



UNIVERSITÀ
DI PAVIA

DIPARTIMENTO DI SCIENZE POLITICHE E SOCIALI

CORSO DI LAUREA IN SVILUPPO ECONOMICO E RELAZIONI INTERNAZIONALI

TITOLO

L'IMPATTO DELLO SMART WORKING SULLA PRODUTTIVITÀ DEL LAVORO:
EVIDENZE ECONOMETRICHE DAL REGNO UNITO

Relatore:

Chiar.mo Prof. Francesco Scervini

Correlatore:

Chiar.mo Prof. Emanuele Forlani

Tesi di laurea di Giulio Adamo

Matricola n. 558497

ANNO ACCADEMICO 2025/26

Sommario

Introduzione	4
1. Literature review	8
2. Dati e strategia empirica	20
2.1. Definizione e costruzione delle variabili	21
2.1.1. Criteri di campionamento e selezione del sottocampione.....	22
2.1.2. La variabile dipendente: costruzione della proxy di produttività oraria	23
2.2. Le variabili esplicative: margine intensivo ed estensivo del WFH.....	25
2.2.1. Il ruolo e l'esclusione di <i>jbflex7</i>	26
2.2.2. L'esclusione di <i>wkhomewil</i>	28
2.2.3. Costruzione delle variabili armonizzate finali	29
2.3. Variabili di controllo	30
2.3.1. Età (<i>dvage</i>)	30
2.3.2. Genere (<i>sex</i>)	32
2.3.3. Livello di istruzione (<i>hiqua1_dv</i>)	34
2.3.4. Settore industriale (<i>jbsic07_cc</i>)	36
2.3.5. Contesto geografico e residenziale (<i>country</i> e <i>urban_dv</i>)	38
2.3.6. Numero di figli minori di 16 anni (<i>nchresp</i>)	40
2.3.7. Frequenza di utilizzo di internet (<i>netpusenew</i>)	42
3. Strategia empirica e analisi causale.....	46
3.1. Obiettivo dell'analisi e sfide metodologiche.....	46
3.2. La stima preliminare: <i>pooled OLS</i>	48
3.2.1. La strategia di stima a blocchi progressivi.....	49
3.2.2. Limiti del modello <i>pooled OLS</i> e correzione degli errori.....	50
3.3. Il modello panel a effetti fissi individuali.....	52
3.4. Risultati attesi e ipotesi di ricerca	55
3.4.1. Ipotesi 1: il premio di produttività nel modello <i>pooled OLS</i>	55
3.4.2. Ipotesi 2: la riduzione dell'effetto e il ruolo della non-significatività nel modello Panel (<i>Fixed Effects</i>)	56
4. Analisi empirica e risultati del modello	58
4.1. Il modello di benchmark: stime ad effetti fissi (FE) individuali	58
4.1.1. Analisi del margine estensivo: l'impatto dell'accesso al WFH	59
4.1.2. Il ruolo dei controlli: shock temporali e impatto di internet	62

4.1.3.	Analisi del margine intensivo: l'impatto della frequenza di WFH..	63
4.1.4.	Presentazione dei risultati empirici e assenza di significatività statistica	65
4.1.5.	Ragioni teoriche della non significatività diffusa	66
4.2.	Stima dei modelli OLS (margine estensivo e intensivo)	67
4.2.1.	Il premio di produttività nel margine estensivo	69
4.2.2.	La non-linearità nel margine intensivo.....	70
4.2.3.	Il ruolo delle caratteristiche demografiche e del capitale umano ...	73
4.2.4.	Interpretazione critica del quadro OLS: l'effetto selezione	74
4.3.	OLS e FE: il confronto metodologico	75
Conclusioni.....		78
Bibliografia.....		81
Sitografia.....		83

Introduzione

Negli ultimi decenni, le economie avanzate hanno assistito a una profonda riorganizzazione delle proprie strutture occupazionali, guidata dalla transizione digitale e dalla diffusione pervasiva delle tecnologie dell'informazione e della comunicazione (ICT). Questo processo di modernizzazione ha progressivamente allentato il tradizionale vincolo della compresenza fisica all'interno degli spazi aziendali, dimostrando come una quota crescente di mansioni d'ufficio e ad alto valore cognitivo possa essere svolta efficacemente anche a distanza. Tuttavia, se fino al 2020 il lavoro da remoto (*Work From Home*, WFH) rappresentava una pratica contrattuale di nicchia e limitata a poche figure professionali, lo shock esogeno e impreveduto della pandemia di COVID-19 ne ha determinato un'adozione massiccia, improvvisa e generalizzata. Nel giro di poche settimane, lo smart working è stato trasformato da opzione di welfare secondaria a realtà strutturale e strumento di continuità operativa per milioni di lavoratori dipendenti in tutto il mondo.

