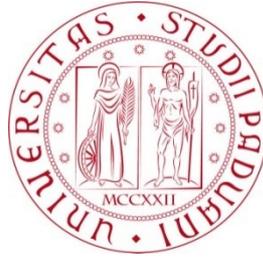


Università degli Studi di Padova
Dipartimento di Scienze Statistiche
Corso di Laurea Magistrale in
Scienze Statistiche



**La transizione lavoro - non lavoro. Analisi differenziale in
Italia prima e durante la pandemia di Covid-19**

Relatore: Prof. Gianpiero Dalla Zuanna
Dipartimento di Scienze Statistiche

Correlatore: Prof. Adriano Paggiaro
Dipartimento di Scienze Statistiche

Laureando: Riccardo Fusinato
Matricola N 1183869

Anno Accademico 2020/2021

*Mettiti in cammino anche se l'ora non ti piace.
Quando arriverai l'ora sarà comunque gradita.
(Proverbio africano)*

A Silvia e alla mia famiglia

INDICE

INTRODUZIONE E SOMMARIO	1
Il mercato del lavoro in Italia prima e durante la pandemia di Covid-19.....	1
La domanda di ricerca.....	3
La scelta della fonte dei dati.....	5
Struttura della tesi.....	5
CAPITOLO 1: LA RILEVAZIONE CONTINUA FORZE LAVORO	7
1.1 Il disegno campionario	9
1.2 Le tecniche di rilevazione	11
1.3 La struttura del questionario	12
1.4 Soggetti e variabili di interesse	17
CAPITOLO 2: DEFINIZIONE DEL SOTTOCAMPIONE DI RIFERIMENTO	19
2.1 Abbinamento longitudinale	20
2.2 Composizione socio-demografica alla prima rilevazione	22
2.3 Caratteristiche lavorative alla prima rilevazione.....	24
CAPITOLO 3: ANALISI DESCRITTIVE	27
3.1 Flussi per gli individui del gruppo pre-pandemia.....	28
3.2 Flussi per gli individui del gruppo durante la pandemia	29
3.3 Analisi disaggregata dei flussi trimestrali.....	30
3.3.1 Flussi per periodo e genere	30
3.3.2 Flussi per periodo e cittadinanza.....	33
3.3.3 Flussi per periodo e ripartizione geografica di residenza.....	35
CAPITOLO 4: ANALISI DEI DETERMINANTI DELLA PERDITA DEL LAVORO	39
4.1 Il modello logistico	40
4.2 Modello logistico senza interazioni	43
4.3 Modello logistico con interazioni con genere e cittadinanza.....	45
4.4 Modello logistico con interazioni con genere, cittadinanza e periodo.....	48
CAPITOLO 5: ANALISI CON PROPENSITY-SCORE MATCHING	51
5.1 Il propensity-score matching.....	52
5.2 Analisi della transizione lavorativa su soggetti comparabili.....	55

5.2.1 Periodo come trattamento.....	55
5.2.2 Cittadinanza come trattamento	56
5.2.3 Ripartizione geografica come trattamento.....	58
CAPITOLO 6: ULTERIORI EVIDENZE DAI DATI.....	61
6.1 Applicazione di diversi metodi di matching su differenti outcome	61
6.2 Analisi su definizione alternativa degli stati lavorativi.....	68
CAPITOLO 7: CONCLUSIONI E POSSIBILI SVILUPPI.....	71
BIBLIOGRAFIA	75
RINGRAZIAMENTI.....	77
APPENDICE	81
A1. Flussi trimestrali disaggregati.....	81
A2. Distribuzione caratteristiche pre-matching.....	84

INTRODUZIONE E SOMMARIO

L'analisi del mercato del lavoro in Italia è un tema che da molto tempo suscita interesse anche a livello dei singoli individui, e non più solamente a livello macro-economico e politico-sociale. I cittadini, infatti, si trovano davanti ad una situazione in cui, da qualche anno, trovare un lavoro non è più un'assicurazione sul suo mantenimento per tutto l'arco della vita lavorativa.

Non troppi anni fa si è assistito a quella che è stata la peggior crisi economica della storia dopo quella derivante dal crollo della borsa di New York del 1929. La crisi del 2008 ha infatti avuto lunghi strascichi sul mercato del lavoro italiano per diversi anni e ha mostrato segnali di ripresa solamente a partire dal 2013.

Così come recita il rapporto sul mercato del lavoro 2020 redatto congiuntamente da Ministero del lavoro e delle Politiche Sociali, ISTAT, INPS, INAIL e ANPAL, gli shock, indotti nel 2020 dalla pandemia e dalle necessarie misure di contrasto, mostrano attraverso i dati macroeconomici caratteri straordinari e pervasivi sul mercato del lavoro che ne hanno profondamente intaccato la dinamica precedente.

Ecco allora perché con questa tesi si è voluto andare ad analizzare la dinamica della transizione lavoro-non lavoro prima e durante la pandemia di Covid-19.

Il mercato del lavoro in Italia prima e durante la pandemia di Covid-19

Nel rapporto annuale 2019, al capitolo relativo al mercato del lavoro e il capitale umano, e nel report sul mercato del lavoro del 2019, l'ISTAT afferma che a cinque anni dall'avvio della ripresa economica (2013-2014), il mercato del lavoro italiano mostrava un sostanziale miglioramento, superando i livelli occupazionali pre-crisi e

riducendo progressivamente la forza lavoro non utilizzata nel sistema produttivo, che permaneva tuttavia ancora su livelli elevati. L'occupazione ha raggiunto in quell'anno il massimo storico di 23.4 milioni di unità nel secondo trimestre, con un tasso di disoccupazione in discesa stabile.

La ripresa dell'occupazione è riuscita però solo parzialmente a ridurre le vulnerabilità e i divari, che si erano acuiti durante la fase recessiva. Tra i dipendenti è aumentata infatti l'incidenza dei lavori a termine e tra gli indipendenti quella degli autonomi senza dipendenti. Contestualmente si sono ridotte le forme di lavoro permanente a tempo pieno, mentre è fortemente aumentato il part-time involontario, soprattutto per la componente femminile. Ha continuato ad ampliarsi il divario generazionale a favore della classe d'età più adulta e si è accentuata la disuguaglianza territoriale tra Centro-Nord e Sud del paese.

L'espansione dell'occupazione meno qualificata ha inoltre accentuato la segmentazione del mercato del lavoro, con la concentrazione dei lavoratori stranieri in occupazioni caratterizzate da bassi skill, da minori tutele e da retribuzioni più basse.

La ricomposizione del lavoro in favore di forme e settori relativamente più vulnerabili ha riguardato anche i giovani, meno presenti nel mercato del lavoro, ma sempre più istruiti, acuendo il problema del mismatch tra domanda e offerta di lavoro e, in questo contesto, la presenza di sovraistruzione, ovvero di occupati con un titolo di studio più alto rispetto a quello richiesto per la mansione svolta, ha rappresentato un fattore di ostacolo alla piena valorizzazione del capitale umano.

Ci si è quindi affacciati al 2020 in una sostanziale situazione in cui vi era sì un aumento costante dell'occupazione e una diminuzione della disoccupazione significativa, ma con le disuguaglianze evidenziate e quindi non in modo omogeneo per tutta la penisola e per tutti gli individui.

I medesimi rapporti ISTAT sul 2020, infatti, affermano che la pandemia da COVID-19 si è innestata su una situazione sociale caratterizzata da forti disuguaglianze, più ampie di quelle esistenti al momento della crisi del 2008-2009.

Un aspetto particolarmente critico in tale anno è stato rappresentato dall'elevato tasso di irregolarità dell'occupazione, più alto tra le donne, nel Mezzogiorno, tra i lavoratori molto giovani e tra quelli più anziani. È un segmento del mercato del lavoro strutturalmente debole e più esposto, nella particolare situazione seguita al

diffondersi dell'epidemia, a causa delle difficoltà di accesso agli ammortizzatori sociali e dell'impossibilità di giustificare formalmente nel lockdown gli spostamenti per motivi di lavoro. Inoltre, lo shock organizzativo familiare provocato dal lockdown ha potenzialmente interessato tutti i nuclei con figli minori ed entrambi i genitori, o l'unico genitore, occupati/o, a causa delle criticità legate alla carenza dei servizi per la prima infanzia e alle susseguenti difficoltà di conciliare i tempi di vita. Guardando strettamente ai numeri, nel 2020 si è assistito ad una sostanziale stagnazione degli occupati mensili nei primi due mesi dell'anno e successivamente ad una forte caduta, soprattutto tra marzo e aprile, quando la pandemia da Covid-19 ha colpito con la prima ondata.

Segni di ripresa si sono visti a partire da luglio ed erano ancora presenti a novembre, benché deboli, dovuti probabilmente alla forma più leggera delle misure di lockdown, agli interventi a sostegno delle attività economiche e, più in generale, al dispiegarsi di processi di apprendimento e adattamento. La perdita di occupazione registrata tra febbraio e giugno 2020 (-542 mila unità) è stata per metà recuperata tra luglio e novembre.

Le flessioni si sono concentrate soprattutto tra i dipendenti a termine e, in misura inferiore, tra gli indipendenti, a fronte di un incremento dello stock di dipendenti a tempo indeterminato. Questo provvisorio bilancio risente della natura dei provvedimenti di sostegno dell'occupazione, con la presenza di un'ampia gamma di ammortizzatori sociali, scaricandosi di più sulle ore lavorate, che hanno svolto un ruolo di "spugna" sui livelli di occupazione.

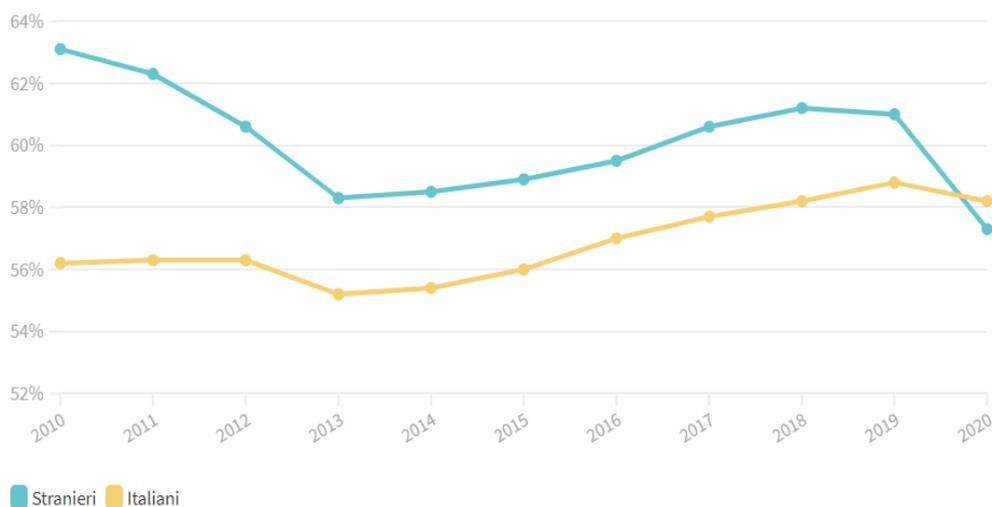
La domanda di ricerca

Nel giugno 2021, Enrico Di Pasquale e Chiara Tronchin di Fondazione Moressa, istituto di studi e ricerche nato nel 2002 allo scopo di diffondere la conoscenza del valore economico degli stranieri in Italia, hanno pubblicato uno studio trasversale su chi pagasse maggiormente la crisi da pandemia.

Da questo studio, operato sui dati relativi all'occupazione nel 2020, è risultato che le più colpite dalla crisi dovuta alla pandemia sono state le donne straniere. Sul totale dei posti persi tra il 2019 ed il 2020 oltre un terzo risultava attribuibile alla componente straniera e il 24% da imputarsi alle sole donne straniere. Gli stranieri

in generale hanno visto una diminuzione dell'occupazione maggiore rispetto agli autoctoni e per la prima volta il loro dato, pari al 57.3% sulla popolazione 15-64 anni, è sceso sotto al dato degli italiani, pari al 58.2%.

Figura 1: Serie storica del tasso di occupazione (15-64 anni) per cittadinanza dal 2010 al 2020.



Fonte: lavoce.info, rielaborazione dati ISTAT

Lo studio riporta come questo sia stato dovuto al fatto che le donne straniere siano più spesso occupate in lavori precari e che quindi anche misure di contenimento come il blocco dei licenziamenti e la cassa integrazione abbiano potuto apportare ben poco aiuto.

Stante questo fenomeno osservato, ci si è chiesti se si potesse osservare un'associazione del genere e della cittadinanza anche sulla perdita del lavoro analizzata in modo longitudinale, quindi seguendo uno stesso individuo per un certo periodo di tempo, andando ad indagare l'associazione anche con altre variabili, che potrebbero aver giocato un ruolo importante sulla transizione lavorativa degli individui inizialmente occupati.

In secondo luogo, sono state prese le mosse dallo studio di Paggiaro (2013) su Demographic Research che aveva analizzato le evidenze empiriche riguardanti la crisi del 2008 su italiani e stranieri. Tale studio ci ha spinti a chiederci se fosse possibile anche un effetto di composizione del campione sui differenti tassi di perdita del lavoro.

In particolare, ci si concentrerà su tre variabili che a priori ci aspettiamo possano aver giocato un ruolo importante sui tassi di transizione: il periodo di osservazione (prima o durante la pandemia), la cittadinanza e l'area geografica di residenza (con un confronto tra il Centro-Nord ed il Sud del paese).

La scelta della fonte dei dati

Necessitando di una fonte dei dati che permettesse un abbinamento longitudinale degli individui per poterne osservare le transizioni lavorative ad un certo periodo di distanza tra due rilevazioni, la scelta fisiologica è ricaduta sulla Rilevazione Continua delle Forze Lavoro (RCFL) operata da diversi anni da parte dell'Istituto Nazionale di Statistica (ISTAT).

Tale rilevazione, infatti, presenta molti vantaggi relativi alle analisi che si vogliono effettuare ed in particolare mette a disposizione tutta una serie di variabili ricostruite dall'ente relative alla condizione lavorativa e segue gli individui per un periodo di 15 mesi, permettendo abbinamenti a 3, 12 e 15 mesi, a seconda di quali occasioni di intervista si vanno ad abbinare, ricostruendo una sorta di setting di tipo panel.

Le variabili necessarie a permettere l'abbinamento non sono condivise nei dataset pubblici e sono state rese disponibili per questo elaborato grazie alla convenzione fra ISTAT e Dipartimento di Scienze Statistiche dell'Università degli Studi di Padova. Per le analisi che verranno eseguite, si è scelto un abbinamento a tre mesi.

Struttura della tesi

L'elaborato seguirà questo schema: è stata fin qui esposta la situazione di partenza da cui è stato preso spunto per formulare le domande di ricerca alle quali si vuole dare risposta nei capitoli seguenti. Inoltre, sono state esposte le domande di ricerca e una motivazione alla scelta della fonte utilizzata per i dati. Nel *Capitolo 1* viene presentata in modo più dettagliato, con una breve cronistoria e con i dettagli tecnici di campionamento e rilevazione, la fonte dei dati, cioè la RCFL. Alla fine del capitolo vengono anche presentati gli individui di interesse per le analisi e le variabili che si vanno a mantenere. Nel *Capitolo 2* viene in primo luogo mostrata la metodologia

utilizzata per eseguire l'abbinamento longitudinale degli individui, con una breve descrizione di quanti vengano rilevati in entrambi i casi. Successivamente vengono descritte le caratteristiche socio-demografiche e lavorative degli individui nei due periodi, prima e dopo la pandemia. A partire dal *Capitolo 3* si entra nel vivo delle analisi, con un'analisi descrittiva dei flussi di transizione da uno stato all'altro della condizione occupazionale e successivamente con delle analisi disaggregate per periodo e altre variabili di interesse, in modo da osservare se queste possano essere discriminanti nella probabilità di transitare da uno stato all'altro. Al *Capitolo 4* vengono mantenuti solamente gli individui che alla prima rilevazione risultavano essere occupati e si analizza, per mezzo di modelli di regressione, la loro probabilità di perdere il lavoro a tre mesi. Nel *Capitolo 5* viene proposto un tipo di analisi in cui si è voluto andare a vedere se quanto trovato nelle analisi precedenti reggesse anche dopo aver abbinato individui simili per caratteristiche socio-demografiche e lavorative. Tali analisi sono state svolte con una procedura di tipo propensity-score matching. Nel *Capitolo 6* vengono presentate alcune analisi eseguite al fine di testare la robustezza dei risultati. Nel *Capitolo 7*, infine, si traggono le conclusioni rispetto ai risultati delle analisi e vengono delineati possibili sviluppi e migliorie che si potrebbero apportare a partire dalle analisi eseguite.

CAPITOLO 1

LA RILEVAZIONE CONTINUA FORZE LAVORO

La Rilevazione Continua Forze Lavoro (RCFL) è un tipo di indagine eseguito dall'Istituto Nazionale di Statistica (ISTAT) che ha la finalità di ottenere informazioni sulla situazione lavorativa, sulla ricerca di lavoro e sugli atteggiamenti verso il mercato del lavoro della popolazione in età lavorativa. Rappresenta, in sostanza, la principale fonte di informazione statistica sul mercato del lavoro, sulla base della quale vengono derivate le stime ufficiali di occupati e disoccupati, nonché le informazioni sui principali aggregati dell'offerta di lavoro (Gazzelloni, 2006).

L'indagine è eseguita in Italia fin dal 1959 ed in principio veniva svolta con cadenza trimestrale nei mesi di gennaio, aprile, luglio e ottobre. Nel corso degli anni ha subito diverse modifiche sia nei metodi di rilevazione che nei contenuti, per adeguarsi sempre meglio alla continua evoluzione della società e del mercato del lavoro. Inoltre, alcune modifiche importanti sono state introdotte per motivi di armonizzazione con le rilevazioni a livello europeo. La modifica più importante è avvenuta con regolamento comunitario 577/98, entrato pienamente in uso a partire dal 2004, che prevede uno svolgimento da effettuarsi non più in una specifica settimana del trimestre per tutti, ma per tutte le 52 settimane di un anno mediante una distribuzione uniforme del campione nelle settimane, motivo per cui si parla di rilevazione "continua" e non più di rilevazione "trimestrale". Altri regolamenti importanti sono stati il 1575/2000 ed il 1897/2000 che definiscono rispettivamente le variabili target dell'indagine e il concetto di persona in cerca di occupazione. Tali innovazioni hanno comportato una radicale riorganizzazione del disegno di campionamento che verrà esposto nei paragrafi successivi. Infine, con il regolamento Ue 2019/1700, operativo dal 1° gennaio 2021, la rilevazione viene ulteriormente aggiornata stabilendo requisiti più dettagliati e vincolanti per le

statistiche europee su persone e famiglie. In particolare, cambiano le definizioni per l'identificazione della famiglia, dove la coabitazione rimane requisito fondamentale, ma ad esso si aggiunge anche il criterio della condivisione del reddito o delle spese (concetto dell'*housekeeping*), rendendo non più determinante l'esistenza di un legame di parentela o affettivo tra i membri della famiglia e per l'identificazione della condizione di occupato.

Tali modifiche non interesseranno i dati da noi utilizzati in quanto ci si è concentrati su anni non ancora toccati dalla modifica di quest'ultimo regolamento comunitario. I rispondenti all'intervista sono tutti gli individui tra i 15 e gli 89 anni appartenenti alle famiglie di fatto il cui intestatario risiede nel comune selezionato. Vengono perciò esclusi dalla popolazione i membri permanenti delle convivenze (cioè ospizi, istituti religiosi, caserme, ecc). Se uno di essi è assente o impegnato, è possibile fissare un altro appuntamento nell'arco del tempo previsto per la rilevazione oppure i quesiti vengono posti ad un altro familiare che risponde per conto dell'assente.

Fanno parte della rilevazione anche le persone di cittadinanza straniera, purché regolarmente iscritte alle anagrafi comunali italiane.

Per famiglia di fatto, prima dell'introduzione dell'ultimo regolamento comunitario citato poc'anzi e quindi relativamente ai dati che verranno da noi analizzati, si intende un insieme di persone legate o meno da vincoli di parentela o affettivi, dimoranti abitualmente nella stessa abitazione. Non rientrano tra questi coloro che vivono presso lo stesso domicilio per motivi puramente economici o momentanei come colf, badanti, baby-sitter ed ospiti, mentre vi rientrano lavoratori e studenti fuori sede assenti e dimoranti in un'altra abitazione esclusivamente per motivi di lavoro o studio.

Tra una rilevazione e la successiva, i componenti della famiglia possono variare per uscite da parte di alcuni componenti o per l'ingresso di nuovi componenti non presenti alle rilevazioni precedenti. Un criterio per l'uscita dalla famiglia è il cambio di residenza del componente. Tale criterio può anche non sussistere ad esempio nel caso di individui che lasciano la famiglia anche senza aver ancora effettuato il cambio di residenza, come studenti che vivono altrove mantenendosi da soli o i figli sposati che vanno a vivere con i consorti in un altro luogo.

1.1 Il disegno campionario

L'indagine viene condotta solo su una parte della popolazione. Ogni famiglia del campione viene estratta dalle Liste Anagrafiche dei Comuni campione (LAC) e rappresenta simbolicamente l'insieme teorico di famiglie aventi le stesse caratteristiche.

Ogni anno vengono intervistate circa 250mila famiglie complessivamente, per un totale di 600mila individui.

Il piano di campionamento è a due stadi, con stratificazione delle unità di primo stadio.

Le unità di primo stadio sono i comuni, mentre le unità di secondo stadio sono le famiglie.

