, le stime effettuate sono tutte affette da problematiche di eteroschedasticità e non corretta specificazione della forma funzionale. Pertanto, è necessario intervenire e correggere tali questioni. Tale situazione potrebbe però mantenersi anche con un'informazione modificata.

CAPITOLO 4. CONCLUSIONI

La crisi pandemica ha sottolineato ancora una volta la necessità di incrementare l'attività di monitoraggio per prevedere l'evolversi di un fenomeno.

I modelli presentati hanno lo scopo di fornire un'analisi dettagliata della tendenza dell'epidemia di Covid-19 in Italia, resa possibile dall'ampia disponibilità di dati che sono croce e delizia per svolgere questo tipo di lavoro. Infatti, se da una parte la numerosità di dati permette di guadagnare efficienza, l'affidabilità di questi ultimi potrebbe essere a rischio, dato che provengono da numerose fonti e spesso si incorre in correzioni, rettifiche importanti e omissioni di registrazioni.

In definitiva, però, il modello fornisce una rappresentazione della nostra realtà, confermando molteplici trend. Inoltre, le IC sono in relazione negativa con il numero di tamponi effettuali ogni giorno e quelli con esito positivo, a confermare la tendenza positiva dell'evoluzione della malattia in quest'ultimo periodo di pandemia. Il numero di ricoverati gravi è in relazione positiva con il totale dei positivi, dei malati che si curano in casa, dei ricoverati e dei deceduti, infatti, un aumento del numero dei positivi al virus sicuramente si tradurrà in un aumento di malati gravi o comunque di soggetti che hanno bisogno di maggiore assistenza in un centro ospedaliero. Stupisce la relazione positiva, invece, con il numero dei guariti, ma potrebbe essere la conferma del fatto che il numero dei positivi sia aumentato tanto e allo stesso tempo anche la malattia colpisce con meno forza i

soggetti, e che quindi il numero dei guariti aumenti nonostante aumenti anche quello delle terapie intensive. Tale effetto potrebbe essere dovuto anche all'efficacia della progressiva campagna di vaccinazione e all'avanzamento della buona stagione, che mitigano l'aggressività del virus.

L'econometria intesa come scienza sociale che si propone di confrontare la teoria economica con i dati storici a disposizione, al fine di prevedere e/o suggerire l'esistenza di relazioni ipotetiche fra le variabili cui si riferiscono i dati, si conferma un potente strumento a disposizione di economisti e statisti, ma anche di politici, per poter assumere decisioni rilevanti in situazioni di estrema delicatezza e senza precedenti come questa del Coronavirus.

Bibliografia e Sitografia

G. ALBI, L. PARESCHI, M. ZANELLA (2020), *Control with uncertain data of socially structured compartmental epidemic models*, Cornell University, pp. 26

E. BERTUZZO, L. MARI, D. PASSETTO, S. MICCOLI, R. CASAGRANDE, M. GATTO & A. RINALDO (2020), *The geography of COVID-19 spread in Italy and implications for the relaxation of confinement measures*, Nature Communications, pp. 11

G. C. CALAFIORE, C. NOVARA, C. POSSIERI (2020), *A Modified SIR Model for the COVID-19 Contagion in Italy*, Cornell University, pp.6

G. DE NATALE, V. RICCIARDI, G. DE LUCA, D. DE NATALE, G. DI MEGLIO, A. FERRAGAMO, V. MARCHITELLI, A. PICCOLO, A. SCALA, R. SOMMA, E. SPINA, C. TROISE (2020), *The COVID-19 Infection in Italy: A Statistical Study of an Abnormally Severe Disease*, Journal of Clinical Medicine, pp. 19

F. DELLA ROSSA, D. SALZANO, A. DI MEGLIO, F. DE LELLIS, M. CORAGGIO, C. CALABRESE, A. GUARINO, R. CARDONA-RIVERA, P. DE LELLIS, D. LIUZZA, F. LO IUDICE, G. RUSSO, M. DI BERNARDO (2020), *A network model of Italy shows that intermittent regional strategies can alleviate the COVID-19 epidemic*, Nature Communications, pp. 9

A. DI MATTEO, M. COLANIERI (2020), *Modelling the COVID-19 epidemic and implementation of population-wide interventions in Italy*, Nature Medicine, pp.32

G. GIORDANO, F. BLANCHINI, R. BRUNO, P. COLANIERI, A. DI FILIPPO, D. GIULIANI, M. M. DICKSON, G. ESPA, F. SANTI (2020), *Modelling and predicting the spatio-temporal spread of COVID-19 in Italy*, BMC Infectious Diseases, pp. 10

G. PALOMBA (2008), *Panel Data*, pp. 22

F. PARINO, L. ZINO, M. PORFIRI, A. RIZZO (2021), *Modelling and predicting the effect of social distancing and travel restrictions on COVID-19 spreading*, The Royal Society, pp.10

C. PIZZUTI, A. SOCIEVOLE, B. PRASSE, P. V. MIEGHEM (2020), *Network-based Prediction of COVID-19 Epidemic Spreading in Italy*, PubMed Central, pp. 24

<https://www.salute.gov.it/portale/nuovocoronavirus/dettaglioContenutiNuovoCoronavirus.jsp?lingua=italiano&id=5351&area=nuovoCoronavirus&menu=vuoto>

https://it.wikipedia.org/wiki/Gestione_della_pandemia_di_COVID-19_in_Italia

<https://www.blog.it/post/riassunto-covid-19>

<https://www.humanitas.it/malattie/covid-19/>

Ringraziamenti

Al termine di questo lavoro, vorrei fare alcuni ringraziamenti, anche se non sarà facile farlo in poche righe.

*Ringrazio anzitutto il Professore **Giulio Palomba**, il mio relatore, per avermi mostrato la sua massima disponibilità e avermi fornito consigli e incoraggiamenti nei momenti di sconforto.*

*Ringrazio **Verusca** e a **Gaia**, le mie compagne di avventure, che hanno mosso con me i primi passi in università. E **Olga**, per aver condiviso tutte le preoccupazioni e le emozioni della fine di questo percorso: ha saputo ascoltarmi in ogni situazione ostica, mi è stata vicina in ogni fase critica, ma soprattutto è riuscita sempre a trovare le parole giuste per darmi un'ottima carica positiva.*

*A **Francesca**, all'amicizia che ci ha legate dopo il liceo. Mi ha ascoltata e supportata, sebbene le diverse strade che abbiamo preso ci hanno tenuto per molto tempo lontane.*

***Andrea**, la mia metà, ti sono riconoscente per tutto quello che hai fatto e fai per me, ti ringrazio per la capacità che hai sempre di alleggerire le mie ansie. Grazie anche alla tua famiglia, che si è sempre preoccupata per me in questi anni.*

*Grazie alle mie **nonne**, per aver gioito ad ogni mio traguardo, facendomi sentire la loro "reginetta". Ai miei **nonni**, che da lassù mi hanno costantemente protetta.*

*Un ringraziamento speciale va alla mia splendida **Famiglia**. Alla mia premurosa mamma, che con la sua dolcezza mi ha confortato nei momenti in cui sembrava*

impossibile. Al mio rassicurante e forte papà, che mi ha sempre incoraggiata e ha saputo tirar fuori la mia grinta per affrontare questo percorso. Al mio grande fratello Mirco, che non me l'ha mai dimostrato troppo, ma infondo non ha mai smesso di credere ed essere fiero di sua sorella.

*E infine grazie a **me**, che non mi sono mai arresa. Alla mia vittoria e alle prossime che raggiungerò. Mi auguro di indossare sempre il fantastico sorriso che mi dona questo momento.*