Per le unità di primo stadio, a partire dagli 8100 Comuni Italiani, si opera una stratificazione all'interno di ciascuna provincia sulla base della dimensione demografica dei comuni. Tale stratificazione conduce all'identificazione di due tipologie di comuni:

- Comuni Autorappresentativi (AR): Comprendono tutti i capoluoghi di provincia e i comuni con una popolazione residente superiore ad una prefissata soglia, variabile da provincia a provincia. Tutti i comuni AR entrano a far parte in maniera permanente del campione. Quelli con dimensione demografica maggiore entrano nella rilevazione tutte le settimane, mentre gli altri AR una sola settimana al mese.
- Comuni Non Autorappresentativi (NAR): Comprendono tutti i comuni che non superano la soglia demografica prefissata per provincia. Solo alcuni NAR entrano nel campione ed essi vengono ulteriormente stratificati e da ciascuno strato viene estratto un solo comune campione con probabilità proporzionale alla dimensione demografica. Ognuno dei comuni NAR viene coinvolto nella rilevazione una settimana al mese, secondo uno schema di associazione casuale delle settimane ai comuni del campione.

Individuati i comuni coinvolti nell'indagine, si procede all'estrazione dalle LAC di un campione casuale semplice di unità di secondo stadio, le famiglie. Tale campione si compone di gruppi di quattro famiglie, detti quartine, in numero eguale a quello delle famiglie da estrarre nel comune stesso. Il primo elemento estratto (a sorte)

dalla quartina diventa la famiglia base, mentre i successivi elementi rappresentano delle “riserve” da utilizzare in caso di indisponibilità o rifiuto della famiglia base.

Il campione ha una struttura a panel ruotato, cioè per ogni famiglia vengono svolte quattro rilevazioni nell’arco di 15 mesi secondo uno schema di rotazione 2-2-2, per il quale una famiglia viene intervistata per due trimestri consecutivi, segue due trimestri consecutivi di pausa in cui non viene intervistata e successivamente viene nuovamente intervistata per altri due trimestri consecutivi.

In ogni trimestre il campione dell’indagine è formato perciò da quattro sotto-campioni detti gruppi di rotazione. In ogni occasione di rilevazione un gruppo si troverà alla prima intervista, un altro alla seconda, un altro alla terza ed infine un altro ancora alla quarta ed ultima intervista. Le famiglie alla quarta rilevazione successivamente escono definitivamente dal campione.

Da questo schema ne consegue che in via teorica il 50% delle famiglie sono reintervistate a distanza di 3 mesi, e il 50% a distanza di 12 mesi, mentre a tre e cinque trimestri vi è teoricamente una sovrapposizione del campione del 25%, a meno delle mancate risposte.

Figura 2: Schema di rotazione delle famiglie

Trimestre	Gruppi di rotazione					
I trimestre anno a	A4	B3		E2	F1	
II trimestre anno a		B4	C3		F2	G1
III trimestre anno a			C4	D3		G2 H1
IV trimestre anno a				D3	E3	H2 I1
I trimestre anno a+1				E4	F3	I2 J1
II trimestre anno a+1					F4	G3 J2 K1

Fonte: Gazzelloni (2006)

Per una stessa famiglia, e di concerto per uno stesso individuo, si dispone quindi in via teorica di un’osservazione al trimestre t, una al trimestre t+1, una al trimestre t+4, perciò ad un anno dalla prima, ed una al trimestre t+5, cioè ad un anno dalla seconda. Questo permetterà di creare per alcune variabili di interesse dei flussi trimestrali e/o annuali. Tale struttura longitudinale nei campioni trasversali non rispecchia un vero e proprio panel, in quanto le reinterviste avvengono solamente

per gli individui della popolazione “compresente”, residente cioè nella medesima abitazione della prima rilevazione.

Mentre le famiglie, da un trimestre all’altro, ruotano, i comuni campione rimangono sempre gli stessi nel tempo.

1.2 Le tecniche di rilevazione

La rilevazione avviene tramite l’impiego di una tecnica mista. La prima intervista viene effettuata con tecnica CAPI (Computer Assisted Personal Interview), in modo da favorire il coinvolgimento dell’interlocutore e permettere un primo contatto di persona tra l’intervistatore e l’intervistato, con il primo che ha la possibilità di illustrare passo passo le domande e verificare personalmente gli eventuali errori. Il rilevatore è supportato da un questionario elettronico che permette di controllare in tempo reale l’ammissibilità delle risposte, evidenziando eventuali incoerenze.

Tutte le interviste successive alla prima, per quelle famiglie che dispongono di un telefono, vengono invece somministrate con tecnica CATI (Computer Assisted Telephone Interview), dove l’intervistatore pone le domande agli interlocutori a mezzo telefonico e si avvale dello stesso questionario elettronico per seguirne il percorso e controllare le risposte.

Esistono due casi in cui le interviste non avvengono con CATI ma proseguono con CAPI: se l’intestatario della scheda di famiglia è straniero o qualora le famiglie non dispongano di telefono o vi siano situazioni particolari che evidenzino la necessità della presenza del rilevatore.

Nel periodo durante la pandemia di Covid-19, in considerazione dell’emergenza sanitaria e delle misure adottate dal Governo italiano per il suo contenimento, l’ISTAT ha intrapreso delle azioni per sopperire all’impossibilità o alla difficoltà oggettiva di effettuare le interviste di persona. Per questi motivi, anche nel caso di interviste CAPI il rilevatore contattava direttamente le famiglie telefonicamente ed era la famiglia stessa, nel caso, ad esprimere la preferenza per l’intervista faccia a faccia che avveniva solo previo consenso e nel rispetto dei protocolli sanitari stanziati.

1.3 La struttura del questionario

Il questionario elettronico è strutturato in sezioni, ognuna caratterizzata da uno specifico obiettivo.

La prima sezione (SG) è una scheda generale nella quale vengono registrate le informazioni anagrafiche, lo stato civile, informazioni sulla cittadinanza.

Questo avviene per tutti i componenti familiari.

Successivamente per i componenti con almeno 15 anni di età viene registrato dalla scheda generale anche il titolo di studio.

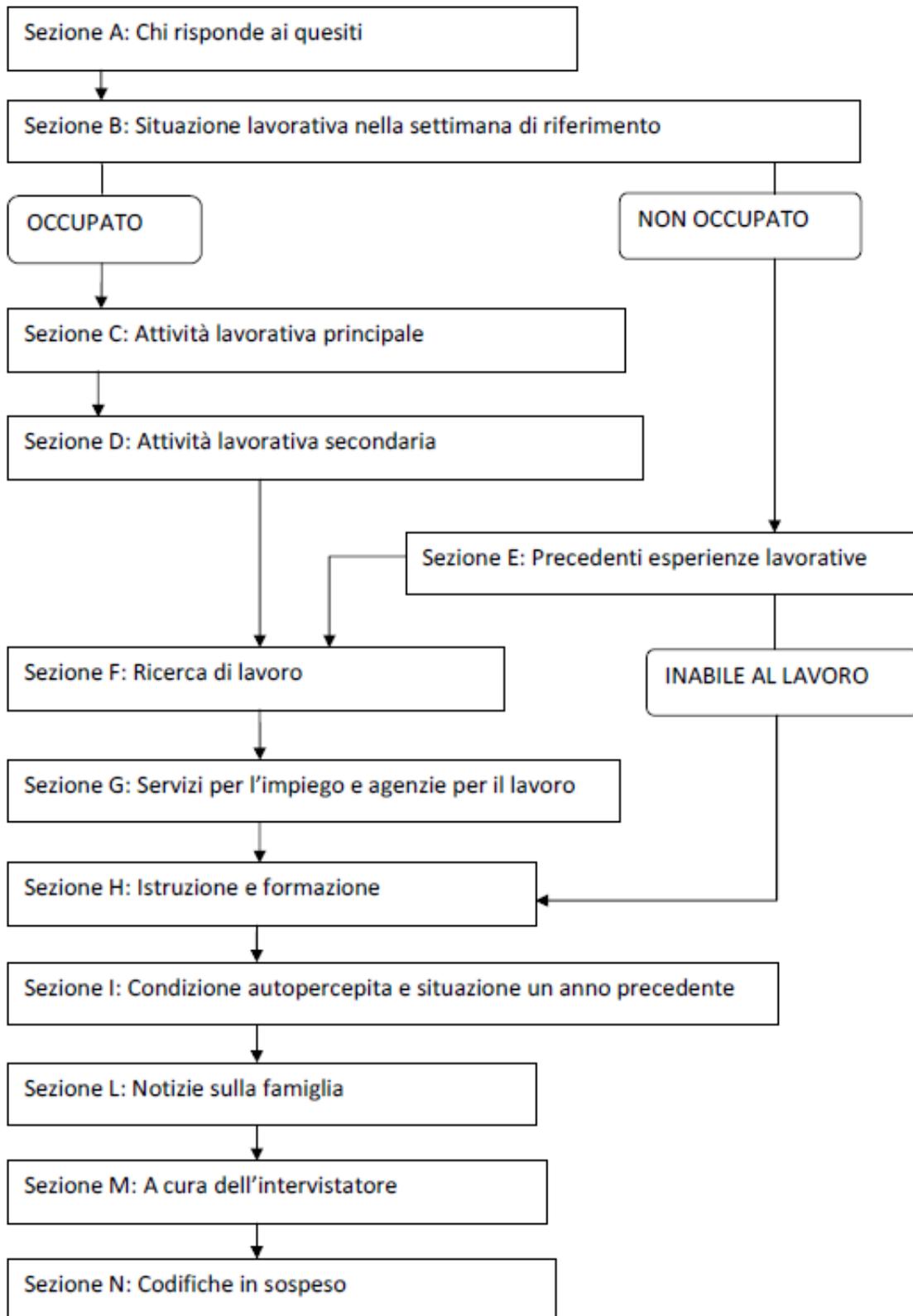
A questi quesiti risponde solamente l'intestatario, maggiorenne, della scheda, che dopo aver indicato il numero dei componenti della famiglia di fatto, fornisce le informazioni ad essa relative.

Conclusa la scheda generale, si passa al questionario individuale al quale rispondono solamente gli individui di almeno 15 anni compiuti, mentre per i minori di tale età il questionario si intende concluso con le sole informazioni della scheda generale.

Dopo la scheda generale, il questionario presenta nove sezioni, dalla A alla I, che vengono ripetute per ogni componente in età lavorativa, due sezioni di chiusura dell'intervista familiare, L ed M, e una sezione per la gestione delle codifiche in sospenso, N.

Nella figura alla pagina seguente è possibile vedere il diagramma di flusso del questionario con le varie sezioni.

Figura 3: Diagramma di flusso del questionario individuale



Fonte: Gazzelloni (2006)

Ogni intervistato ha un percorso personalizzato all'interno del questionario a seconda delle risposte che fornisce che consente di limitare i quesiti che gli vengono sottoposti da un minimo di 10, quando si tratta di un individuo disabile, ad un massimo di novanta, qualora l'intervistato sia un occupato con secondo lavoro.

Analizzando le sezioni così come riportate in Figura 3, si nota come nella sezione A vengano riportate le informazioni personali dell'intervistato e la data dell'intervista; nella sezione B si determini la condizione lavorativa dell'intervistato, passo cruciale per poter poi proseguire lungo la giusta direttrice; alle sezioni C e D rispondono solamente gli occupati e si raccolgono informazioni circa l'attività lavorativa principale ed eventualmente la secondaria; alla sezione E rispondono i non occupati (esclusi gli inabili al lavoro) e si indagano le precedenti esperienze lavorative; le sezioni F e G (precluse agli inabili al lavoro) riguardano la ricerca del lavoro ed eventualmente l'ausilio in tale ricerca per mezzo di agenzie del lavoro e servizi per l'impiego; la sezione H raccoglie informazioni sul percorso di istruzione e formazione dell'intervistato; la sezione I riguarda la situazione auto percepita nel momento dell'intervista e all'anno precedente; la sezione L ricostruisce ulteriori informazioni sulla famiglia per le successive interviste; le sezioni M ed N sono riservate al rilevatore, in cui può segnare note su come si è svolta l'intervista o eventuali codifiche in sospeso.

Per dare seguito ai cambiamenti richiesti dal nuovo regolamento comunitario, per le interviste a partire dal primo trimestre 2021 il questionario è stato predisposto in una nuova forma in cui alcuni quesiti vengono posti secondo una diversa sequenza rispetto al vecchio questionario. L'occasione è stata proficua anche per perfezionare il questionario sia razionalizzando alcune informazioni che già venivano raccolte, sia introducendo nuovi quesiti per approfondire temi più specifici come possono essere il percorso migratorio o il lavoro autonomo. Nel complesso, il nuovo questionario ha portato alla conferma di 164 quesiti, alla modifica di 89, all'eliminazione di 86 e all'introduzione di 48 nuove domande.

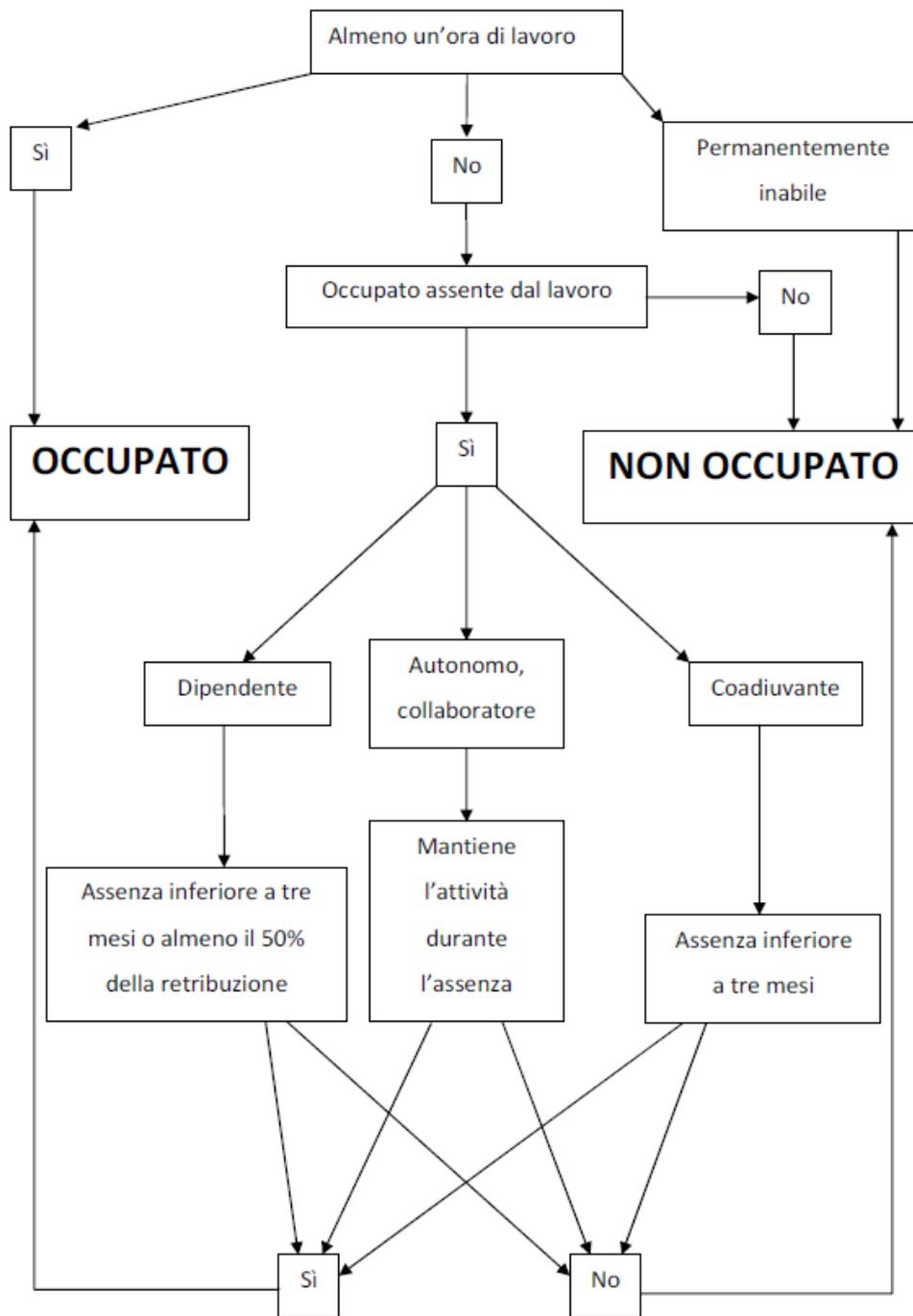
Visto che il fulcro di quello che si andrà ad analizzare sono le possibili situazioni rispetto alla condizione occupazionale divise in Occupazione, Disoccupazione ed Inattività, vale la pena soffermarsi sulla loro definizione¹:

¹ Definizione come riportata dal glossario di Veneto Lavoro. Sito raggiungibile al seguente link: <https://www.venetolavoro.it/glossario-e-strumenti>

- Occupati: Persone di 15 anni e oltre che nella settimana di riferimento presentano una delle seguenti caratteristiche: hanno svolto almeno un'ora di lavoro in una qualsiasi attività che preveda un corrispettivo monetario o in natura; hanno svolto almeno un'ora di lavoro non retribuito nella ditta di un familiare nella quale collaborano abitualmente; sono assenti dal lavoro (per ferie, cassa integrazione o malattia). I dipendenti assenti dal lavoro sono considerati occupati se l'assenza non supera i tre mesi o se durante l'assenza continuano a percepire almeno la metà della retribuzione. Coloro che sono lavoratori indipendenti, assenti dal lavoro, continuano ad essere considerati occupati se, durante l'assenza, mantengono comunque l'attività. Le condizioni riportate prescindono dalla sottoscrizione di un contratto di lavoro e pertanto gli occupati stimati tramite la RCFL comprendono anche le forme di lavoro irregolare.
- Disoccupati: Persone non occupate tra i 15 ed i 74 anni che abbiano effettuato almeno un'azione attiva di ricerca di lavoro nelle quattro settimane precedenti quella di riferimento e che sono disponibili a lavorare (o ad avviare un'attività autonoma) entro le due settimane successive. Rientrano nella categoria anche quanti inizieranno un lavoro entro tre mesi dalla settimana di riferimento e sarebbero disponibili a lavorare entro le due settimane successive qualora fosse possibile anticipare l'inizio del lavoro.
- Inattivi: Persone che non fanno parte delle forze lavoro e quindi non classificate come occupate o disoccupate. In tale categoria rientrano perciò: coloro che non hanno cercato un lavoro e non sono disponibili a lavorare, coloro che pur non avendo cercato un lavoro si sono comunque dichiarati disponibili a lavorare, coloro che hanno cercato un lavoro ma che non sono disponibili ad iniziare a lavorare entro due settimane dall'intervista.

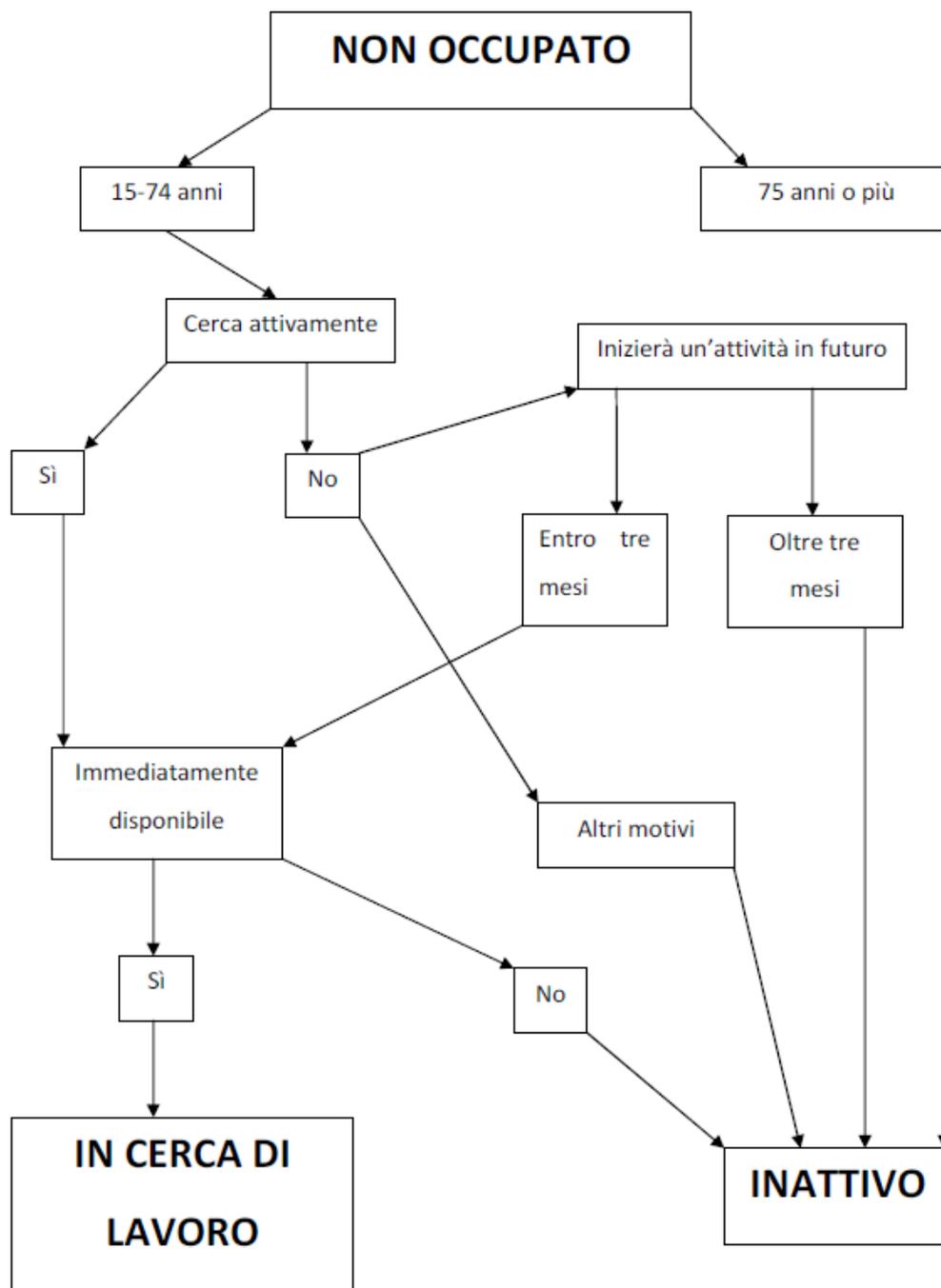
Per comprendere meglio secondo quali percorsi avviene l'assegnazione dell'individuo ad una delle tre categorie, vengono in aiuto due diagrammi di flusso, dal manuale "La rilevazione sulle forze di lavoro: contenuti, metodologie, organizzazione" a cura di S.Gazzelloni (2006), riportati in Figure 4 e 5.

Figura 4: Diagramma di flusso per l'individuazione degli occupati



Fonte: Gazzelloni (2006)

Figura 5: Diagramma di flusso per l'individuazione dei disoccupati e degli inattivi



Fonte: Gazzelloni (2006)

1.4 Soggetti e variabili di interesse

Il dataset, su cui si basano le analisi che verranno espone nel seguito della tesi, è composto dalle osservazioni di tutti quegli individui che avevano risposto ad entrambe le interviste ad un trimestre di distanza per i 9 trimestri considerati, dal IV trimestre del 2018 al IV trimestre del 2020. Per ognuno di questi individui si

hanno a disposizione le variabili rilevate dal questionario e le variabili ricostruite da ISTAT per sintetizzare le informazioni che si possono dedurre dal questionario.

Di tutte le molteplici variabili a disposizione, stante la natura delle analisi che si vogliono svolgere, solo alcune sono davvero di interesse e vengono perciò mantenute. In particolare, ci si concentra sulla variabile COND3, ricostruita da ISTAT sulla base delle risposte alla sezione B del questionario, che descrive la condizione lavorativa degli individui. Tale variabile presenta tre modalità, nello specifico:

- Occupati: COND3=1;
- In cerca di lavoro (Disoccupati): COND3=2;
- Inattivi: COND3=3.

Sfruttando la struttura longitudinale del campione, a partire da questa variabile si potranno creare i flussi di transizione da una condizione all'altra, vero fulcro dell'analisi.

In particolare per produrre anche delle analisi disaggregate in base al valore di alcune altre variabili di contesto, sono state mantenute nel dataset anche variabili relative alla situazione socio-demografica degli individui (Classe di età, ripartizione geografica di residenza, stato civile, numero di componenti del nucleo familiare, cittadinanza) e variabili relative al lavoro svolto dagli occupati (Posizione ricoperta, settore lavorativo, tipo di contratto, tipo di orario, esperienza nel lavoro attuale, situazione lavorativa l'anno precedente).

Sfruttando la possibilità di collegare gli individui facenti parte di una stessa famiglia, sono state create delle variabili che per ogni individuo ricostruissero anche delle caratteristiche familiari (presenza di minori di 15 anni nel nucleo familiare, presenza di maggiori di 64 anni nel nucleo familiare, presenza di diplomati, presenza di laureati, presenza di disoccupati e presenza di altri occupati).

CAPITOLO 2

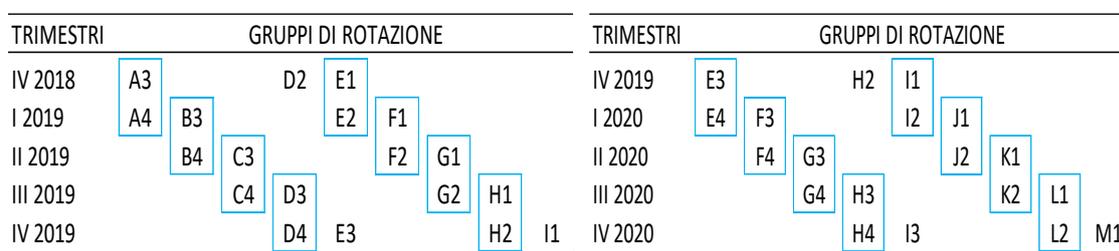
DEFINIZIONE DEL SOTTOCAMPIONE DI RIFERIMENTO

A partire dai dati della Rilevazione Continua delle Forze Lavoro (RCFL), è stato studiato lo schema di rotazione delle famiglie per ottenere un campione di individui che risultassero essere stati intervistati per due occasioni consecutive ad un trimestre di distanza. In particolare, volendo andare ad indagare le transizioni lavorative nel periodo pre-pandemico e durante la pandemia di Covid-19, sono stati presi in considerazione gli individui intervistati dal IV trimestre 2018 al IV trimestre 2020.

Con tali trimestri è stato possibile individuare i gruppi di rotazione da abbinare poiché nel primo caso avessero almeno una delle due rilevazioni, la prima o la seconda, nell'anno 2019, considerato periodo pre-pandemico, e nel secondo avessero almeno una delle due rilevazioni nell'anno 2020, considerato per intero periodo pandemico, dal momento che i primi casi di contagio sono stati registrati verso la fine di Gennaio 2020 e i primi decreti del Consiglio dei Ministri riguardanti limitazioni alle attività lavorative sono stati emanati nei primi giorni di Marzo.

Dalla figura sottostante (Figura 6) è possibile apprezzare quali gruppi di rotazione sono stati tra loro abbinati per ottenere il campione su cui sono state successivamente svolte le analisi. Una parte degli individui del gruppo pre-pandemico saranno presenti anche tra coloro che sono stati intervistati durante la pandemia, poiché proprio per lo schema di rotazione del campione questi erano alle prime due occasioni di intervista nel 2019 e alle seconde due occasioni nel 2020.

Figura 6: Schema di rotazione delle famiglie. Evidenziati i gruppi abbinati (a sinistra per il gruppo prima della pandemia, a destra per il gruppo durante la pandemia).



Fonte: Rielaborazione da Gazzelloni (2006)

2.1 Abbinamento longitudinale

Per identificare univocamente gli individui tra le varie occasioni di intervista non è presente una variabile che funga da codice. Dalla totalità delle variabili a disposizione è però possibile ricavare un insieme di queste ultime che componga una chiave di abbinamento. In particolare, conoscendo l'anno di estrazione (variabile ANNOES), la regione di residenza (REG), la provincia/città metropolitana di residenza (PROVCM), il comune di residenza (CODCOM), la quartina (CODQUA) e il codice della famiglia all'interno della quartina (CODFAM) è possibile identificare le varie famiglie nelle diverse occasioni di intervista. Una volta all'interno della famiglia è possibile abbinare gli individui facendo uso del codice individuale all'interno della famiglia (INDIV) e della data di nascita (ricomposta a partire dalle variabili che identificano giorno, SG19_1, mese, SG19_2, e anno di nascita, SG19_3) come ulteriore controllo di correttezza dell'abbinamento. Tali variabili, che hanno reso possibile l'abbinamento, sono state a nostra disposizione grazie alla convenzione aperta fra ISTAT e Dipartimento di Scienze Statistiche dell'Università degli Studi di Padova, poiché non sono generalmente rese pubbliche nei dataset scaricabili dalla banca dati aperta alla popolazione.

Così come riporta lo schema in figura 6, per abbinare gli individui in due occasioni di intervista ad un trimestre di distanza si devono abbinare quelli alla prima wave con quelli alla seconda nel trimestre successivo e quelli nella terza wave con quelli alla quarta nel trimestre successivo.

Dal momento che da una rilevazione all'altra un individuo potrebbe non voler più far parte del campione o potrebbe uscirne per cambi di residenza in comuni differenti o uscita dalla famiglia, è fisiologico che tra un'intervista e l'altra si possano perdere degli individui o addirittura intere famiglie. Per questo motivo, dopo aver eseguito l'abbinamento, sia per il gruppo del periodo pre-pandemia che per il gruppo del periodo pandemico, si è andato a vedere quale fosse la percentuale di famiglie che si andavano a perdere e la percentuale di individui rispondenti ad entrambe le interviste.

Per il periodo pre-pandemico, in media per ogni gruppo di rotazione sono state intervistate 16479 famiglie e le famiglie presenti ad entrambe le occasioni di intervista erano in media 15043 per ogni gruppo di rotazione. In media, ad un trimestre di distanza, si perdeva quindi poco meno del 10% delle famiglie della prima rilevazione (8.7%). Il numero medio di individui è invece pari a 36962 dei quali se ne perde, a tre mesi di distanza, il 9.7%.

Cifre simili si osservano anche per il numero medio di famiglie e di individui intervistati nel periodo pandemico, con una percentuale di intere famiglie o individui persi al follow-up comparabili e sempre nell'ordine del 9-10%.

Dalle tabelle seguenti è possibile osservare nel dettaglio il numero di famiglie e di individui alla prima rilevazione per ogni gruppo di rotazione e le percentuali di quante e quanti comparivano anche ad un trimestre di distanza, per i due periodi considerati.

Tabella 1: Famiglie ed individui intervistati prima della pandemia.

Gruppo di rotazione	Prima occasione di intervista considerata	Num. Famiglie prima rilevazione	Num. Famiglie presenti a un trimestre di distanza	% famiglie rispondenti a entrambe le occasioni	Num. Individui prima rilevazione	Num. Individui presenti ad un trimestre di distanza	% individui intervistati due volte sul totale degli individui alla prima rilevazione
A	Terza intervista	15875	14895	93.8%	35236	32527	92.3%
B	Terza intervista	16136	14919	92.5%	36007	32741	90.9%
C	Terza intervista	15991	14740	92.2%	35480	32039	90.3%
D	Terza intervista	15814	14628	92.5%	34780	31455	90.4%
E	Prima intervista	16952	15511	91.5%	38704	34649	89.5%
F	Prima intervista	17138	15507	90.5%	39229	34667	88.4%
G	Prima intervista	17032	15062	88.4%	38434	32904	85.6%
H	Prima intervista	16892	15078	89.3%	37824	32753	86.6%
Totale		131830	120340		295694	263735	

Fonte: Rielaborazione dati RCFL

Tabella 2: Famiglie ed individui intervistati durante la pandemia

Gruppo di rotazione	Prima occasione di intervista considerata	Num. Famiglie prima rilevazione	Num. Famiglie presenti a un trimestre di distanza	% famiglie rispondenti a entrambe le occasioni	Num. Individui prima rilevazione	Num. Individui presenti ad un trimestre di distanza	% individui intervistati due volte sul totale degli individui alla prima rilevazione
E	Terza intervista	15873	14416	90.8%	35214	31405	89.2%
F	Terza intervista	15558	14188	91.2%	34698	31154	89.8%
G	Terza intervista	15103	13947	92.3%	33350	30186	90.5%
H	Terza intervista	15780	14602	92.5%	34614	31377	90.6%
I	Prima intervista	16904	14741	87.2%	38113	32501	85.3%
J	Prima intervista	14333	12502	87.2%	32445	27691	85.3%
K	Prima intervista	16257	14758	90.8%	36345	31734	87.3%
L	Prima intervista	15926	14283	89.7%	35418	30743	86.8%
Totale		125734	113437		280197	246791	

Fonte: Rielaborazione dati RCFL

Avendo perciò considerato tutti gli individui facenti parti delle famiglie abbinate, il campione è stato ulteriormente pulito andando ad eliminare gli individui che erano presenti solamente ad una delle due interviste ed eliminando successivamente tutti quegli individui che alla prima intervista non avevano un'età compresa tra i 15 ed i 64 anni, età lavorativa, o che compivano il sessantacinquesimo anno di età nel corso del trimestre tra le due interviste.

Così facendo si hanno a disposizione per le analisi le osservazioni per 150490 individui nel periodo pre-pandemico e le osservazioni per 137536 individui nel periodo durante la pandemia.

Nelle successive analisi tali individui saranno utilizzati tutti per le analisi descrittive ed in seguito verranno selezionati solamente gli individui che alla prima rilevazione risultavano essere occupati per andare ad indagare la transizione verso la perdita del lavoro.

2.2 Composizione socio-demografica alla prima rilevazione

In entrambi i periodi, il campione è composto maggiormente da individui adulti, cioè di età compresa tra i 35 ed i 54 anni ed in particolare maggiormente nella classe 45-54 anni. I giovani intervistati, sia prima che durante la pandemia, compongono poco meno del 30% del campione.

Più della metà dei rispondenti in entrambi i periodi risulta essere di genere femminile, con un divario in numeri assoluti rispetto ai maschi di circa cinquemila unità.

Si ha una prevalenza di individui residenti al Nord Italia, di poco sotto alla metà del campione (46.90% sia prima che durante la pandemia).

Circa il 90% degli individui in entrambi i periodi risulta avere cittadinanza italiana, con gli individui di cittadinanza straniera che sono invece in maggioranza extra-comunitari, con un 31-32% di stranieri provenienti invece da paesi dell'Unione Europea (Ue27).

La maggior parte degli individui possiede un titolo di studio medio, cioè un diploma di scuola superiore di 3 o 5 anni, anche se vi è un'alta percentuale di quanti possiedono al massimo la licenza media (rispettivamente 41.30% e 39.70% degli intervistati nei due periodi).

Tabella 3: Distribuzione delle caratteristiche socio-demografiche per periodo di osservazione

Variabile	Livello	Prima della pandemia	Durante la pandemia
Età	Giovani	29.20%	29.12%
	Adulti	46.56%	45.51%
	Over54	24.23%	25.37%
Genere	Maschi	48.45%	48.22%
	Femmine	51.55%	51.78%
Cittadinanza	Italiani	89.80%	90.24%
	Stranieri UE	3.16%	3.14%
	Stranieri extra-UE	7.04%	6.62%
Titolo di studio	Alto	16.34%	17.14%
	Medio	42.36%	43.16%
	Nessuno/Basso	41.30%	39.70%
Residenza	Nord	46.90%	46.70%
	Centro	18.19%	18.62%
	Sud	34.91%	34.68%

2.3 Caratteristiche lavorative alla prima rilevazione

Viste le caratteristiche socio-demografiche, uno sguardo va dato anche alle caratteristiche lavorative degli intervistati nei due periodi.

Tabella 4: Distribuzione delle caratteristiche lavorative per periodo di osservazione

Variabile	Livello	Prima della pandemia	Durante la pandemia
Posizione lavorativa	Dirigente	1.66%	1.66%
	Quadro	5.23%	5.28%
	Impiegato	33.62%	34.72%
	Operaio	36.81%	36.09%
	Apprendista	0.58%	0.58%
	Lavoratore in casa per conto impresa	0.03%	0.02%
	Imprenditore	1.10%	1.06%
	Libero professionista	5.53%	5.50%
	Lavoratore in proprio	13.10%	12.87%
	Socio di cooperativa	0.13%	0.13%
	Coadiuvante azienda familiare	1.31%	1.27%
	Co.co.co	0.44%	0.45%
	Prestazione occasionale	0.45%	0.38%
Settore lavorativo	Agricoltura, sivecoltura e pesca	4.15%	4.19%
	Industria	19.59%	19.70%
	Costruzioni	6.01%	5.96%
	Commercio	13.63%	13.47%
	Alberghi e ristoranti	6.59%	6.08%
	Trasporto e magazzinaggio	4.76%	4.81%
	Informazione e comunicazione	2.33%	2.35%
	Attività finanziarie e assicurative	2.59%	2.68%
	Attività immobiliari, servizi alle imprese	10.96%	10.90%
	Amministrazione pubblica	6.00%	6.07%
	Istruzione, sanità e altri servizi sociali	15.82%	16.47%
Altri servizi collettivi e personali	7.56%	7.31%	

Nei due periodi risultavano essere occupati rispettivamente 87772 e 79508 individui intervistati, pari al 58.32% e al 57.80% dei rispettivi campioni.

I settori maggiormente rappresentati sono quelli dell'industria, dell'istruzione, sanità o servizi sociali e del commercio, con percentuali simili nei due periodi considerati.

Anche le posizioni ricoperte maggiormente risultano essere le stesse. Operai, impiegati o lavoratori in proprio sono infatti i tipi di inquadramento più frequenti nei due periodi.

Infine, sia gli intervistati prima della pandemia che quelli intervistati dopo, risultano essere impiegati mediamente nello stesso posto di lavoro da 13 anni.

CAPITOLO 3

ANALISI DESCRITTIVE

Grazie alla struttura longitudinale del dataset si può seguire in maniera sintetica ma comunque esaustiva come un individuo si muove tra uno stato e l'altro delle condizioni lavorative oppure se rimane nello stesso stato.

Dal momento che saremo interessati solamente alla perdita del lavoro senza guardare in modo dettagliato verso quale tipo di non occupazione si sposta l'individuo che perde il lavoro nel periodo di osservazione, per fare questo si è guardato principalmente alla variabile COND3.

Tabella 5: Frequenze assolute e percentuali situazione lavorativa intervistati per gruppo di appartenenza.

Stato occupazionale	Periodo pre pandemia		Periodo pandemia	
	Frequenza	Percentuale	Frequenza	Percentuale
Occupato	87772	58.32	79508	57.81
In cerca di lavoro	10166	6.76	8156	5.93
Inattivo	52552	34.92	49872	36.26
Totale	150490	100.00	137536	100.00

Da tabella 5 è possibile osservare come si distribuivano gli individui intervistati alla prima intervista considerata. In particolare, ricordando che sono stati mantenuti solamente individui di età compresa tra i 15 ed i 64 anni nel periodo di rilevazione, si osserva come, sia nel periodo prima che durante la pandemia, la quota maggiore di individui sia occupata. Alta è anche la quota di soggetti che ISTAT classifica come inattivi ed in particolare possiamo notare come in partenza tra i due gruppi vi sia una più bassa quota di occupati ed una più alta di inattivi per il gruppo di quanti

sono stati seguiti durante la pandemia, anche se con differenze di pochi punti percentuali.

Si vogliono ora esaminare i flussi trimestrali relativi a questa variabile e per fare questo si guarda congiuntamente alla variabile COND3 nelle due occasioni di rilevazione. Dal momento che già in partenza il campione era stato pulito dalla presenza di individui che risultavano rispondenti solamente a una delle due rilevazioni, non verranno perse in questo momento ulteriori osservazioni.

Si vedranno questi flussi distintamente per i due gruppi analizzati. In particolare, lungo la diagonale della tabella si posizioneranno gli individui che, a distanza di tre mesi, risulteranno essere nello stesso stato occupato alla prima rilevazione e che quindi non avranno subito transizioni. Fuori dalla diagonale si osservano invece gli spostamenti da uno stato all'altro in senso positivo (da uno stato di non occupazione ad uno di occupazione) ed in senso negativo (da uno stato di iniziale occupazione ad uno di ricerca di lavoro o inattività), transizioni alle quali saremo maggiormente interessati nel prosieguo delle analisi.

I valori delle percentuali si possono intendere come probabilità condizionate di transizione in un certo stato a tre mesi di distanza, dato il fatto di essere in un determinato stato alla prima rilevazione.

3.1 Flussi per gli individui del gruppo pre-pandemia

Si vanno ora ad osservare ed analizzare i flussi di transizione relativi agli individui seguiti nel 2019, cioè il periodo prima dell'insorgere della pandemia di Covid-19.

Da tabella 6 si possono trarre alcune conclusioni che andiamo a descrivere di seguito.

Tabella 6: Flussi trimestrali di COND3 nel periodo pre-pandemia.

Prima rilevazione	Seconda Rilevazione			Totale
	Occupato	In cerca di lavoro	Inattivo	
Occupato	84172 95.9	1171 1.33	2429 2.77	87772 100.00
In cerca di lavoro	1482 14.58	4530 44.56	4154 40.86	10166 100.00
Inattivo	1989 3.78	3448 6.56	47115 89.65	52552 100.00
Totale	87643 58.24	9149 6.08	53698 35.68	150490 100.00

In generale, per gli occupati la probabilità di risultare inoccupati a 3 mesi è minore del 5%. Chi ha un lavoro tende a mantenerlo e se lo lascia ne trova uno nuovo nell'arco dei tre mesi e risulta quindi ancora occupato, oppure tende maggiormente ad essere inattivo piuttosto che alla ricerca di un nuovo lavoro.

Per i disoccupati, cioè quanti sono classificati come "In cerca di lavoro", meno della metà (44.56%) risulta essere ancora tale a tre mesi di distanza. Poco meno del 15% riesce a trovare un lavoro e passa quindi tra le fila degli occupati, mentre una sostanziale quota superiore al 40% smette anche di cercare il lavoro o di essere disponibile a lavorare e diventa inattivo.

Per gli inattivi, la probabilità di trovare un lavoro nell'arco dei tre mesi è di poco inferiore al 4% e anche la quota di quanti iniziano la ricerca di un lavoro è bassa, pari al 6.56%. Quindi, la grande maggioranza, quasi 9 inattivi su 10, a tre mesi tendono a rimanere nella propria condizione di partenza.

3.2 Flussi per gli individui del gruppo durante la pandemia

Così come fatto per il gruppo precedente, anche per gli intervistati durante la pandemia si vanno a seguire i flussi di transizione. In tabella 7 si possono osservare quali sono state le dinamiche nei tre mesi.

Tabella 7: Flussi trimestrali di COND3 nel periodo durante la pandemia.

Prima rilevazione	Seconda Rilevazione			
	Occupato	In cerca di lavoro	Inattivo	Totale
Occupato	75703 95.21	1081 1.36	2724 3.43	79508 100.00
In cerca di lavoro	1140 13.98	3033 37.19	3983 48.84	8156 100.00
Inattivo	2092 4.19	3158 6.33	44622 89.47	49872 100.00
Totale	78935 57.39	7272 5.29	51329 37.32	137536 100.00

Anche in questo caso, così come osservato per il gruppo precedente, la maggior parte di chi ha un lavoro alla prima rilevazione tende a mantenerlo o a lasciarlo per un altro che trova nell'arco dei tre mesi. Di quanti perdono il lavoro è

sostanzialmente uguale la probabilità di diventare disoccupati, mentre risulta essere più alta la probabilità di diventare inattivo, pari al 3.43%.

Si fa più difficile trovare un nuovo lavoro per i disoccupati che in circa il 37% dei casi rimangono tali e in poco meno del 14% dei casi invece riescono a trovare un lavoro. Quasi la metà dei disoccupati, a tre mesi di distanza si ritrova a divenire inattivo (48.84%, poco meno di 8 punti percentuali in più rispetto al gruppo precedente).

Per gli inattivi rimane alta la probabilità di restare fermo allo stesso stato, con però una probabilità di poco più alta di trovare un lavoro (4.19%) e leggermente più bassa (6.33%) di iniziare la ricerca di un lavoro senza però trovarlo.

3.3 Analisi disaggregata dei flussi trimestrali

Viste le probabilità di transitare da uno stato all'altro della condizione lavorativa in generale, senza un controllo per altre caratteristiche degli individui, può risultare interessante anche un'analisi disaggregata dei flussi in base a specifiche caratteristiche che saranno di interesse anche per le successive analisi.

Vengono perciò proposte le suddivisioni relative al genere, alla cittadinanza e alla ripartizione geografica di residenza. In tutti e tre i casi si distinguerà anche per il periodo di osservazione e quindi ciascuna suddivisione sarà eseguita una volta sugli individui seguiti nel periodo prima della pandemia e una volta sugli individui seguiti durante la pandemia.

3.3.1 Flussi per periodo e genere

Nel periodo prima della pandemia si avevano 72908 maschi e 77582 femmine di età compresa tra i 15 ed i 64 anni, mentre durante la pandemia i numeri erano rispettivamente di 66324 e 71212 unità.

Nelle tabelle che vanno dalla 8 alla 11 si possono osservare i flussi trimestrali disaggregati per il genere.

Tabella 8: Flussi trimestrali di COND3 per i maschi nel periodo pre-pandemia.

	Seconda Rilevazione			
Prima rilevazione	Occupato	In cerca di lavoro	Inattivo	Totale
Occupato	47233 96.45	612 1.25	1124 2.3	48969 100.00
In cerca di lavoro	789 15.35	2517 48.97	1834 35.68	5140 100.00
Inattivo	872 4.64	1598 8.5	16329 86.86	18799 100.00
Totale	48894 67.06	4727 6.48	19287 26.45	72908 100.00

Tabella 9: Flussi trimestrali di COND3 per i maschi nel periodo durante la pandemia.

	Seconda Rilevazione			
Prima rilevazione	Occupato	In cerca di lavoro	Inattivo	Totale
Occupato	42500 96.06	574 1.3	1168 2.64	44242 100.00
In cerca di lavoro	582 14.66	1632 41.12	1755 44.22	3969 100.00
Inattivo	971 5.36	1519 8.39	15623 86.25	18113 100.00
Totale	44053 66.42	3725 5.62	18546 27.96	66324 100.00

Tabella 10: Flussi trimestrali di COND3 per le femmine nel periodo pre-pandemia.

	Seconda Rilevazione			
Prima rilevazione	Occupato	In cerca di lavoro	Inattivo	Totale
Occupato	36939 95.2	559 1.44	1305 3.36	38803 100.00
In cerca di lavoro	693 13.79	2013 40.05	2320 46.16	5026 100.00
Inattivo	1117 3.31	1850 5.48	30786 91.21	33753 100.00
Totale	38749 49.95	4422 5.7	34411 44.35	77582 100.00

Tabella 11: Flussi trimestrali di COND3 per le femmine nel periodo durante la pandemia.

	Seconda Rilevazione			
Prima rilevazione	Occupato	In cerca di lavoro	Inattivo	Totale
Occupato	33203 94.15	507 1.44	1556 4.41	35266 100.00
In cerca di lavoro	558 13.33	1401 33.46	2228 53.21	4187 100.00
Inattivo	1121 3.53	1639 5.16	28999 91.31	31759 100.00
Totale	34882 48.98	3547 4.98	32783 46.04	71212 100.00

Da queste tabelle è possibile fare un doppio confronto. Il primo lo si può fare tra individui di diverso genere per lo stesso periodo di osservazione e il secondo può essere operato per individui dello stesso genere nei due differenti periodi di osservazione.

Guardando le tabelle 8 e 10 possiamo confrontare le probabilità di transizione tra maschi e femmine prima della pandemia. Si può osservare come, per quanti risultavano inizialmente occupati, le femmine avessero una probabilità più alta dei maschi di transitare verso la disoccupazione (1.44% vs. 1.25%) e una ancor più alta probabilità di transitare all'inattività (3.36% vs. 2.30%). Le femmine facevano registrare anche una maggior probabilità di passare all'inattività da una situazione iniziale di disoccupazione (46.16% vs. 35.68%) e una minor probabilità di trovare un lavoro se erano alla ricerca (13.79% vs. 15.35%). Le medesime conclusioni si possono trarre per le femmine inattive rispetto ai maschi inattivi.

Da tabelle 9 e 11 si può osservare invece la situazione nel periodo durante la pandemia. Le dinamiche di transizione differenti tra maschi e femmine permangono le medesime, con delle percentuali differenti ma divari simili.

Andando invece ora a confrontare gli individui dello stesso genere nei due periodi differenti, si osserva come, sia i maschi sia le femmine del periodo durante la pandemia, abbiano fatto registrare delle probabilità più elevate di transitare a stati di disoccupazione o inattività. Di concerto diminuiva anche la probabilità di trovare lavoro nel trimestre.

Se per i maschi l'aumento di probabilità di perdere il lavoro era nell'ordine di circa 0.4 punti percentuale, era invece molto più grande il divario per le femmine confrontando prima della pandemia con durante la pandemia. In particolare, durante la pandemia la probabilità per le femmine di perdere il lavoro aumenta di 1.05 punti percentuali, interamente dovuti all'aumento della probabilità di transitare verso l'inattività.

3.3.2 Flussi per periodo e cittadinanza

Dal campione a disposizione, nel 2019 erano residenti in Italia 4756 cittadini stranieri comunitari e 10599 extra-comunitari, numeri che nel 2020 sono passati rispettivamente a 4321 e 9107 individui.

In questo caso si vanno ad unire i cittadini stranieri, senza quindi distinguere per comunitari ed extra-comunitari, poiché nelle analisi successive si è scelto di considerare congiuntamente tutti gli stranieri.

Nelle tabelle seguenti (dalla 12 alla 15) vengono perciò riportati i flussi trimestrali della condizione lavorativa per italiani e stranieri nei due periodi. In appendice verranno riportate anche le tabelle con stranieri comunitari ed extra-comunitari disaggregati.

Tabella 12: Flussi trimestrali di COND3 per i cittadini italiani nel periodo pre-pandemia.

Prima rilevazione	Seconda Rilevazione			
	Occupato	In cerca di lavoro	Inattivo	Totale
Occupato	75480 96.01	951 1.21	2187 2.78	78618 100.00
In cerca di lavoro	1213 14.14	3635 42.36	3733 43.5	8581 100.00
Inattivo	1753 3.66	3061 6.39	43122 89.96	47936 100.00
Totale	78446 58.05	7647 5.66	49042 36.29	135135 100.00

Tabella 13: Flussi trimestrali di COND3 per i cittadini italiani nel periodo durante la pandemia.

	Seconda Rilevazione			
Prima rilevazione	Occupato	In cerca di lavoro	Inattivo	Totale
Occupato	68529 95.47	885 1.23	2364 3.29	71778 100.00
In cerca di lavoro	949 13.59	2499 35.8	3533 50.61	6981 100.00
Inattivo	1845 4.07	2781 6.13	40723 89.8	45349 100.00
Totale	71323 57.47	6165 4.97	46620 37.56	124108 100.00

Tabella 14: Flussi trimestrali di COND3 per i cittadini stranieri nel periodo pre-pandemia.

	Seconda Rilevazione			
Prima rilevazione	Occupato	In cerca di lavoro	Inattivo	Totale
Occupato	8692 94.95	220 2.4	242 2.64	9154 100.00
In cerca di lavoro	269 16.97	895 56.47	421 26.56	1585 100.00
Inattivo	236 5.11	387 8.38	3993 86.5	4616 100.00
Totale	9197 59.9	1502 9.78	4656 30.32	15355 100.00

Tabella 15: Flussi trimestrali di COND3 per i cittadini stranieri nel periodo durante la pandemia.

	Seconda Rilevazione			
Prima rilevazione	Occupato	In cerca di lavoro	Inattivo	Totale
Occupato	7174 92.81	196 2.54	360 4.66	7730 100.00
In cerca di lavoro	191 16.26	534 45.45	450 38.3	1175 100.00
Inattivo	247 5.46	377 8.34	3899 86.2	4523 100.00
Totale	7612 56.69	1107 8.24	4709 35.07	13428 100.00

Guardando le tabelle 12 e 14 possiamo fare un confronto delle probabilità di italiani e stranieri nel 2019, mentre guardando le tabelle 13 e 15 si può fare il medesimo confronto per gli individui seguiti nel 2020.

Nel periodo precedente la pandemia gli occupati stranieri presentavano una minor probabilità di mantenere il proprio lavoro, nell'ordine di circa un punto percentuale (94.95% vs. 96.01% degli italiani). In particolare, gli occupati stranieri presentavano il doppio della probabilità degli italiani di transitare verso la disoccupazione (2.4% vs. 1.21%), ma con una più bassa probabilità di transitare verso l'inattività (2.64% vs. 2.78%).

Gli stranieri inoltre sembrava si "scoraggiassero" meno, dove per "scoraggiarsi" si intende transitare dalla disoccupazione all'inattività, poiché presentavano una probabilità più alta di rimanere nella condizione di disoccupato (56.47% vs. 42.36%) e avevano anche una probabilità più alta di trovare un lavoro (16.97% vs. 14.14%).

Analizzando la medesima dinamica nel periodo della pandemia, si evidenziano le stesse differenze anche se aumentano le probabilità di perdere il lavoro e diminuiscono quelle di rimanere nello stato di partenza sia per italiani che per stranieri.

Confrontando perciò i cittadini italiani nel 2019 con quelli nel 2020 si nota come aumenti di circa mezzo punto percentuale la probabilità di perdere il lavoro e che in particolare quanti perdono il lavoro transitino maggiormente verso uno stato di inattività (3.29% vs 2.78%).

Risultava anche essere maggiore la probabilità di scoraggiarsi se si partiva da una situazione di disoccupazione, infatti un disoccupato ha una probabilità del 50.61% di diventare inattivo nel periodo della pandemia rispetto al 43.5% prima della pandemia.

Per gli stranieri, al confronto tra i due periodi, nel 2020 raddoppia la probabilità di transitare dall'occupazione all'inattività (4.66% vs. 2.64%) ed aumenta sensibilmente la probabilità di diventare inattivi dopo una iniziale situazione di disoccupazione (38.3% vs. 26.56%), mentre di concerto diminuisce la probabilità di trovare un lavoro.

3.3.3 Flussi per periodo e ripartizione geografica di residenza

Dalle analisi descrittive di partenza ricordiamo che, nel 2019 e nel 2020, risultavano essere residenti al Nord Italia rispettivamente 70583 e 64235 individui. I residenti

al Nord risultavano essere la maggioranza degli individui in entrambi i periodi osservati, con i residenti al Centro Italia pari rispettivamente a 27378 e 25604 e quelli residenti al Sud e nelle Isole pari a 52529 e 47697.

Per omogeneità con le analisi successive, si vanno ad unificare i residenti al Centro con i residenti al Nord, distinguendo quindi solamente per due ripartizioni: Centro-Nord e Sud.

Nelle tabelle seguenti dalla 16 alla 19 si riportano quindi i flussi trimestrali della condizione lavorativa per residenti al Centro-Nord e al Sud nei due periodi. In appendice verranno riportate anche le tabelle con residenti al Centro e al Nord disaggregati.

Tabella 16: Flussi trimestrali di COND3 per i residenti al Centro-Nord nel periodo pre-pandemia.

	Seconda Rilevazione			
Prima rilevazione	Occupato	In cerca di lavoro	Inattivo	Totale
Occupato	61929 96.58	719 1.12	1476 2.30	64124 100.00
In cerca di lavoro	973 18.80	2289 44.22	1914 36.98	5176 100.00
Inattivo	1147 4.00	1399 4.88	26115 91.12	28661 100.00
Totale	64049 65.38	4407 4.50	29505 30.12	97961 100.00

Tabella 17: Flussi trimestrali di COND3 per i residenti al Centro-Nord nel periodo durante la pandemia.

	Seconda Rilevazione			
Prima rilevazione	Occupato	In cerca di lavoro	Inattivo	Totale
Occupato	55749 95.84	697 1.20	1720 2.96	58166 100.00
In cerca di lavoro	743 17.56	1575 37.23	1913 45.21	4231 100.00
Inattivo	1187 4.33	1390 5.07	24865 90.61	27442 100.00
Totale	57679 64.20	3662 4.08	28498 31.72	89839 100.00

Tabella 18: Flussi trimestrali di COND3 per i residenti al Sud nel periodo pre-pandemia.

Prima rilevazione	Seconda Rilevazione			
	Occupato	In cerca di lavoro	Inattivo	Totale
Occupato	22243 94.06	452 1.91	953 4.03	23648 100.00
In cerca di lavoro	509 10.20	2241 44.91	2240 44.89	4990 100.00
Inattivo	842 3.52	2049 8.58	21000 87.90	23891 100.00
Totale	23594 44.92	4742 9.03	24193 46.06	52529 100.00

Tabella 19: Flussi trimestrali di COND3 per i residenti al Sud nel periodo durante la pandemia.

Prima rilevazione	Seconda Rilevazione			
	Occupato	In cerca di lavoro	Inattivo	Totale
Occupato	19954 93.50	384 1.80	1004 4.70	21342 100.00
In cerca di lavoro	397 10.11	1458 37.15	2070 52.74	3925 100.00
Inattivo	905 4.03	1768 7.88	19757 88.08	22430 100.00
Totale	21256 44.56	3610 7.57	22831 47.87	47697 100.00

Dalle tabelle 16 e 18 si possono osservare le differenze nelle probabilità di transizione tra i residenti al Centro-Nord e quelli al Sud per il periodo pre-pandemico.

In particolare, per quanti risultavano essere occupati alla prima rilevazione, una quota consistente lo era anche alla seconda, con una differenza di poco meno di tre punti percentuali a sfavore dei residenti al Sud. Questi ultimi avevano anche una probabilità maggiore di transitare alla disoccupazione (1.91% vs. 1.12%) o all'inattività (4.03% vs. 2.30%).

Per i disoccupati, invece, la probabilità di rimanere nel medesimo stato a tre mesi di distanza era simile. Dove vi sono differenze marcate è nella probabilità di trovare un lavoro, con i residenti al Sud penalizzati (10.20% vs. 18.80% dei residenti al Centro-

Nord), facendo registrare anche una probabilità di transitare all'inattività maggiore di 7.91 punti percentuali.

Operando i medesimi confronti tra residenti al Centro-Nord e al Sud durante la pandemia, le dinamiche che si osservano sono le medesime del periodo precedente, ma con probabilità di transitare ad una condizione peggiore rispetto a quella di partenza più alte.

Confrontando invece i residenti al Centro-Nord nei due periodi, per gli occupati si assiste ad una simile probabilità di transitare alla disoccupazione, mentre aumenta di 0.66 punti percentuali la probabilità di perdere il lavoro e diventare inattivi. Guardando ai disoccupati risulta leggermente inferiore la probabilità di trovare un lavoro per quanti sono stati seguiti durante la pandemia ed una maggior probabilità di transitare invece all'inattività con una differenza di 8.23 punti percentuali.

Andando a vedere le stesse dinamiche nei due periodi per i residenti al Sud si osservano una minor probabilità di transitare verso la disoccupazione per quanti inizialmente occupati ed un aumento della probabilità di diventare inattivi di 0.67 punti percentuali. Per i disoccupati aumenta molto la probabilità di scoraggiarsi e transitare all'inattività, in particolare di 7.85 punti percentuali con più della metà dei disoccupati alla prima rilevazione durante la pandemia che, a tre mesi di distanza, vedono peggiorare la propria condizione (52.74%).

CAPITOLO 4

ANALISI DEI DETERMINANTI DELLA PERDITA DEL LAVORO

La serie di analisi bivariate finora svolte, miravano semplicemente a descrivere i flussi di transizione da uno stato all'altro della condizione occupazionale e come questi cambiassero in relazione ad alcune caratteristiche socio-demografiche degli individui, prese singolarmente.

Se l'obiettivo è però quello di considerare simultaneamente tali caratteristiche ed eventualmente una più ampia gamma delle variabili a disposizione, il passo successivo sta nell'utilizzare una logica di regressione e adattare un modello statistico ai dati.

Dal momento che ci si concentrerà sulla perdita del lavoro e quindi sull'aggregazione delle possibili transizioni dallo stato di occupato allo stato di disoccupato o di inattivo, la variabile risposta assume la forma di una variabile binaria con valore 1 nel caso in cui l'individuo occupato dopo tre mesi non lo sia più, valore 0 se invece l'individuo occupato alla prima intervista manterrà tale caratteristica anche alla seconda rilevazione.

Ci si concentra solamente su quanti alla prima rilevazione in entrambi i periodi risultavano essere occupati e il dataset degli occupati è poi stato ulteriormente depurato eliminando le osservazioni di quanti presentavano dei valori mancanti all'interno delle variabili da porre in regressione. Alla fine di queste operazioni si è ottenuto un campione di 130697 occupati alla prima rilevazione.

Tra tutti i possibili modelli per l'analisi del fenomeno in oggetto, stante la natura della variabile risposta, il più adatto all'analisi è il modello di regressione logistica, applicato proprio nei casi in cui la variabile dipendente è di tipo dicotomico.

4.1 Il modello logistico

Il modello logistico ha avuto le sue prime applicazioni negli studi biomedici, ad esempio per valutare come la presenza/assenza di alcune malattie fosse legata a dei comportamenti o stili di vita specifici. In seguito, la sua applicazione si diffuse anche a molti altri ambiti come le scienze sociali, le ricerche di mercato e la finanza. Per mezzo del modello di regressione logistica è possibile determinare la probabilità che un individuo con determinate caratteristiche appartenga ad una determinata classe, nel caso delle nostre analisi la classe di quanti perdono il lavoro a tre mesi dalla prima intervista (Salvan et al., 2020).

Dal momento che la probabilità è strettamente vincolata ad assumere valori che siano compresi nell'intervallo (0,1), un modello di regressione semplice non fa al caso delle analisi poiché potrebbe produrre risultati al di fuori di tale intervallo, non ammissibili. Tramite la sua funzione di legame, invece, il modello di regressione logistica permette di ottenere risultati nel dominio della variabile risposta poiché è una funzione $g: [0,1] \rightarrow \mathbb{R}$ monotona crescente.

Questo tipo di modelli è quindi descritto dalla funzione:

$$\text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k = X\beta$$

Dove

$$\text{logit}(p) = \ln\left(\frac{p}{1-p}\right)$$

è la funzione logit e p la probabilità che l'evento "perdita del lavoro" si verifichi. Una volta stimato il modello con il metodo della massima verosimiglianza ed ottenute le stime $\hat{\beta}$ dei coefficienti, per ottenere la stima della probabilità di perdere il lavoro dato il valore di alcune covariate si applica la funzione inversa del logit

$$\hat{p} = \frac{e^{x\hat{\beta}}}{1 + e^{x\hat{\beta}}}$$

Dopo aver brevemente esposto la teoria relativa al modello di regressione logistica, di seguito vengono esposte le variabili inserite nel modello. Tali variabili sono state ricostruite a partire da altre variabili che già ISTAT aveva ricostruito o direttamente dalle risposte al questionario.

- ETA: Età alla prima rilevazione. Si distinguono tre classi: Giovani, individui dai 15 ai 34 anni, gli Adulti, con età compresa tra i 35 ed i 54 anni e gli Over54, cioè individui di età compresa tra i 55 ed i 64 anni.
- RIP3: Ripartizione geografica di residenza. È la classica suddivisione in tre modalità operata da ISTAT delle regioni italiane. Si ha il Nord (Valle d'Aosta, Piemonte, Liguria, Lombardia, Trentino-Alto Adige, Emilia-Romagna, Veneto e Friuli-Venezia Giulia), il Centro (Toscana, Umbria, Marche e Lazio) e Sud (Abruzzo, Molise, Campania, Basilicata, Puglia, Calabria, Sicilia e Sardegna).
- STUD: Titolo di studio posseduto dall'individuo. Viene suddiviso in tre modalità a partire dalla variabile ISTAT TISTUD. Le modalità sono Basso (Nessun titolo di studio o fino alla licenza media), Medio (diploma di scuola superiore di 3 o 5 anni) e Alto (laurea triennale o superiore o titolo equivalente).
- CAT12: Settore lavorativo. Proviene da una rielaborazione della variabile CAT12 ISTAT che riporta 12 modalità, qui però ristrette a 10: Agricoltura, Industria in senso stretto, Costruzioni, Commercio, Ricettive (Alberghi e ristoranti), Logistica (Trasporto e magazzinaggio), Servizi pubblici (che comprendono servizi di informazione e comunicazione, amministrazione pubblica e difesa assicurazione sociale obbligatoria), Imprenditoria (in cui rientrano le attività immobiliari, i servizi alle imprese e altre attività imprenditoriali), Istruzione, Sanità e servizi sociali ed infine Altri servizi alla persona.
- DETIND: Tipo di contratto. Si distingue in contratto a tempo indeterminato e contratto a tempo determinato.
- PIEPAR: Orario di lavoro. Si distingue in full-time e part-time.
- DURLAV: Esperienza lavorativa in anni presso il luogo di lavoro della prima rilevazione. Viene raggruppata in classi a seconda che fosse inferiore a 2 anni, dai 3 ai 5 anni, dai 6 ai 10 anni, dagli 11 ai 20 anni o maggiore di 20 anni.

- COMP: Numero dei componenti del nucleo familiare. Possono essere 1, 2, 3, 4 o 5 e più.
- MIN15: Presenza di minori di 15 anni nel nucleo familiare.
- MAG64: Presenza di maggiori di 64 anni nel nucleo familiare.
- UNOC; Presenza di un altro occupato oltre all'individuo nel nucleo familiare.
- STRANIERO: Variabile indicatrice della cittadinanza dell'individuo. Vale 1 per gli individui con cittadinanza straniera (UE o extra-UE) e 0 per gli individui con cittadinanza italiana.
- SESSO: Variabile indicatrice del genere dell'individuo. Vale 1 per le femmine.
- ANNO: Variabile a due modalità, 2019 e 2020, che indica di quale gruppo fa parte l'individuo, se di quello intervistato prima della pandemia o durante la pandemia.

Essendo le variabili sopra riportate dei fattori, si deve decidere quali modalità prendere come riferimento. L'individuo baseline ha perciò le seguenti caratteristiche: è un maschio adulto, residente al Nord in un nucleo familiare composto da 3 persone, senza minori di 15 anni o maggiori di 64 anni e senza altri occupati, con titolo di studio medio, impiegato a tempo indeterminato full-time nel settore dell'industria dove ha dai 6 ai 10 anni di esperienza, facente parte del gruppo degli individui seguiti nel periodo precedente la pandemia.

Per l'interpretazione dei coefficienti, perciò, ci si deve sempre ricordare delle caratteristiche dell'individuo di riferimento. Operando la funzione inversa del logit sulla stima dell'intercetta si trova la probabilità di perdere il lavoro per l'individuo di riferimento. Andando a sommare i coefficienti relativi ad altri valori delle variabili considerate diversi da quelli di riferimento e applicando su quel valore la funzione inversa del logit si ottiene la probabilità di perdere il lavoro dell'individuo che ha le medesime caratteristiche dell'individuo baseline fatta eccezione per le variabili di cui si è preso in considerazione un diverso valore.

A differenza di un modello lineare, quindi, il coefficiente stimato non è di per sé l'effetto sulla probabilità, ma è l'effetto rispetto al valore di riferimento della variabile.

Come prima analisi, per capire quali fossero le variabili associate con la perdita del lavoro, viene stimato un modello senza interazioni. A partire da quello verranno poi

inseriti i termini di interazione con il genere, la cittadinanza ed il periodo. Di seguito si riportano i risultati.

4.2 Modello logistico senza interazioni

Si riportano in tabella 20 le stime dei coefficienti del modello di regressione.

Guardando al valore dell'intercetta ed applicando la trasformazione inversa del logit, così come riportata nel paragrafo teorico, si può asserire che per l'individuo baseline la probabilità di perdere il lavoro è pari al 2%.

Su tale probabilità agiscono in modo differente diverse variabili tra quelle considerate. In particolare, la deviazione dalle caratteristiche di riferimento che comporta il maggior aumento di probabilità di perdita del lavoro è data dall'aver un contratto a tempo determinato anziché indeterminato. Gli individui con tale tipo di contratto registrano infatti una probabilità pari al 10.43%, maggiore rispetto all'individuo di riferimento di più di 8 punti percentuali al variare di una sola caratteristica.

Ad avere un ruolo di peggioramento della probabilità di perdita del lavoro vi è anche l'età. Sia gli individui più giovani rispetto alla baseline che quelli più vecchi presentano probabilità maggiori di perdere il lavoro, pari rispettivamente al 2.22% e al 3.50%.

Analizzando come varia la probabilità cambiando solo la ripartizione di residenza si assiste ad una differenza non significativa tra i residenti al Nord (baseline) e residenti al Centro Italia, mentre risulta essere un fattore di svantaggio risiedere al Sud. I residenti al Sud con le medesime caratteristiche socio-demografiche e lavorative dei residenti al Nord risultano penalizzati di 1.09 punti percentuali sulla probabilità di perdere il lavoro, che sale al 3.09%.

Tabella 20: Stime del modello logistico (** p-value<0.01, * p-value<0.05).

VARIABILE	LIVELLO	COEF	STD.ERROR	P-VALUE	SIGNIF.
Età	Giovani	0.106	0.036	0.003	**
	Over54	0.573	0.041	<0.001	**
Residenza	Centro	0.051	0.038	0.159	
	Sud	0.446	0.022	<0.001	**
Titolo di studio	Medie o meno	0.226	0.032	<0.001	**
	Laurea o più	-0.203	0.044	<0.001	**
Ambito lavoro	Agricoltura	-0.188	0.069	0.007	**
	Costruzioni	-0.720	0.058	<0.001	**
	Commercio	-0.275	0.074	<0.001	**
	Ricettive	-0.630	0.061	<0.001	**
	Logistica	0.352	0.054	<0.001	**
	Servizi Pubblici	-0.555	0.078	<0.001	**
	Imprenditoria	-0.980	0.077	<0.001	**
	Istruzione/Sanità	-0.642	0.065	<0.001	**
Altri servizi alla persona	-0.638	0.058	<0.001	**	
Tipo contratto	Determinato	1.740	0.035	<0.001	**
	Part-time	0.214	0.033	<0.001	**
Esperienza	<2 anni	0.608	0.052	<0.001	**
	3-5 anni	0.134	0.056	0.017	*
	11-20 anni	-0.330	0.062	<0.001	**
	>20 anni	-0.159	0.063	0.012	*
Componenti nucleo	Uno	-0.142	0.052	0.007	**
	Due	-0.023	0.041	0.580	
	Quattro	-0.038	0.037	0.304	
	Cinque o più	0.025	0.052	0.631	
Caratteristiche nucleo	Minori 15 anni	0.002	0.037	0.949	
	Maggiori 64 anni	0.141	0.066	0.033	*
	Altro occupato	-0.018	0.030	0.546	
Cittadinanza	Straniero	-0.098	0.042	0.020	*
Sesso	Femmina	0.227	0.032	<0.001	**
Anno	2020	0.251	0.028	<0.001	**
	Costante	-3.890	0.080	<0.001	**

L'istruzione è associata positivamente alla probabilità di perdere il lavoro. In particolare, rispetto all'individuo di riferimento che possiede un titolo di studio medio, chi possiede un titolo di studio basso o nessun titolo ha una maggior probabilità di perdere il lavoro, pari al 2.5%, mentre chi ha un titolo di studio alto ha una probabilità più bassa e pari all'1.64%. Si evince quindi come il miglioramento

dell'istruzione comporti anche un miglioramento rispetto alla condizione lavorativa, con la probabilità di perdere il lavoro che diminuisce.

Il settore lavorativo di impiego, se diverso dall'industria, risulta in una diminuzione della probabilità di perdita del lavoro compresa tra gli 0.33 e gli 1.24 punti percentuali. Fa eccezione il settore della logistica, con gli impiegati in tale settore che registrano una probabilità maggiore rispetto agli impiegati nell'industria.

Come per l'istruzione, anche guardando all'esperienza lavorativa si osserva un effetto che porta a diminuire la probabilità di perdere il lavoro all'aumentare degli anni trascorsi presso il medesimo luogo di lavoro.

Infine, osserviamo come sono associate alla perdita del lavoro tre variabili molto importanti per le analisi: il genere, la cittadinanza e il periodo di osservazione. Le femmine con le medesime caratteristiche dei maschi risultano essere svantaggiate poiché vedono la propria probabilità di perdere il lavoro aumentare fino al 2.5%. Gli stranieri con caratteristiche simili agli italiani di riferimento sembrano invece avere un outcome migliore, con una probabilità pari all'1.82%, ma bisogna ricordare che gli stranieri con caratteristiche simili a quelle dell'individuo di riferimento sono una piccolissima minoranza. La variabile di maggior interesse, quella relativa al periodo di osservazione, conferma quanto si era intuito con le analisi descrittive ed in particolare la sua associazione positiva con la probabilità di perdere il lavoro a tre mesi. Gli individui seguiti durante la pandemia che differivano dall'individuo di riferimento solamente per questa caratteristica hanno infatti avuto una probabilità più alta di 0.56 punti percentuali (pari a 2.56%).

4.3 Modello logistico con interazioni con genere e cittadinanza

Per esplorare maggiormente le determinanti della perdita del lavoro, si è deciso di adattare anche un modello con interazioni. Inizialmente si era stimato il modello con tutte le interazioni tra genere e le altre variabili, cittadinanza e le altre variabili e genere e cittadinanza con le altre variabili. La quasi totalità dei coefficienti di interazione risultava non essere significativa e perciò si è reso necessario operare una ricerca di specificazione del modello provando ad inserire i termini di interazione non tutti insieme, ma un po' alla volta. Sono risultate essere significative le interazioni tra il genere e l'età, il genere e la presenza di minori di 15 anni, il

genere e l'orario di lavoro, la cittadinanza e l'età, la cittadinanza e il tipo di contratto e la cittadinanza e l'esperienza presso il posto di lavoro in cui si è occupati alla prima rilevazione.

Nella tabella seguente sono riportate le stime dei coefficienti del modello che verranno successivamente commentate.

Tabella 21: Stime del modello logistico con interazioni per genere e cittadinanza (** p-value<0.01, * p-value<0.05).

VARIABILE	LIVELLO	COEF	STD.ERROR	P-VALUE	SIGNIF.
Età	Giovani	0.154	0.053	0.004	**
	Over54	0.792	0.058	<0.001	**
Residenza	Centro	0.051	0.038	0.184	**
	Sud	0.441	0.032	<0.001	**
Titolo di studio	Medie o meno	0.233	0.032	<0.001	**
	Laurea o più	-0.217	0.044	<0.001	**
Ambito lavoro	Agricoltura	-0.136	0.070	0.054	
	Costruzioni	-0.684	0.059	<0.001	**
	Commercio	-0.219	0.075	0.004	**
	Ricettive	-0.593	0.062	<0.001	**
	Logistica	0.383	0.055	<0.001	**
	Servizi Pubblici	-0.525	0.079	<0.001	**
	Imprenditoria	-0.964	0.078	<0.001	**
	Istruzione/Sanità	-0.614	0.065	<0.001	**
Altri servizi alla persona	-0.614	0.060	<0.001	**	
Tipo contratto	Determinato	1.833	0.039	<0.001	**
	Part-time	0.343	0.053	<0.001	**
Esperienza	<2 anni	0.536	0.058	<0.001	**
	3-5 anni	0.109	0.063	0.086	
	11-20 anni	-0.331	0.068	<0.001	**
	>20 anni	-0.171	0.067	0.011	*
Componenti nucleo	Uno	-0.131	0.053	0.013	*
	Due	-0.014	0.041	0.738	
	Quattro	-0.031	0.037	0.395	
	Cinque o più	0.038	0.052	0.466	
Caratteristiche nucleo	Minori 15 anni	-0.249	0.053	<0.001	**
	Maggiori 64 anni	0.144	0.066	0.030	*
	Altro occupato	-0.039	0.030	0.200	
Cittadinanza	Straniero	0.155	0.120	0.197	
Sesso	Femmina	0.232	0.057	<0.001	**
Anno	2020	0.250	0.028	<0.001	**
Interazioni	Femmina, Giovane	-0.033	0.065	0.615	
	Femmina, Over54	-0.307	0.078	<0.001	**
	Femmina, Part-time	-0.233	0.065	<0.001	**
	Femmina, Minori 15 anni	0.479	0.065	<0.001	**
	Straniero, Giovane	-0.213	0.089	0.017	*
	Straniero, Over54	-0.494	0.122	<0.001	**
	Straniero, Determinato	-0.596	0.087	<0.001	**
	Straniero, <2 anni	0.385	0.130	0.003	**
	Straniero, 3-5 anni	0.097	0.141	0.489	
	Straniero, 11-20 anni	0.076	0.173	0.659	
	Straniero, >20 anni	-0.002	0.384	0.995	
	Costante		-3.944	0.089	<0.001

Analizzando le stime dei coefficienti del modello, si vede che permangono le significatività riscontrate nel modello senza interazione con i medesimi segni ed intensità simili.

La presenza di minori di 15 anni all'interno del nucleo familiare, non significativa nel modello senza interazioni, diventa ora significativa ed assume una dinamica opposta tra maschi e femmine. Se infatti per i maschi che differiscono dall'individuo di riferimento per tale caratteristica comporta una diminuzione della probabilità che la porta ad essere pari ad 1.42%, contro l'1.81% dell'individuo baseline, per le femmine la presenza di minori di 15 anni nel nucleo familiare comporta un aumento della probabilità di perdere il lavoro, pari al 2.84%.

Risulta significativa anche l'interazione tra il genere e l'età, con le femmine che registrano una diminuzione della probabilità di perdere il lavoro quando sono più vecchie rispetto alla baseline, mentre i maschi registravano un aumento. Le femmine ultracinquantenni con tutte le altre caratteristiche pari ai valori di riferimento hanno una probabilità di perdere il lavoro pari al 3.64%, mentre i maschi hanno una probabilità pari al 3.91%.

Anche l'orario di lavoro ha una dinamica diversa per maschi e femmine, con i primi che erano colpiti negativamente dal lavorare part-time, con una probabilità che aumentava rispetto al full-time, mentre per le femmine il part-time porta ad una diminuzione della probabilità di perdere il lavoro che risulta essere la medesima dei maschi, 2.53%.

L'essere stranieri non risulta più in un aumento significativo della probabilità di perdere il lavoro rispetto all'individuo di riferimento, mentre sono significativi alcuni termini di interazione tra la cittadinanza e alcune caratteristiche del lavoro. In particolare, al confronto con gli italiani con le stesse caratteristiche, gli stranieri che hanno un contratto di lavoro a tempo determinato sono meno svantaggiati dei maschi con lavoro a tempo determinato. I primi registrano una probabilità di perdere il lavoro pari al 6.90%, mentre i secondi pari al 10.33%.

L'esperienza lavorativa invece è un fattore che penalizza maggiormente gli stranieri rispetto agli italiani, con gli stranieri con meno di 2 anni di esperienza che hanno una maggior probabilità di perdere il lavoro rispetto agli italiani con meno di 2 anni di esperienza, stanti i valori delle altre variabili pari ai livelli di riferimento.

4.4 Modello logistico con interazioni con genere, cittadinanza e periodo

In ultima battuta, nel modello vengono inseriti anche i termini di interazione con il periodo, vero fattore di interesse delle analisi per capire se e come la pandemia abbia influito sulla perdita del lavoro.

Anche in questo caso, come nel precedente, inizialmente erano state inserite tutte le interazioni con il periodo, ma vista la non significatività della maggior parte dei coefficienti, sono state inserite delle interazioni mirate di interesse. Perciò, dopo aver mantenuto le interazioni significative al passaggio precedente, sono state inserite nel modello le interazioni tra periodo e genere, periodo e cittadinanza, periodo e tipo di contratto.

In tabella 22 vengono riportate le stime dei coefficienti del modello dal quale è possibile osservare come permangano le osservazioni fatte ai due paragrafi precedenti, con l'anno che però, in questo caso, non ha più un effetto marginalmente sulla probabilità di perdere il lavoro, ma lo ha in relazione alle altre variabili poste in interazione.

In particolare, gli stranieri durante la pandemia hanno avuto una probabilità significativamente maggiore rispetto agli italiani nello stesso periodo con le medesime caratteristiche pari al livello di riferimento. Per gli stranieri la probabilità è stata pari al 2.87%, contro il 2.22% degli italiani.

Anche le femmine durante la pandemia hanno registrato delle probabilità più alte di perdere il lavoro, sia rispetto alle femmine prima della pandemia, che rispetto ai maschi. Infatti, per le femmine durante la pandemia la probabilità si attestava al 2.93%, contro il 2.49% del periodo prima della pandemia e il 2.22% dei maschi, sempre considerando i valori delle altre variabili pari al livello dell'individuo di riferimento.

Infine, il tipo di contratto. Se già prima della pandemia avere un contratto a tempo determinato portava ad avere una probabilità superiore al 10% di perdere il lavoro a tre mesi rispetto all'individuo di riferimento con contratto a tempo indeterminato, durante la pandemia il peso di un contratto a tempo determinato è stato ancora più importante. Quanti durante la pandemia possedevano un contratto a tempo determinato, stanti i valori delle altre variabili ai livelli di riferimento, hanno infatti visto aumentare la propria probabilità di perdere il lavoro fino al 13.39%.

Tabella 22: Stime del modello logistico con interazioni per genere, cittadinanza e periodo (** p-value<0.01, * p-value<0.05).

VARIABILE	LIVELLO	COEF	STD.ERROR	P-VALUE	SIGNIF.
Età	Giovani	0.152	0.053	0.004	**
	Over54	0.794	0.058	<0.001	**
Residenza	Centro	0.051	0.038	0.185	
	Sud	0.443	0.032	<0.001	**
Titolo di studio	Medie o meno	0.232	0.032	<0.001	**
	Laurea o più	-0.217	0.044	<0.001	**
Ambito lavoro	Agricoltura	-0.139	0.071	0.048	*
	Costruzioni	-0.684	0.059	<0.001	**
	Commercio	-0.221	0.075	0.003	**
	Ricettive	-0.595	0.062	<0.001	**
	Logistica	0.384	0.055	<0.001	**
	Servizi Pubblici	-0.528	0.079	<0.001	**
	Imprenditoria	-0.946	0.078	<0.001	**
	Istruzione/Sanità	-0.615	0.065	<0.001	**
	Altri servizi alla persona	-0.618	0.060	<0.001	**
Tipo contratto	Determinato	1.744	0.048	<0.001	**
	Part-time	0.343	0.053	<0.001	**
Esperienza	<2 anni	0.527	0.058	<0.001	**
	3-5 anni	0.109	0.063	0.086	
	11-20 anni	-0.330	0.068	<0.001	**
	>20 anni	-0.169	0.067	0.011	*
Componenti nucleo	Uno	-0.129	0.053	0.014	*
	Due	-0.014	0.041	0.729	
	Quattro	-0.031	0.037	0.391	
	Cinque o più	0.039	0.051	0.458	
Caratteristiche nucleo	Minori 15 anni	-0.251	0.053	<0.001	**
	Maggiori 64 anni	0.146	0.066	0.027	*
	Altro occupato	-0.039	0.030	0.206	
Cittadinanza	Straniero	0.040	0.126	0.749	
Sesso	Femmina	0.178	0.063	0.005	**
Anno	2020	0.063	0.053	0.231	
Interazioni	Femmina, Giovane	-0.031	0.065	0.638	
	Femmina, Over54	-0.308	0.078	<0.001	**
	Femmina, Part-time	-0.233	0.065	<0.001	**
	Femmina, Minori 15 anni	0.480	0.065	<0.001	**
	Straniero, Giovane	-0.210	0.089	0.019	*
	Straniero, Over54	-0.497	0.122	<0.001	**
	Straniero, Tempo determinato	-0.591	0.087	<0.001	**
	Straniero, <2 anni	0.393	0.130	0.003	**
	Straniero, 3-5 anni	0.091	0.141	0.517	
	Straniero, 11-20 anni	0.065	0.173	0.707	
	Straniero, >20 anni	-0.007	0.384	0.986	
	2020, Tempo determinato	0.173	0.056	0.002	**
	2020, Straniero	0.224	0.077	0.004	**
	2020, Femmina	0.107	0.055	0.054	*
Costante		-3.847	0.092	<0.001	**

CAPITOLO 5

ANALISI CON PROPENSITY-SCORE MATCHING

Vista l'analisi dei dati con metodi di regressione, si vuole ora verificare se i risultati ottenuti per alcune covariate siano effettivamente dovuti all'impatto che queste variabili hanno avuto sul fenomeno o non insista anche un effetto legato alla composizione del campione.

Il modello logistico, infatti, come tutti i modelli di regressione, impone sui dati un modello parametrico, supponendo quindi di sapere quale sia la forma della relazione tra la variabile risposta e le variabili esplicative ed inoltre utilizza tutte le osservazioni a disposizione, senza curarsi del fatto che i soggetti possano essere molto diversi tra loro.

In questo capitolo si vuole indagare quali siano i tassi di transizione alla perdita del lavoro in tre diversi casi e se all'interno di questi casi le differenze siano significative: nel gruppo di quanti sono stati seguiti prima della pandemia e durante la pandemia, nel gruppo dei cittadini italiani e dei cittadini stranieri ed infine nel gruppo dei residenti al nord e dei residenti al centro-sud.

In tutti e tre i casi, si può pensare alla variabile che divide i gruppi (periodo, cittadinanza e ripartizione di residenza rispettivamente) come ad un trattamento dicotomico, anche se impropriamente detto perché non si è in presenza di uno studio randomizzato in cui è stato il ricercatore a scegliere il trattamento e ad assegnare gli individui ai gruppi dei trattati e dei controlli.

Per mezzo dell'abbinamento statistico (*matching*) è però possibile creare un gruppo di controllo "a posteriori" composto da soggetti non trattati che siano i più simili possibile ai trattati.

Questo metodo permette quindi di ricreare un ambiente di tipo sperimentale.

5.1 Il propensity-score matching

Se l'obiettivo è quello di conoscere la differenza di risultato tra chi ha subito il trattamento e chi no, indicato con Y_i il risultato e con D_i il trattamento, un modo per farlo è lavorare sulle medie

$$E[Y_i|D_i = 1] - E[Y_i|D_i = 0]$$

Il problema di fondo sta nel fatto che in uno stesso individuo si può osservare solamente uno dei due risultati, perché una volta assegnato o meno l'individuo al trattamento, non è più possibile conoscere quale sarebbe stato il risultato nell'altro caso, detto controfattuale. Si ragiona perciò in termini di risultati potenziali, risultati che si ottengono a seconda dell'assegnazione o meno al trattamento. Quanto scritto prima assume allora la forma seguente:

$$E[Y_{1i}|D_i = 1] - E[Y_{0i}|D_i = 0]$$

dove il primo addendo è il valore atteso del risultato di aver subito il trattamento dato che si è stati assegnati ad esso, mentre il secondo addendo è il valore atteso del risultato del non aver subito il trattamento dato che non si è stati assegnati al trattamento, entrambe quantità osservabili.

Con una manipolazione matematica che va ad aggiungere e togliere il controfattuale $E[Y_{0i}|D_i = 1]$, cioè il valore atteso del risultato se non avessero ricevuto il trattamento quanti lo hanno invece ricevuto, risultato non osservabile, si giunge alla situazione iniziale del problema:

$$\begin{aligned} E[Y_i|D_i = 1] - E[Y_i|D_i = 0] &= \\ &= E[Y_{1i}|D_i = 1] - E[Y_{0i}|D_i = 1] + E[Y_{0i}|D_i = 1] - E[Y_{0i}|D_i = 0] \end{aligned}$$

$$E[Y_i|D_i = 1] - E[Y_i|D_i = 0] = E[Y_{1i} - Y_{0i}|D_i = 1] + E[Y_{0i}|D_i = 1] - E[Y_{0i}|D_i = 0]$$

dove il primo addendo al secondo membro è definito *Average Treatment effect on the Treated* (ATT, effetto medio del trattamento sui trattati) e la differenza di

secondo e terzo addendo è il *Selection Bias*, la distorsione dovuta alle diverse condizioni di partenza di trattati e controlli. Se il selection bias fosse nullo, e quindi i soggetti trattati e non trattati in partenza fossero simili, la differenza osservata nel gruppo dei trattati e dei non trattati andrebbe a coincidere con l'effetto che si vorrebbe stimare.

L'utilizzo dell'abbinamento si basa su quella che è definita *Conditional Independence Assumption* (CIA), assunto secondo il quale il selection bias è eliminabile se ci si condiziona su tutte le variabili osservabili. A parità di queste ultime, si può assumere che l'assegnazione al trattamento sia avvenuta in modo casuale (Martini & Sisti, 2009).

L'abbinamento richiede perciò di definire quali unità sono simili fra loro e quali non lo sono. Il problema nel fare questo è il numero eccessivo di caratteristiche da tenere in considerazione per stabilire la somiglianza, cosa che lo rende spesso praticamente impossibile.

Per ovviare a tale problema, Rosenbaum e Rubin hanno introdotto l'uso del *propensity score*. Il teorema del propensity score afferma quanto segue. Se si assume che i risultati potenziali siano indipendenti dall'assegnazione al trattamento condizionatamente ad un vettore multivariato di covariate X_i , allora i risultati potenziali sono indipendenti dall'assegnazione al trattamento anche condizionatamente ad una funzione delle covariate, che prende appunto il nome di propensity score.

$$E[Y_{ij}|P(D_i = 1|X_i), D_i = 1] = E[Y_{ij}|P(D_i = 1|X_i), D_i = 0] = E[Y_{ij}|P(D_i = 1|X_i)]$$

Il propensity score è una probabilità, la probabilità di un'unità di essere assegnata al trattamento date le sue caratteristiche osservabili prima del trattamento.

$$P(D_i = 1|X_i) = E[D_i|X_i]$$

Laddove trattati e non trattati hanno un supporto comune, il propensity score racchiude tutta l'informazione rilevante per applicare la CIA (Angrist & Pischke, 2008).

Per costruire il propensity score, il primo passo è utilizzare tutte le unità (trattate e non trattate) per stimare un modello probabilistico che metta in relazione il trattamento D, in questo caso variabile dipendente, con le variabili osservabili che si pensa possano influenzare l'assegnazione al trattamento, l'insieme delle X. Il modello usato è un modello di regressione per dati binari, solitamente con funzione di legame logit o, in alcuni casi in alternativa, probit.

Per operare l'abbinamento delle unità trattate e non trattate, l'idea di base è quella di abbinare a ciascun trattato il non trattato con il propensity score più simile. Per definire quanto simile debba essere, esistono diverse metodologie:

- *Nearest neighbor*: è il metodo più semplice e intuitivo. Consiste nell'abbinare ad ogni unità trattata quell'unità non trattata che ha il valore di propensity score numericamente più vicino. La selezione può avvenire sia con reimmissione dell'unità non trattata, che quindi potrebbe essere abbinata a più trattati, sia senza;
- *Caliper*: funziona in modo simile al nearest neighbor, ma restringe l'abbinamento alle unità non trattate il cui propensity score cade in un intorno ristretto del propensity score dell'unità trattata, di raggio λ . A differenza del nearest neighbor, con questo metodo possono essere scartate anche delle unità trattate qualora non si trovasse nessuna unità non trattata con propensity score che cada nell'intervallo $P \pm \lambda$;
- *Kernel*: ad ogni unità trattata sono abbinate tutte le unità non trattate, pesate in modo inversamente proporzionale alla distanza del loro propensity score da quello dell'unità trattata. Si chiama Kernel poiché per pesare in modo proporzionale alla distanza viene usata una funzione nucleo. Tale abbinamento impone la scelta di pesi tali che i loro valori decrescano tanto più è grande la distanza del propensity score da quello dell'unità trattata. Ha il vantaggio di usare tutta l'informazione disponibile.

Per valutare poi la qualità dell'abbinamento si va a verificare il miglioramento del bilanciamento della distribuzione delle variabili utilizzate per calcolare il propensity score, eventualmente anche utilizzando dei t-test sulle differenze.

5.2 Analisi della transizione lavorativa su soggetti comparabili

Le variabili utilizzate per calcolare il propensity-score sono state le medesime già utilizzate nei modelli logistici, con l'aggiunta di qualche altra variabile quale lo stato civile (celibi/nubili, sposati/conviventi, divorziati/vedovi), la presenza di diplomati, la presenza di laureati e la presenza di disoccupati nel nucleo familiare e la situazione lavorativa dichiarata dall'individuo relativa all'anno precedente (occupato, disoccupato, inattivo o altro).

I risultati che vengono commentati sono relativi al matching eseguito con metodo *nearest neighbor*. Gli altri metodi di matching citati al paragrafo precedente sono stati utilizzati per fare un'analisi di robustezza dei risultati, riportata in un capitolo a parte.

È stata valutato in modo preponderante la transizione alla perdita del lavoro in generale. L'analisi delle singole transizioni alla disoccupazione o all'inattività è stata eseguita come analisi di robustezza ed è riportata nel medesimo capitolo in cui si trattano gli altri metodi di matching.

Tutte le analisi si sono svolte separatamente per maschi e femmine, vista la diversa dinamica emersa dalle analisi descrittive disaggregate per genere e confermate dalla significatività dei coefficienti per il genere nei modelli logistici.

In particolare, poi, per cittadinanza e ripartizione geografica, le analisi sono state differenziate anche per periodo.

In tutti e tre i casi trattati, a seguito del matching il bilanciamento dei due gruppi risultava essere migliore rispetto a quello di partenza.

5.2.1 Periodo come trattamento

Il campione di partenza risultava essere già abbastanza bilanciato (si vedano le tabelle con le distribuzioni delle variabili per l'abbinamento in appendice), cosa che ci si aspettava dal momento che un drastico cambiamento della composizione della popolazione da un anno all'altro è improbabile.

In tabella 23 si possono vedere le stime ottenute prima e dopo il matching per maschi e femmine rispetto alla transizione dallo stato di occupazione.

Tabella 23: Tassi di transizione alla perdita del lavoro per genere e periodo con e senza appaiamento.

Genere	Campione	Durante pandemia	Pre-pandemia	Differenza	p-value
Maschi	Tutti gli occupati	4.52	4.02	0.50	<0.001
	Appaiati	4.52	3.70	0.82	<0.001
	Num. Occupati	32,823	36,072		
	Num. Occupati appaiati	32,823	32,823		
Femmine	Tutti gli occupati	5.90	4.94	0.96	<0.001
	Appaiati	5.90	4.54	1.36	<0.001
	Num. Occupate	29,470	32,332		
	Num. Occupate appaiate	29,407	29,470		

Andando a vedere cosa succede alle probabilità di transizione, si nota come sia per i maschi sia per le femmine, da una iniziale situazione in cui si hanno probabilità più alte di perdere il lavoro per gli individui durante la pandemia, a seguito dell'abbinamento tali differenze diventino ancora più marcate. Tale risultato conforta sul fatto che l'innalzamento della probabilità di perdere il lavoro a tre mesi durante la pandemia, piuttosto che prima, non sia dovuto ad un effetto di composizione del campione.

In particolare, poi, confrontando maschi e femmine tra prima e durante la pandemia si nota come l'aumento sia stato più cospicuo per le seconde rispetto ai primi.

5.2.2 Cittadinanza come trattamento

Osservando la distribuzione delle caratteristiche di maschi e femmine italiani e stranieri nei due anni (tabelle in appendice), si nota come vi sia un notevole sbilanciamento di partenza. Gli stranieri del campione risultano generalmente essere più giovani e meno istruiti, risultano risiedere maggiormente al Nord Italia in famiglie numerose o da soli. Hanno più spesso minori di 15 anni nel nucleo familiare ed è più raro che vi siano altri occupati oltre all'intervistato. Lavorano maggiormente con contratti a tempo determinato e part-time, soprattutto le femmine, ed anche i settori di impiego sono differenti: i maschi stranieri sono più spesso impiegati in agricoltura, industria e costruzioni, mentre le femmine straniere in più della metà dei casi sono impiegate nei servizi alla persona, con una grande differenza in tale settore rispetto alle femmine italiane.

In tabella 24 si possono apprezzare le variazioni delle stime dei tassi di transizione alla perdita del lavoro per maschi e femmine nei due periodi con un confronto rispetto alla cittadinanza.

Tabella 24: Tassi di transizione alla perdita del lavoro per periodo, genere e cittadinanza con e senza appaiamento.

Periodo	Genere	Campione	Stranieri	Italiani	Differenza	p-value
Prima Pandemia	Maschi	Tutti gli occupati	4.51	3.96	0.55	0.104
		Appaiati	4.51	5.67	-1.16	0.015
		Num. Occupati	4,215	31,857		
		Num. Occupati appaiati	4,215	4,215		
	Femmine	Tutti gli occupati	6.30	4.75	1.55	<0.001
		Appaiati	6.30	8.28	-1.98	0.001
		Num. Occupate	3,780	28,552		
		Num. Occupate appaiate	3,780	3,780		
Durante pandemia	Maschi	Tutti gli occupati	5.81	4.36	1.45	<0.001
		Appaiati	5.81	6.36	-0.55	0.326
		Num. Occupati	3,630	29,193		
		Num. Occupati appaiati	3,630	3,630		
	Femmine	Tutti gli occupati	9.19	5.51	3.68	<0.001
		Appaiati	9.19	10.08	-0.89	0.233
		Num. Occupate	3,166	26,304		
		Num. Occupate appaiate	3,166	3,166		

Dal matching per i maschi si può notare come una iniziale situazione che vede gli stranieri svantaggiati rispetto agli italiani, con quote più alte degli italiani e differenze significative tra le due, che aumentano nel periodo della pandemia fino a quasi triplicare, a seguito dell'abbinamento in realtà non sembrano soffrire di questo svantaggio. Le differenze, infatti, cambiano addirittura di segno, ad indicare che sarebbero invece gli italiani a soffrire di più a parità di condizioni, ma non risultano essere significativamente diverse da zero.

Si può quindi asserire che, confrontando occupati italiani e stranieri con caratteristiche comparabili, non vi siano differenze significative rispetto alla probabilità di perdere il lavoro e che le differenze iniziali siano quindi dovute molto ad un effetto di composizione dei due gruppi.

Guardando alle femmine, il discorso può essere il medesimo rispetto a quello fatto per i maschi, notando però come in questo caso la situazione di differenza iniziale sia molto più marcata e si aggravi particolarmente durante la pandemia. Tali differenze tra italiane e straniere sono infatti circa il triplo di quelle che si osservano

tra i maschi (1.55% vs 0.55% prima della pandemia, 3.68% vs. 1.45% durante la pandemia).

5.2.3 Ripartizione geografica come trattamento

Sono note le differenze di dinamiche lavorative che esistono tra il Centro-Nord ed il Sud Italia. In particolare, guardando alla distribuzione delle caratteristiche degli individui che abitano nelle due zone (tabelle in appendice) si può notare come i residenti al Centro-Nord in generale siano lievemente più giovani e più istruiti. Inoltre, si osserva come siano maggiormente celibi/nubili e vivono quindi in nuclei familiari con meno componenti, dove sono presenti quote minori di disoccupati e una maggior presenza di un altro o più occupati. Al Centro-Nord risultano risiedere anche più stranieri rispetto al Sud. Guardando al lavoro, i residenti al Centro-Nord sono maggiormente impiegati con contratti a tempo indeterminato e full-time e lavorano maggiormente in settori quali l'industria, mentre i residenti al Sud sono solitamente più impiegati in settori come l'agricoltura, le strutture ricettive ed i servizi pubblici.

Tabella 25: Tassi di transizione alla perdita del lavoro per periodo, genere e residenza con e senza appaiamento.

Periodo	Genere	Campione	Sud	Nord	Differenza	p-value
Prima Pandemia	Maschi	Tutti gli occupati	6.15	3.17	2.98	<0.001
		Appaiati	6.15	4.07	2.08	<0.001
		Num. Occupati	10,325	25,747		
		Num. Occupati appaiati	10,325	10,325		
	Femmine	Tutti gli occupati	7.39	4.17	3.22	<0.001
		Appaiati	7.39	4.69	2.70	<0.001
		Num. Occupate	7,651	24,681		
		Num. Occupate appaiate	7,651	7,651		
Durante pandemia	Maschi	Tutti gli occupati	6.58	3.70	2.88	<0.001
		Appaiati	6.58	4.64	1.94	<0.001
		Num. Occupati	9,368	23,455		
		Num. Occupati appaiati	9,368	9,368		
	Femmine	Tutti gli occupati	8.30	5.18	3.12	<0.001
		Appaiati	8.30	5.70	2.60	<0.001
		Num. Occupate	6,843	22,627		
		Num. Occupate appaiate	6,843	6,843		

Per i maschi seguiti prima della pandemia, risulta una situazione di partenza per la quale i residenti al Centro-Nord hanno tassi più bassi dei residenti al Sud rispetto

alla perdita del lavoro. A seguito dell'appaiamento, tali differenze diminuiscono perché aumentano i tassi per i residenti al Centro-Nord con caratteristiche simili ai residenti al Sud e tali differenze rimangono significative, ad indicare sì un effetto della differente composizione del campione, ma non completamente dovuta a questo.

Analizzando i maschi durante la pandemia, si assiste alla medesima dinamica a cui si assisteva prima della pandemia, con le differenze che risultano ancora statisticamente significative.

Guardando alle femmine prima della pandemia, si possono trarre le stesse conclusioni che sono state tratte per i maschi. In particolare, si hanno situazioni di partenza ancora più svantaggiose sia per le residenti al Centro-Nord che per le residenti al Sud, ad indicare un maggior propensione delle femmine a perdere il lavoro.

Andando a guardare quanto succede nel periodo durante la pandemia, la situazione di differenza tra residenti al Centro-Nord e al Sud sembra migliorare per le seconde, con differenze più contenute, ma comunque significative.

CAPITOLO 6

ULTERIORI EVIDENZE DAI DATI

Dal momento che esistono diversi metodi di matching sul propensity score e che con i dati a disposizione era possibile non solo valutare la transizione alla perdita del lavoro in generale, ma anche separatamente le transizioni verso la disoccupazione e verso l'inattività, tali analisi sono state svolte in aggiunta a quanto discusso nei risultati, in modo da fare un'analisi di robustezza. Sempre con i dati a disposizione era inoltre possibile esplorare delle diverse variabili come risposta al fenomeno oggetto di studio, così come esposto in questo capitolo su una possibile variabile alternativa per dare uno spunto su futuri ampliamenti del lavoro.

6.1 Applicazione di diversi metodi di matching su differenti outcome

Per tutte e tre le variabili considerate come trattamento, sono stati utilizzati altri tre metodi di matching: nearest neighbor con caliper, stima del propensity-score con modello probit e matching nearest neighbor con caliper, kernel matching.

Vengono riportate di seguito le tabelle con i risultati debitamente commentate, ad iniziare dall'uso del periodo come trattamento.

Tabella 26: Tassi di transizione verso i diversi tipi di perdita del lavoro per periodo. Tutti i maschi occupati e maschi occupati appaiati.

Transizione	Campione	Durante pandemia	Pre-pandemia	Differenza	p-value
Disoccupazione	Tutti gli occupati	1.55	1.49	0.06	0.503
	Nearest-neighbor	1.55	1.33	0.22	0.018
	NN caliper	1.56	1.33	0.23	0.016
	NN caliper probit	1.56	1.37	0.19	0.043
	Kernel	1.55	1.35	0.20	0.031
Inattività	Tutti gli occupati	2.97	2.53	0.44	<0.001
	Nearest-neighbor	2.97	2.37	0.60	<0.001
	NN caliper	2.97	2.38	0.59	<0.001
	NN caliper probit	2.97	2.42	0.55	<0.001
	Kernel	2.97	2.41	0.56	<0.001
Perdita del lavoro	Tutti gli occupati	4.52	4.02	0.50	0.001
	NN caliper	4.53	3.71	0.82	<0.001
	NN caliper probit	4.53	3.79	0.74	<0.001
	Kernel	4.52	3.75	0.77	<0.001
Numerosità	Tutti gli occupati	32,823	36,072		
	Nearest-neighbor	32,823	32,823		
	NN caliper	32,548	32,548		
	NN caliper probit	32,552	32,552		

Guardando la perdita del lavoro in generale, i diversi metodi di matching confermano quanto si è ottenuto tramite il nearest neighbor. Analizzando le transizioni separate a disoccupazione e inattività si nota come la significatività delle differenze relative alla perdita del lavoro in generale siano dovute in larga parte alla transizione verso l'inattività, con differenze sempre significative a sfavore di quanti sono stati intervistati durante la pandemia, mentre la transizione alla disoccupazione presenti differenze più contenute e significative ad un livello più basso.

Tabella 27: Tassi di transizione verso i diversi tipi di perdita del lavoro per periodo. Tutte le femmine occupate e femmine occupate appaiate.

Transizione	Campione	Durante pandemia	Pre-pandemia	Differenza	p-value
Disoccupazione	Tutti gli occupati	1.53	1.55	-0.02	0.873
	Nearest-neighbor	1.53	1.40	0.13	0.171
	NN caliper	1.54	1.40	0.14	0.168
	NN caliper probit	1.54	1.39	0.15	0.129
	Kernel	1.53	1.39	0.14	0.150
Inattività	Tutti gli occupati	4.37	3.39	0.98	<0.001
	Nearest-neighbor	4.37	3.14	1.23	<0.001
	NN caliper	4.40	3.16	1.24	<0.001
	NN caliper probit	4.41	3.13	1.28	<0.001
	Kernel	4.37	3.15	1.22	<0.001
Perdita del lavoro	Tutti gli occupati	5.90	4.94	0.96	<0.001
	NN caliper	5.94	4.56	1.38	<0.001
	NN caliper probit	5.94	4.52	1.42	<0.001
	Kernel	5.90	4.54	1.36	<0.001
Numerosità	Tutti gli occupati	29,470	32,332		
	Nearest-neighbor	29,407	29,470		
	NN caliper	29,168	29,168		
	NN caliper probit	29,167	29,167		

Quanto detto precedentemente per i maschi può essere ripetuto per le femmine. In questo caso però le differenze relative alla transizione alla disoccupazione non risultano essere significative, dando ancora maggior peso alla transizione all'inattività sulla significatività delle differenze relative alla perdita del lavoro in generale.

Passando all'uso della cittadinanza come trattamento, gli altri metodi di matching utilizzati confermano la non significatività delle differenze durante la pandemia a seguito del matching, sia per i maschi che per le femmine. Si nota una maggior incertezza nel periodo prima della pandemia, con i metodi con caliper in disaccordo rispetto ai metodi kernel e nearest neighbor senza caliper.

Tabella 28: Tassi di transizione verso i diversi tipi di perdita del lavoro per cittadinanza prima della pandemia. Tutti i maschi occupati e maschi occupati appaiati.

Transizione	Campione	Stranieri	Italiani	Differenza	p-value
Disoccupazione	Tutti gli occupati	2.54	1.35	1.19	<0.001
	Nearest-neighbor	2.54	2.21	0.33	0.316
	NN caliper	2.51	2.00	0.51	0.159
	NN caliper probit	2.58	1.98	0.60	0.101
	Kernel	2.54	2.17	0.37	0.168
Inattività	Tutti gli occupati	1.97	2.61	-0.64	0.006
	Nearest-neighbor	1.97	3.46	-1.49	<0.001
	NN caliper	2.00	2.94	-0.94	0.014
	NN caliper probit	1.98	2.82	-0.84	0.025
	Kernel	1.97	3.60	-1.63	<0.001
Perdita del lavoro	Tutti gli occupati	4.51	3.96	0.55	0.104
	NN caliper	4.51	4.94	-0.43	0.417
	NN caliper probit	4.56	4.80	-0.24	0.643
	Kernel	4.51	5.76	-1.25	0.001
Numerosità	Tutti gli occupati	4,215	31,857		
	Nearest-neighbor	4,215	4,215		
	NN caliper	3,298	3,298		
	NN caliper probit	3,335	3,335		

Tabella 29: Tassi di transizione verso i diversi tipi di perdita del lavoro per cittadinanza prima della pandemia. Tutte le femmine occupate e femmine occupate appaiate.

Transizione	Campione	Stranieri	Italiani	Differenza	p-value
Disoccupazione	Tutti gli occupati	2.70	1.40	1.30	<0.001
	Nearest-neighbor	2.70	2.70	0.00	1.000
	NN caliper	2.84	2.68	0.16	0.730
	NN caliper probit	2.96	2.37	0.59	0.191
	Kernel	2.70	2.92	-0.22	0.490
Inattività	Tutti gli occupati	3.60	3.35	0.25	0.444
	Nearest-neighbor	3.60	5.58	-1.98	<0.001
	NN caliper	3.88	5.08	-1.20	0.040
	NN caliper probit	3.87	5.17	-1.30	0.026
	Kernel	3.60	5.37	-1.77	<0.001
Perdita del lavoro	Tutti gli occupati	6.30	4.75	1.55	<0.001
	NN caliper	6.72	7.76	-1.04	0.156
	NN caliper probit	6.83	7.54	-0.71	0.328
	Kernel	6.30	8.29	-1.99	<0.001
Numerosità	Tutti gli occupati	3,780	28,552		
	Nearest-neighbor	3,780	3,780		
	NN caliper	2,499	2,499		
	NN caliper probit	2,534	2,534		

Tabella 30: Tassi di transizione verso i diversi tipi di perdita del lavoro per cittadinanza durante la pandemia. Tutti i maschi occupati e maschi occupati appaiati.

Transizione	Campione	Stranieri	Italiani	Differenza	p-value
Disoccupazione	Tutti gli occupati	2.70	1.41	1.29	<0.001
	Nearest-neighbor	2.70	2.34	0.36	0.331
	NN caliper	2.37	1.75	0.62	0.109
	NN caliper probit	2.44	1.87	0.57	0.143
	Kernel	2.70	2.54	0.16	0.603
Inattività	Tutti gli occupati	3.11	2.95	0.16	0.599
	Nearest-neighbor	3.11	4.02	-0.91	0.04
	NN caliper	3.07	3.82	-0.75	0.124
	NN caliper probit	3.25	3.64	-0.39	0.423
	Kernel	3.11	4.01	-0.90	0.010
Perdita del lavoro	Tutti gli occupati	5.81	4.36	1.45	<0.001
	NN caliper	5.44	5.57	-0.13	0.815
	NN caliper probit	5.69	5.51	0.18	0.773
	Kernel	5.81	6.55	-0.74	0.105
Numerosità	Tutti gli occupati	3,630	29,193		
	Nearest-neighbor	3,630	3,630		
	NN caliper	2,799	2,799		
	NN caliper probit	2,831	2,831		

Tabella 31: Tassi di transizione verso i diversi tipi di perdita del lavoro per cittadinanza durante la pandemia. Tutte le femmine occupate e femmine occupate appaiate.

Transizione	Campione	Stranieri	Italiani	Differenza	p-value
Disoccupazione	Tutti gli occupati	2.65	1.40	1.25	<0.001
	Nearest-neighbor	2.65	3.51	-0.86	0.050
	NN caliper	2.80	3.52	-0.72	0.186
	NN caliper probit	2.81	3.57	-0.76	0.160
	Kernel	2.65	2.83	-0.18	0.603
Inattività	Tutti gli occupati	6.54	4.11	2.43	<0.001
	Nearest-neighbor	6.54	6.57	-0.03	0.956
	NN caliper	7.22	5.51	1.71	0.023
	NN caliper probit	6.95	6.67	0.28	0.713
	Kernel	6.54	5.78	0.76	0.168
Perdita del lavoro	Tutti gli occupati	9.19	5.51	3.68	<0.001
	NN caliper	10.03	9.03	1.00	0.270
	NN caliper probit	9.76	10.24	-0.48	0.607
	Kernel	9.19	8.61	0.58	0.363
Numerosità	Tutti gli occupati	3,166	26,304		
	Nearest-neighbor	3,166	3,166		
	NN caliper	2,104	2,104		
	NN caliper probit	2,100	2,100		

Anche in questo caso, così come osservato per il periodo come trattamento, si nota che la spinta maggiore alle significatività viene data dalle differenze rispetto alla transizione all'inattività, mentre nella totalità dei casi le differenze relative alle transizioni verso la disoccupazione non risultano essere mai significative a seguito del matching.

Infine, anche per quanto riguarda l'uso come trattamento della ripartizione geografica, i diversi metodi di matching utilizzati confermano i risultati ottenuti con il metodo proposto al capitolo precedente.

Tabella 32: Tassi di transizione verso i diversi tipi di perdita del lavoro per ripartizione di residenza prima della pandemia. Tutti i maschi occupati e maschi occupati appaiati.

Transizione	Campione	Sud	Nord	Differenza	p-value
Disoccupazione	Tutti gli occupati	2.16	1.22	0.94	<0.001
	Nearest-neighbor	2.16	1.50	0.66	<0.001
	NN caliper	2.00	1.49	0.51	0.007
	NN caliper probit	2.05	1.36	0.69	<0.001
	Kernel	2.16	1.66	0.50	0.003
Inattività	Tutti gli occupati	3.99	1.95	2.04	<0.001
	Nearest-neighbor	3.99	2.57	1.42	<0.001
	NN caliper	3.78	2.53	1.25	<0.001
	NN caliper probit	3.69	2.45	1.24	<0.001
	Kernel	3.99	2.75	1.24	<0.001
Perdita del lavoro	Tutti gli occupati	6.15	3.17	2.98	<0.001
	NN caliper	5.78	4.01	1.77	<0.001
	NN caliper probit	5.74	3.81	1.93	<0.001
	Kernel	6.15	4.41	1.74	<0.001
	Numerosità	Tutti gli occupati	10,325	25,747	
Nearest-neighbor		10,325	10,325		
NN caliper		9,494	9,494		
NN caliper probit		9,491	9,491		

Tabella 33: Tassi di transizione verso i diversi tipi di perdita del lavoro per ripartizione di residenza prima della pandemia. Tutte le femmine occupate e femmine occupate appaiate.

Transizione	Campione	Sud	Nord	Differenza	p-value
Disoccupazione	Tutti gli occupati	2.29	1.32	0.97	<0.001
	Nearest-neighbor	2.29	1.54	0.75	0.001
	NN caliper	2.20	1.45	0.75	<0.001
	NN caliper probit	2.19	1.37	0.82	<0.001
	Kernel	2.29	1.63	0.66	<0.001
Inattività	Tutti gli occupati	5.10	2.84	2.26	<0.001
	Nearest-neighbor	5.10	3.15	1.95	<0.001
	NN caliper	5.00	3.12	1.88	<0.001
	NN caliper probit	4.97	3.25	1.72	<0.001
	Kernel	5.10	3.54	1.56	<0.001
Perdita del lavoro	Tutti gli occupati	7.39	4.17	3.22	<0.001
	NN caliper	7.20	4.56	2.64	<0.001
	NN caliper probit	7.16	4.61	2.55	<0.001
	Kernel	7.39	5.16	2.23	<0.001
	Numerosità	Tutti gli occupati	7,651	24,681	
Nearest-neighbor		7,651	7,651		
NN caliper		7,540	7,540		
NN caliper probit		7,542	7,542		

Tabella 34: Tassi di transizione verso i diversi tipi di perdita del lavoro per ripartizione di residenza durante la pandemia. Tutti i maschi occupati e maschi occupati appaiati.

Transizione	Campione	Sud	Nord	Differenza	p-value
Disoccupazione	Tutti gli occupati	2.17	1.30	0.87	<0.001
	Nearest-neighbor	2.17	1.53	0.64	0.001
	NN caliper	2.05	1.52	0.53	0.008
	NN caliper probit	2.07	1.53	0.54	0.007
	Kernel	2.17	1.80	0.37	0.041
Inattività	Tutti gli occupati	4.41	2.40	2.01	<0.001
	Nearest-neighbor	4.41	3.11	1.30	<0.001
	NN caliper	4.17	3.02	1.15	<0.001
	NN caliper probit	4.25	2.98	1.27	<0.001
	Kernel	4.41	3.36	1.05	<0.001
Perdita del lavoro	Tutti gli occupati	6.58	3.70	2.88	<0.001
	NN caliper	6.22	4.54	1.68	<0.001
	NN caliper probit	6.32	4.51	1.81	<0.001
	Kernel	6.58	5.16	1.42	<0.001
Numerosità	Tutti gli occupati	9,368	23,455		
	Nearest-neighbor	9,368	9,368		
	NN caliper	8,685	8,685		
	NN caliper probit	8,652	8,652		

Tabella 35: Tassi di transizione verso i diversi tipi di perdita del lavoro per ripartizione di residenza durante la pandemia. Tutte le femmine occupate e femmine occupate appaiate.

Transizione	Campione	Sud	Nord	Differenza	p-value
Disoccupazione	Tutti gli occupati	1.90	1.42	0.48	0.005
	Nearest-neighbor	1.90	1.48	0.42	0.055
	NN caliper	1.89	1.47	0.42	0.052
	NN caliper probit	1.87	1.44	0.43	0.050
	Kernel	1.90	1.59	0.31	0.103
Inattività	Tutti gli occupati	6.40	3.76	2.64	<0.001
	Nearest-neighbor	6.40	4.22	2.18	<0.001
	NN caliper	6.26	4.10	2.16	<0.001
	NN caliper probit	6.26	4.18	2.08	<0.001
	Kernel	6.40	4.32	2.08	<0.001
Perdita del lavoro	Tutti gli occupati	8.30	5.18	3.12	<0.001
	NN caliper	8.15	5.57	2.58	<0.001
	NN caliper probit	8.13	5.62	2.51	<0.001
	Kernel	8.30	5.92	2.38	<0.001
Numerosità	Tutti gli occupati	6,843	22,627		
	Nearest-neighbor	6,843	6,843		
	NN caliper	6,757	6,757		
	NN caliper probit	6,744	6,744		

In questo caso, a differenza di quanto succedeva per gli altri due trattamenti considerati, si assiste ad una significatività delle differenze anche per quanto riguarda la transizione alla disoccupazione. Questo può forse suggerirci una ancor diversa dinamica della mobilità lavorativa tra nord e sud del paese.

6.2 Analisi su definizione alternativa degli stati lavorativi

Visti i decreti del Consiglio dei Ministri che andavano a limitare la possibilità di svolgere il proprio lavoro per la chiusura imposta a determinati tipi di lavori, nell'ottica di porre un freno alla catena del contagio, può essere interessante anche vedere i flussi di quanti degli intervistati durante la pandemia, ad un trimestre di distanza dalla prima rilevazione in cui dichiaravano di aver lavorato almeno un'ora, non lo avessero invece più fatto.

La variabile relativa all'aver lavorato almeno un'ora nella settimana di riferimento è infatti una delle variabili cardine per discriminare tra occupati e non occupati. Analizzando questi flussi si possono avere quindi ulteriori conferme di quanto ci si chiedeva inizialmente, cioè se la pandemia avesse avuto un effetto sulla perdita del lavoro degli occupati.

Tabella 36: Flussi trimestrali della risposta alla domanda “Ha lavorato almeno un'ora nella settimana di riferimento?”

Prima rilevazione	Seconda rilevazione			
	Lavorato almeno 1 ora	Non lavorato almeno 1 ora	Permanentemente inabile	Totale
Lavorato almeno 1 ora	57141 81.83	12675 18.15	16 0.02	69832 100.00
Non lavorato almeno 1 ora	9696 14.67	55766 84.35	650 0.98	66112 100.00
Permanentemente inabile	8 0.50	563 35.36	1021 64.13	1592 100.00
Totale	66845 48.60	69004 50.17	1687 1.23	137536 100.00

Da tabella 36 si può notare come il 18.15% di quanti alla prima rilevazione avevano lavorato almeno un'ora nella settimana di riferimento, ad un trimestre di distanza dicano di non averlo fatto. Di questi, risultano avere ancora un lavoro attivo alla seconda rilevazione il 76.63% dei casi e le cause maggiori a giustificare il non aver lavorato, escluse le ferie (motivazione addotta dal 33.90% degli individui), sono la Cassa Integrazione (ordinaria o straordinaria) nel 19.05% dei casi, la ridotta attività dell'impresa nel 9.63% dei casi e altri casi non specificati e non tra quelli codificati nel 20.14% dei casi.

Tali motivazioni possono sì dipendere da un'evoluzione fisiologica del mercato del lavoro, ma se confrontati con i valori del gruppo seguito prima della pandemia ci si accorge che i valori sono di molto più alti. Per questi individui infatti era pari all'11.66% la quota di quanti ad un trimestre di distanza avevano lavorato meno di un'ora nella settimana di riferimento e la maggior parte delle cause erano da imputare alle ferie, che incidono per oltre il 65% delle assenze mentre il ricorso alla CIG era la giustificazione di meno dell'1% delle assenze (0.92%, contro oltre il 19% nel periodo della pandemia) e le assenze relative ad una ridotta attività dell'impresa erano circa un terzo di quelle che si sono registrate nel periodo durante la pandemia (3.38% contro il 9.63%).

Questa ulteriore analisi, quindi, conforta anche sulla bontà degli altri risultati trovati nelle analisi eseguite e apre a diversi possibili scenari di analisi futuri che potrebbero essere indagati per scendere sempre di più nel dettaglio rispetto agli effetti della pandemia sul mercato del lavoro italiano.

CAPITOLO 7

CONCLUSIONI E POSSIBILI SVILUPPI

Come si è detto nell'introduzione a questo lavoro di tesi, l'analisi del mercato del lavoro in Italia è un tema di grande interesse a molti livelli dell'opinione pubblica, da quello macro-economico, a quello politico-sociale, fino al livello del singolo individuo.

Con questo elaborato si è voluto in particolare analizzare la dinamica della transizione lavoro-non lavoro alla luce del periodo straordinario che il paese, ed il mondo in generale, ha dovuto affrontare a partire dal marzo 2020. Il Covid-19 ha infatti colpito in modo duro la situazione lavorativa degli italiani ed in particolare dai rapporti annuali ISTAT 2019 e 2020 si apprendeva del crollo dei tassi di occupazione che fino a fine 2019 erano invece stati stabili o in aumento.

Con le analisi svolte si è quindi voluto indagare cosa avesse inciso sulla transizione verso la perdita del lavoro ed in particolare se ci fosse stata una particolare associazione con la cittadinanza o con la ripartizione geografica di residenza, anche confrontando tra loro individui con caratteristiche socio-demografiche e lavorative simili.

Dopo che nei capitoli 1 e 2 sono state delineate le fonti dei dati, come è stato eseguito l'abbinamento longitudinale degli individui, quali fossero gli individui e le variabili di interesse tra quelle a disposizione, a partire dal capitolo 3 si è entrati nelle vere e proprie analisi per rispondere alle domande di ricerca. Da una prima analisi sui tassi di transizione nei due periodi è emerso come la grande maggioranza degli occupati alla prima rilevazione considerata, lo restassero anche ad un trimestre di distanza. Durante la pandemia risultava però aumentare la percentuale di quanti transitavano all'inattività e diminuire quella di quanti trovavano un lavoro se inizialmente

disoccupati o inattivi. Disaggregando per genere, cittadinanza e ripartizione geografica emergevano delle interessanti differenze nei tassi di transizione. In particolare, le femmine risultavano essere maggiormente penalizzate dal Covid rispetto alla perdita del lavoro. Se infatti già prima della pandemia presentavano percentuali più alte dei maschi nella transizione verso l'inattività o la disoccupazione, durante la pandemia tale divario aumentava ulteriormente. Dinamica simile quella che si osservava per gli stranieri nei confronti degli autoctoni, con entrambi i gruppi che mostravano tassi di transizione all'inattività più alti rispetto a prima della pandemia. Infine, confrontando i residenti al Centro-Nord con quelli residenti al Sud risultava esserci un divario importante a svantaggio degli abitanti del Mezzogiorno sia per i tassi di transizione all'inattività che, in particolare, per i tassi di transizione verso l'occupazione se inizialmente disoccupati o inattivi.

Al capitolo 4 si è voluto analizzare per mezzo della regressione quali fossero le variabili associate alla perdita del lavoro, per avere un'indicazione di quali potessero essere i determinanti del fenomeno in esame. Stanti le caratteristiche dell'individuo di riferimento, così come riportate nel capitolo, la sua probabilità di perdere il lavoro era pari al 2%. Rispetto a tali caratteristiche baseline, risultavano agire in modo significativo aumentando la probabilità in particolare l'essere femmina piuttosto che maschio, il risiedere al Sud piuttosto che al Nord, l'aver un titolo di studio basso o nessun titolo piuttosto che un diploma di scuola superiore, l'essere giovani o avere più di 54 anni piuttosto che essere nella fascia 35-54 anni. Rispetto alle caratteristiche del lavoro, avendo scelto come riferimento i lavoratori nel settore dell'industria, tutti gli altri settori, ad eccezione della logistica e dei trasporti, risultavano diminuire la probabilità di perdere il lavoro, mentre il tipo di contratto determinava il più grande aumento di probabilità, con i lavoratori a tempo determinato che vedevano la quota salire dal 2 ad oltre il 10%. Infine, il periodo di rilevazione indicava che rispetto a quanti erano stati seguiti prima della pandemia, quelli seguiti durante avevano più di mezzo punto percentuale di probabilità in più di perdere il lavoro. In seguito, per ampliare le possibili dinamiche, sono state aggiunte al modello alcune interazioni mirate d'interesse con il periodo, il genere e la cittadinanza che hanno evidenziato come la presenza di minori di 15 anni nel nucleo familiare fosse un fattore penalizzante per le femmine, ma non per i maschi.

L'età più avanzata aveva un effetto che diminuiva la probabilità di perdita del lavoro per le femmine, mentre la aumentava per i maschi. Confrontando italiani e stranieri, i secondi risultavano essere meno influenzati dal tipo di contratto e più influenzati dalla minore o scarsa esperienza lavorativa rispetto ai primi, stanti al valore di riferimento le altre variabili. Infine, facendo interagire il periodo di riferimento con le altre covariate ne risultava una probabilità più alta di perdere il lavoro per gli stranieri durante la pandemia ed un effetto ancora più forte dell'aver un contratto a tempo determinato in tale periodo piuttosto che prima dell'avvento del Covid-19.

Al capitolo 5 si è cambiato tipo di analisi, propendendo per l'uso del matching statistico, volendo indagare se gli effetti di periodo, cittadinanza e ripartizione geografica restassero tali anche a seguito di un confronto tra individui simili per alcune caratteristiche di partenza. Dai risultati si evince come il periodo abbia comportato un aumento significativo dei tassi di transizione alla perdita del lavoro, anche confrontando individui simili, con una differenza tra prima e dopo ancora più alta rispetto a quella ottenuta in assenza di abbinamento. Tale aumento è stato più marcato per le femmine rispetto ai maschi. Confrontando invece italiani e stranieri, se prima della pandemia sia per i maschi che per le femmine a seguito dell'abbinamento erano gli italiani con caratteristiche simili agli stranieri a risultare svantaggiati, durante la pandemia le differenze non risultavano più essere significative, ad indicare che quanto emerso dalle analisi descrittive e dai modelli risente almeno in parte della differente composizione dei due campioni, in particolare per le caratteristiche socio-demografiche e lavorative. Infine, confrontando residenti al Centro-Nord simili ai residenti al Sud, permangono le differenze sui tassi di transizione alla perdita del lavoro, ma risultano essere attenuate rispetto a quanto accadeva senza abbinamento.

Perciò, in conclusione, si può affermare che la pandemia abbia avuto un effetto significativo sulla probabilità di perdere il lavoro a tre mesi e che si sia andata ad innestare su una situazione che già di partenza era diversa per individui diversi. In particolare, ha penalizzato ancora di più le femmine, i giovani, i lavoratori con contratti a termine e con poca esperienza, oltre che i residenti del mezzogiorno che anche a seguito dell'abbinamento con individui del Centro-Nord con caratteristiche simili risultavano avere tassi significativamente maggiori.

Le differenze riscontrate sono risultate essere robuste anche a diversi metodi di matching utilizzati per selezionare gli individui simili e ad una veloce analisi di una differente variabile che si potrebbe utilizzare per studiare il fenomeno, cioè l'aver lavorato meno di un'ora nella settimana di riferimento che ha prodotto risultati in linea, mostrando come durante la pandemia sia aumentata la percentuale di quanti hanno lavorato meno di un'ora per motivi quali la Cassa Integrazione Guadagni o l'assenza di lavoro per calo di produttività aziendale.

Ovviamente il presente lavoro non è esente da critiche e possibili sviluppi. In particolare, nelle analisi è stato considerato come periodo Covid anche il primo trimestre del 2020, che è in realtà stato colpito dalla pandemia in modo minimo dal momento che le conseguenze della prima ondata sono iniziate a farsi pesanti solo sul finire di marzo. Si potrebbe perciò pensare ad un futuro sviluppo eliminando dal campione gli individui con la seconda intervista nel primo trimestre 2020 e di conseguenza quelli con la seconda intervista nel primo trimestre 2019 per mantenere uguali gli abbinamenti. Inoltre, la variabile scelta come variabile risposta non era l'unica che si poteva usare per analizzare il fenomeno. Si potrebbe pensare di ampliare le analisi che sono state citate nell'analisi di robustezza, lavorando di più sul non aver lavorato almeno un'ora o concentrarsi in alternativa sul numero di ore lavorate. Infine, per sfruttare un periodo di osservazione più lungo ed individui che abbiano sperimentato la permanenza durante la pandemia maggiormente, si potrebbe pensare di analizzare flussi annuali, anziché trimestrali, andando ad utilizzare i dati dei trimestri 2021 quando diventeranno disponibili.

Tutti questi possibili sviluppi sono lasciati a futura trattazione.

BIBLIOGRAFIA

- Angrist, J. D., & Pischke, J. S. (2008), *Mostly harmless econometrics*, Princeton University press.
- Di Pasquale, E., & Tronchin, C. (2021), *Donna e straniera: ecco chi paga la crisi da pandemia*, lavoce.info.
- Gazzelloni, S. (a cura di, 2006), *La rilevazione sulle forze di lavoro: contenuti, metodologie, organizzazione*, Metodi e norme n. 32, Istituto Nazionale di Statistica, Roma.
- ISTAT (2019), *Rapporto Annuale 2019. La situazione del Paese*, Roma
- ISTAT (2020), *Il Mercato del Lavoro 2019. Una lettura integrata*, Roma
- ISTAT (2021), *Rapporto Annuale 2020. La situazione del Paese*, Roma
- ISTAT (2020), *Il Mercato del Lavoro 2020. Una lettura integrata*, Roma
- Martini, A., & Sisti, M. (2009). *Valutare il successo delle politiche pubbliche*, Il Mulino, Bologna.
- Paggiaro, A. (2013). *How do immigrants fare during the downturn? Evidence from matching comparable natives*. Demographic research, 28, 229-258.
- Salvan, A., Sartori, N., & Pace, L. (2020), *Modelli Lineari Generalizzati*, Springer, Milano

RINGRAZIAMENTI

Quando in una tesi si arriva a questo “capitolo” è il segno che il percorso è oramai quasi finito. Allora mi fermo a riguardare questo percorso, questi due (+ due) anni che ho impiegato per raggiungere il traguardo, e non posso fare a meno di vederci tanti momenti difficili che mi hanno quasi portato a dire basta, ma anche tantissimi altri momenti belli e densi di ricordi che mi porterò dentro per il resto della vita per i quali non vale la pena, bensì la gioia, di ringraziare le persone che hanno contribuito a formarli.

In primis ringrazio la mia famiglia, mamma e papà su tutti, perché è solo grazie a loro se mi son potuto permettere di arrivare fino a qui senza dovermi preoccupare di qualsiasi altra cosa e perché soprattutto nell’ultimo anno mi hanno fatto sentire la loro vicinanza e non mi hanno mai fatto pesare le scelte difficili che ho dovuto prendere, anche se diverse da quelle che avrebbero preso loro.

Ringrazio il professor Dalla Zuanna per avermi accettato come tesista seppur non mi avesse mai avuto prima come studente ai suoi corsi, il suo aiuto per concludere in tempo questo lavoro è stato fondamentale. Ringrazio il professor Paggiaro che da aprile è stato sempre presenza importante e discreta in questo lavoro. Lo ringrazio per la grandissima disponibilità che ha dimostrato nei miei confronti e anche per avermi lasciato libero di sporcarmi le mani con i dati nel modo che credevo essere migliore e avermi sempre aiutato poi a riordinare il tutto.

Ringrazio i miei compagni Claudia, Filippo e Tommaso per le ore di lezione passate insieme, per le pause pranzo alla Rocca e gli aperitivi da Carlotta.

Martina e Simone per i nostri giornalieri confronti sugli studi e sul mondo e per avermi sempre spronato a fare del mio meglio durante le analisi e la stesura di questa tesi.

Maura perché oltre che una compagna per un certo periodo è stata anche una collega negli uffici dello IOV e in certi momenti è stata anche mia psicologa e confidente. Forse senza il suo aiuto queste parole avrebbero dovuto attendere molti altri mesi prima di essere scritte.

Ringrazio Jacopo, per esserci stato sempre senza dover dire nulla, come solo i veri amici sanno fare.

Infine, ringrazio Silvia. Da quasi quattro anni è al mio fianco e ancora non si è stancata di quel rompiscatole puntiglioso che so essere e del quale io stesso mi stanco ogni tanto. Ha sopportato i miei momenti di pazzia e supportato qualsiasi mia folle idea. Lei è stata àncora e porto sicuro per tutto questo periodo, sapendo farmi vedere la luce in fondo al tunnel e ricordandomi sempre quello che sono in grado di fare ma che spesso dimenticavo. Senza di lei ritrovare la serenità e avere la testa per concludere questo percorso sarebbe stato estremamente più difficile, se non impossibile. Questo mio lavoro è dedicato anche a lei, nella speranza che resti al mio fianco a lungo e sia sempre bussola per le nostre scelte future.

APPENDICE

A1. Flussi trimestrali disaggregati

Tabella A1. Flussi trimestrali di COND3 per i cittadini stranieri comunitari nel periodo pre-pandemia.

Prima rilevazione	Seconda Rilevazione			
	Occupato	In cerca di lavoro	Inattivo	Totale
Occupato	2847 94.9	69 2.3	84 2.8	3000 100.00
In cerca di lavoro	81 17.09	268 56.54	125 26.37	474 100
Inattivo	79 6.16	120 9.36	1083 84.48	1282 100
Totale	3007 63.23	457 9.61	1292 27.17	4756 100

Tabella A2. Flussi trimestrali di COND3 per i cittadini stranieri comunitari nel periodo durante la pandemia.

Prima rilevazione	Seconda Rilevazione			
	Occupato	In cerca di lavoro	Inattivo	Totale
Occupato	2385 92.51	62 2.4	131 5.08	2578 100.00
In cerca di lavoro	59 15.09	173 44.25	159 40.66	391 100
Inattivo	96 7.1	117 8.65	1139 84.25	1352 100
Totale	2540 58.78	352 8.15	1429 33.07	4321 100

Tabella A3. Flussi trimestrali di COND3 per i cittadini stranieri extra-comunitari nel periodo pre-pandemia.

	Seconda Rilevazione			
Prima rilevazione	Occupato	In cerca di lavoro	Inattivo	Totale
Occupato	5845 94.98	151 2.45	158 2.57	6154 100.00
In cerca di lavoro	188 16.92	627 56.44	296 26.64	1111 100
Inattivo	157 4.71	267 8.01	2910 87.28	3334 100
Totale	6190 58.4	1045 9.86	3364 31.74	10599 100

Tabella A4. Flussi trimestrali di COND3 per i cittadini stranieri extra-comunitari nel periodo durante la pandemia.

	Seconda Rilevazione			
Prima rilevazione	Occupato	In cerca di lavoro	Inattivo	Totale
Occupato	4789 92.95	134 2.6	229 4.44	5152 100.00
In cerca di lavoro	132 16.84	361 46.05	291 37.12	784 100
Inattivo	151 4.76	260 8.2	2760 87.04	3171 100
Totale	5072 55.69	755 8.29	3280 36.02	9107 100

Tabella A5. Flussi trimestrali di COND3 per i residenti al Centro nel periodo pre-pandemia.

	Seconda Rilevazione			
Prima rilevazione	Occupato	In cerca di lavoro	Inattivo	Totale
Occupato	16461 96.28	205 1.2	431 2.52	17097 100
In cerca di lavoro	284 16.12	803 45.57	675 38.31	1762 100
Inattivo	339 3.98	518 6.08	7662 89.94	8519 100
Totale	17084 62.4	1526 5.57	8768 32.03	27378 100

Tabella A6. Flussi trimestrali di COND3 per i residenti al Centro nel periodo durante la pandemia.

	Seconda Rilevazione			
Prima rilevazione	Occupato	In cerca di lavoro	Inattivo	Totale
Occupato	15147 95.65	212 1.34	477 3.01	15836 100
In cerca di lavoro	207 14.13	525 35.84	733 50.03	1465 100
Inattivo	376 4.53	509 6.13	7418 89.34	8303 100
Totale	15730 61.44	1246 4.87	8628 33.7	25604 100

Tabella A7. Flussi trimestrali di COND3 per i residenti al Sud nel periodo pre-pandemia.

	Seconda Rilevazione			
Prima rilevazione	Occupato	In cerca di lavoro	Inattivo	Totale
Occupato	22243 94.06	452 1.91	953 4.03	23648 100
In cerca di lavoro	509 10.2	2241 44.91	2240 44.89	4990 100
Inattivo	842 3.52	2049 8.58	21000 87.9	23891 100
Totale	23594 44.92	4742 9.03	24193 46.06	52529 100

Tabella A8. Flussi trimestrali di COND3 per i residenti al Sud nel periodo durante la pandemia.

	Seconda Rilevazione			
Prima rilevazione	Occupato	In cerca di lavoro	Inattivo	Totale
Occupato	19954 93.5	384 1.8	1004 4.7	21342 100
In cerca di lavoro	397 10.11	1458 37.15	2070 52.74	3925 100
Inattivo	905 4.03	1768 7.88	19757 88.08	22430 100
Totale	21256 44.56	3610 7.57	22831 47.87	47697 100

A2. Distribuzione caratteristiche pre-matching

Tabella A9. Distribuzione delle caratteristiche demografiche per genere e periodo

Variabile	Livello	Maschi 2019	Maschi 2020	Femmine 2019	Femmine 2020
Età	Giovani (15-34)	22.2	21.7	19.4	18.2
	Adulti (35-54)	57.6	58.8	58.6	58.2
	Ultracinquantenni (55-64)	20.2	21.6	22.1	23.6
Titolo di Studio	Fino a licenza media	36.2	34.2	23.7	22.5
	Diploma	47.8	48.8	48.5	48.6
	Laurea o superiore	16.0	17.1	27.8	28.9
Stato civile	Celibi/Nubili	34.9	35.7	31.4	31.8
	Sposati/Conviventi	56.9	56.4	51.8	52.1
	Divorziati/Vedovi	8.2	7.9	16.8	16.1
Cittadinanza	Stranieri	11.7	11.1	11.7	10.7
Ripartizione geografica	Nord	52.7	52.2	56.1	56.1
	Centro	18.7	19.2	20.3	20.7
	Sud	28.6	28.5	23.7	23.2

Tabella A10. Distribuzione delle caratteristiche familiari per genere e periodo

Variabile	Livello	Maschi 2019	Maschi 2020	Femmine 2019	Femmine 2020
Numero Componenti Famiglia	1 Componente	13.4	13.3	13.9	13.6
	2 Componenti	17.9	17.8	24.1	24.2
	3 Componenti	28.9	29.2	28.7	29.2
	4 Componenti	30.2	30.3	26.1	26.2
	5+ Componenti	9.7	9.3	7.2	6.8
Caratteristiche altri componenti	Presenza <15 anni	36.1	34.8	33.1	32.4
	Presenza >64 anni	4.5	4.8	4.4	4.6
	Presenza laureati	28.2	29.7	36.0	37.2
	Presenza diplomati	68.4	69.3	69.2	69.5
	Presenza disoccupati	1.5	1.6	1.6	1.5
Altri Occupati	Uno	44.0	44.6	52.9	52.9
	Duo e opiu	11.0	10.6	11.2	10.7

Tabella A11. Distribuzione delle caratteristiche lavorative per genere e periodo

Variabile	Livello	Maschi 2019	Maschi 2020	Femmine 2019	Femmine 2020
Tipo contratto	Tempo determinato	16.6	15.2	17.2	15.7
	Tempo indeterminato	83.4	84.8	82.8	84.3
Tipo orario	Part-time	8.2	8.1	34.5	33.6
	Full-time	91.8	91.9	65.5	66.4
Settore lavorativo	Agricoltura	3.9	3.9	1.7	1.7
	Industria	31.5	31.6	12.7	12.7
	Costruzioni	8.3	8.1	0.8	1.0
	Commercio	10.5	10.7	11.7	11.9
	Alberghi e ristoranti	4.8	4.5	7.3	6.5
	Logistica	8.1	8.1	2.6	2.6
	Servizi Pubblici	14.7	14.8	10.8	11.0
	Imprenditoria	6.4	6.4	10.6	10.2
	Istruzione, Sanità, Servizi sociali	8.2	8.6	29.9	31.0
	Altri servizi alla persona	3.6	3.3	12.1	11.6
Durata del lavoro attuale	Meno di 2 anni	20.7	20.0	20.8	20.1
	3-5 anni	16.1	17.1	16.3	16.5
	6-10 anni	13.9	13.4	14.7	13.6
	11-20 anni	24.5	24.5	25.3	25.9
	Più di 20 anni	24.8	25.1	22.9	24.0
Situazione anno precedente	Occupato	92.9	93.6	91.4	92.1
	Disoccupato	5.4	4.8	5.4	5.1
	Inattivo	1.4	1.4	2.8	2.5
	Altro	0.3	0.2	0.4	0.4

Tabella A12. Distribuzione delle caratteristiche demografiche per periodo, genere e cittadinanza

Variabile	Livello	Maschi 2019		Maschi 2020		Femmine 2019		Femmine 2020	
		Stranieri	Italiani	Stranieri	Italiani	Stranieri	Italiani	Stranieri	Italiani
Età	Giovani (15-34)	30.7	21.0	30.1	20.6	21.9	19.0	20.1	18.0
	Adulti (35-54)	60.1	57.3	60.7	56.3	60.6	58.3	61.2	57.9
	Ultracinquantenni (55-64)	9.2	21.7	9.2	23.1	17.5	22.7	18.7	24.1
Titolo di Studio	Fino a licenza media	56.9	33.4	54.9	31.6	42.6	21.2	42.5	20.1
	Diploma	36.1	49.4	37.9	50.1	42.5	49.3	42.8	49.3
	Laurea o superiore	7.0	17.2	7.2	18.3	14.9	29.5	14.7	30.6
Stato civile	Celibati/Nubili	28.2	35.8	28.3	36.6	27.3	32.0	27.4	32.4
	Sposati/Conviventi	58.6	56.7	59.8	56.0	38.4	53.6	38.6	53.7
	Divorziati/Vedovi	13.2	7.5	11.9	7.4	34.3	14.5	34.0	13.9
Ripartizione geografica	Nord	61.7	51.4	61.8	51.0	58.4	55.7	59.7	55.7
	Centro	23.6	18.1	23.9	18.6	26.2	19.5	27.4	19.8
	Sud e Isole	14.7	30.5	14.3	30.3	15.4	24.8	12.9	24.5

Tabella A13. Distribuzione delle caratteristiche familiari per periodo, genere e cittadinanza

Variabile	Livello	Maschi 2019		Maschi 2020		Femmine 2019		Femmine 2020	
		Stranieri	Italiani	Stranieri	Italiani	Stranieri	Italiani	Stranieri	Italiani
Numero Componenti Famiglia	1 Componente	19.3	12.5	17.5	12.8	30.1	11.8	28.1	11.8
	2 Componenti	14.2	18.4	14.6	18.2	21.2	24.5	22.8	24.4
	3 Componenti	22.0	29.8	22.2	30.1	20.0	29.8	20.1	30.3
	4 Componenti	26.2	30.7	28.2	30.6	18.1	27.2	18.6	27.2
	5+ Componenti	18.3	8.6	17.4	8.3	10.6	6.7	10.4	6.3
Caratteristiche componenti	Presenza <15 anni	47.4	34.6	48.0	33.2	33.9	33.0	33.9	32.3
	Presenza >64 anni	0.9	4.9	1.0	5.2	2.0	4.7	1.8	4.9
	Presenza laureati	14.5	30.0	14.4	31.6	19.4	38.2	19.0	39.4
	Presenza diplomati	49.6	70.9	51.9	71.5	57.2	70.8	56.8	71.0
	Presenza disoccupati	11.4	7.8	9.7	7.1	8.9	6.1	8.7	5.2
Altri Occupati	Uno	33.8	45.3	35.0	45.7	41.2	54.5	42.7	54.1
	Duo e opù	10.7	11.1	9.8	10.8	10.6	11.3	9.2	10.9

Tabella A14. Distribuzione delle caratteristiche lavorative per periodo, genere e cittadinanza

Variabile	Livello	Maschi 2019		Maschi 2020		Femmine 2019		Femmine 2020	
		Stranieri	Italiani	Stranieri	Italiani	Stranieri	Italiani	Stranieri	Italiani
Tipo contratto	Tempo determinato	27.4	15.2	25.9	13.9	18.6	17.0	17.9	15.4
	Tempo indeterminato	72.6	84.8	74.1	86.1	81.4	82.9	82.1	84.6
Tipo orario	Part-time	12.3	7.7	12.1	7.6	42.2	33.5	42.8	32.5
	Full-time	87.7	92.3	87.9	92.4	57.8	66.5	57.2	67.5
Settore lavorativo	Agricoltura	12.4	2.8	12.3	2.8	3.1	1.5	3.3	1.5
	Industria	30.6	31.7	30.1	31.8	8.7	13.2	8.8	13.1
	Costruzioni	14.0	7.6	14.7	7.3	0.3	0.9	0.3	1.1
	Commercio	8.1	10.8	8.0	11.0	5.5	12.6	5.5	12.7
	Alberghi e ristoranti	8.8	4.3	8.2	4.0	12.7	6.6	11.5	5.9
	Logistica	9.1	8.0	10.3	7.9	1.4	2.7	1.7	2.7
	Servizi Pubblici	0.6	16.6	0.7	16.6	0.6	12.1	0.7	12.2
	Imprenditoria	6.6	6.3	6.3	6.5	8.0	11.0	7.7	10.5
	Istruzione, Sanità, Servizi sociali	2.0	9.1	2.3	9.4	8.7	32.7	10.5	33.4
	Altri servizi alla persona	7.8	3.0	7.2	2.9	51.1	6.9	50.1	6.9
Durata del lavoro attuale	Meno di 2 anni	29.1	19.6	27.1	19.1	31.0	19.4	31.5	18.7
	3-5 anni	25.5	14.9	26.7	15.8	29.0	14.6	28.7	15.0
	6-10 anni	22.0	12.8	20.9	12.4	22.7	13.7	19.7	12.8
	11-20 anni	21.3	25.0	22.6	24.7	15.9	26.6	18.5	26.8
	Più di 20 anni	2.0	27.8	2.6	27.9	1.5	25.7	1.6	26.7
Situazione anno precedente	Occupato	91.3	93.2	92.3	93.7	87.5	92.0	87.3	92.7
	Disoccupato	7.3	5.1	6.2	4.6	8.3	5.0	8.3	4.7
	Inattivo	1.0	1.5	1.2	1.5	3.5	2.6	3.8	2.3
	Altro	0.4	0.2	0.3	0.2	0.7	0.4	0.5	0.4

Tabella A15. Distribuzione delle caratteristiche demografiche per periodo, genere e ripartizione geografica di residenza

Variabile	Livello	Maschi 2019		Maschi 2020		Femmine 2019		Femmine 2020	
		Nord	Sud	Nord	Sud	Nord	Sud	Nord	Sud
Età	Giovani (15-34)	22.6	21.1	22.4	19.9	19.6	18.5	18.5	17.2
	Adulti (35-54)	58.6	55.3	57.5	55.0	59.3	56.2	58.9	56.2
	Ultracinquantenni (55-64)	18.8	23.6	20.2	25.1	21.1	25.3	22.6	26.6
Titolo di Studio	Fino a licenza media	34.4	40.6	32.5	38.3	23.4	24.6	22.5	22.6
	Diploma	48.9	45.1	49.8	46.2	49.6	45.1	49.6	45.4
	Laurea o superiore	16.7	14.2	17.7	15.5	27.0	30.3	28.0	32.1
Stato civile	Celibi/Nubili	37.8	27.7	38.7	28.2	32.4	28.3	33.2	27.4
	Sposati/Conviventi	53.5	65.6	52.9	65.4	50.7	55.3	50.5	57.3
	Divorziati/Vedovi	8.7	6.7	8.5	6.4	16.9	16.4	16.3	15.4
Cittadinanza	Stranieri	14.0	6.0	13.3	5.5	13.0	7.6	12.2	6.0

Tabella A16. Distribuzione delle caratteristiche familiari per periodo, genere e ripartizione geografica di residenza

Variabile	Livello	Maschi 2019		Maschi 2020		Femmine 2019		Femmine 2020	
		Nord	Sud	Nord	Sud	Nord	Sud	Nord	Sud
Numero Componenti Famiglia	1 Componente	14.4	10.8	14.4	10.6	14.0	13.6	13.9	12.4
	2 Componenti	19.0	15.3	18.9	15.3	24.7	22.2	25.1	21.4
	3 Componenti	29.1	28.3	29.6	28.4	28.8	28.3	29.2	29.4
	4 Componenti	28.3	34.8	28.4	35.2	25.5	28.0	25.3	29.2
	5+ Componenti	9.3	10.8	8.8	10.6	7.0	7.9	6.6	7.6
Caratteristiche altri componenti	Presenza <15 anni	35.6	37.2	34.4	35.9	33.7	31.2	32.7	31.6
	Presenza >64 anni	4.8	3.7	5.0	4.2	4.1	5.2	4.3	5.4
	Presenza laureati	29.3	25.6	30.5	27.6	35.1	38.8	36.1	40.9
	Presenza diplomati	69.0	67.0	69.6	68.8	70.0	66.6	69.9	67.9
	Presenza disoccupati	1.2	2.2	1.3	2.2	1.3	2.3	1.4	1.9
Altri Occupati	Uno	47.4	35.6	47.9	36.3	54.2	48.7	53.8	50.0
	Duo e opiu	12.5	7.4	11.8	7.8	12.1	8.2	11.4	8.4

Tabella A17. Distribuzione delle caratteristiche lavorative per periodo, genere e ripartizione geografica di residenza

Variabile	Livello	Maschi 2019		Maschi 2020		Femmine 2019		Femmine 2020	
		Nord	Sud	Nord	Sud	Nord	Sud	Nord	Sud
Tipo contratto	Tempo determinato	15.2	20.3	13.7	19.1	15.6	22.3	14.4	19.9
	Tempo indeterminato	84.8	79.7	86.3	80.9	84.4	77.7	85.6	80.1
Tipo orario	Part-time	7.0	11.3	6.8	11.4	35.0	33.0	34.1	32.0
	Full-time	93.0	88.7	93.2	88.6	65.0	67.0	65.9	68.0
Settore lavorativo	Agricoltura	2.5	7.3	2.6	7.0	0.9	4.0	1.0	3.9
	Industria	35.7	21.3	35.5	21.9	14.5	6.6	14.4	6.9
	Costruzioni	8.0	9.0	7.7	8.9	0.9	0.6	1.1	0.6
	Commercio	10.3	11.0	10.4	11.3	12.0	11.0	12.1	11.3
	Alberghi e ristoranti	4.4	5.9	4.1	5.3	7.5	6.6	6.9	5.5
	Logistica	8.2	8.0	8.1	8.2	2.7	2.1	2.7	2.2
	Servizi Pubblici	13.9	16.7	14.3	16.2	10.8	10.8	11.1	10.7
	Imprenditoria	6.4	6.3	6.4	6.4	11.1	9.1	10.6	8.8
	Istruzione, Sanità, Servizi sociali	7.2	10.7	7.5	11.4	27.7	37.0	28.5	39.2
	Altri servizi alla persona	3.5	3.8	3.3	3.4	12.0	12.3	11.7	11.0
Durata del lavoro attuale	Meno di 2 anni	20.6	20.9	19.6	21.0	20.3	22.5	19.7	21.2
	3-5 anni	16.3	15.6	17.7	15.5	16.1	16.6	16.5	16.3
	6-10 anni	14.4	12.7	13.6	12.6	14.9	14.1	13.6	13.6
	11-20 anni	25.1	23.1	25.1	22.8	26.1	22.8	26.5	24.0
	Più di 20 anni	23.6	27.8	23.9	28.1	22.6	24.0	23.7	24.9
Situazione anno precedente	Occupato	94.0	90.3	94.8	90.6	92.4	88.3	92.9	89.5
	Disoccupato	4.1	8.6	3.5	8.2	4.5	8.2	4.3	7.7
	Inattivo	1.6	0.9	1.6	1.0	2.7	3.1	2.5	2.5
	Altro	0.2	0.3	0.2	0.3	0.4	0.5	0.4	0.4