Oggi che l'emergenza sanitaria si è definitivamente conclusa, il mercato del lavoro non è tornato alle abitudini pre-pandemiche. Il WFH non è scomparso, ma si è stabilizzato su livelli storicamente e sistematicamente superiori rispetto al passato, ridefinendo l'organizzazione del tempo e dello spazio urbano. Questo scenario inedito ha riaperto un dibattito cruciale e quanto mai attuale tra accademici, manager e decisori politici: l'impatto del lavoro da remoto sulla produttività del lavoro. La produttività oraria rappresenta una variabile cardinale per la crescita economica di lungo periodo e per la competitività delle imprese; comprendere se e come lo spostamento fisico della prestazione lavorativa influenzi l'efficienza dei fattori produttivi è dunque essenziale per orientare i futuri assetti

contrattuali, la gestione del capitale umano e le stesse politiche di welfare aziendale.

Work from Home over Time in the United States

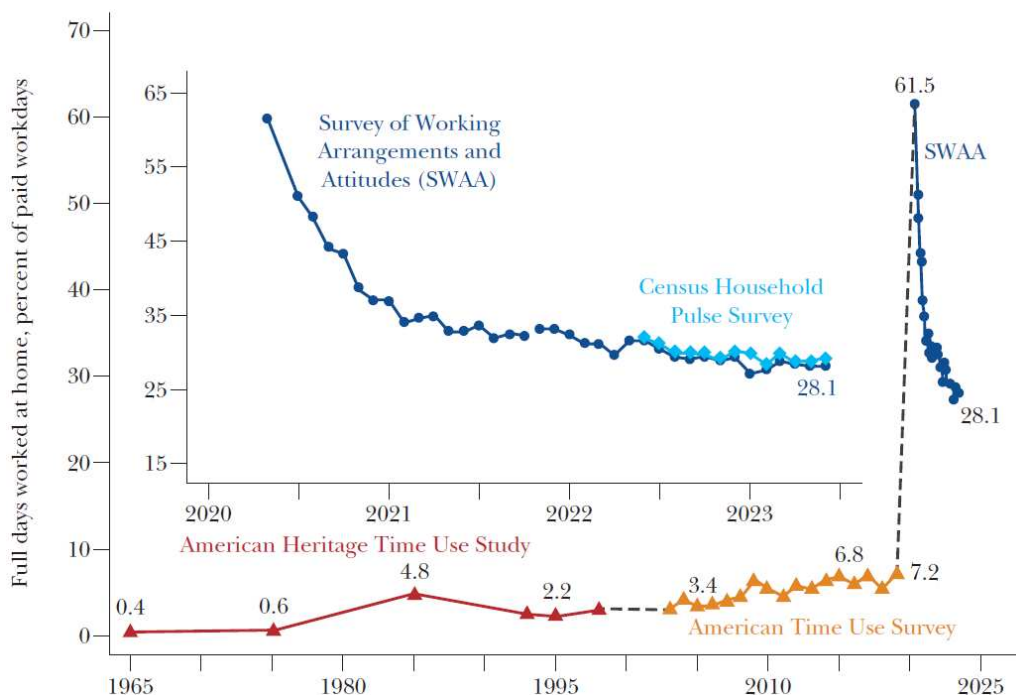


Figura 1 – Evoluzione del lavoro da remoto nel tempo. La figura mostra l'andamento della quota di giorni lavorativi svolti da casa negli Stati Uniti, evidenziando il forte incremento durante la pandemia e la successiva stabilizzazione su livelli significativamente superiori rispetto al periodo pre-pandemico. Fonte: Barrero et al. (2023).

Sotto il profilo strettamente metodologico, tuttavia, lo studio di questa relazione solleva complesse sfide di identificazione causale che la letteratura economica ha cercato a lungo di risolvere. Nella realtà occupazionale, l'adozione dello smart working non avviene in modo casuale o uniforme, ma è il risultato di un doppio processo di selezione e di eleggibilità: da un lato, sono spesso i lavoratori più motivati, disciplinati o con specifiche esigenze di conciliazione tra vita privata e professionale a richiedere formule di flessibilità; dall'altro, sono le imprese operanti nei settori a più alta intensità tecnologica o dotate di migliori prassi

manageriali a concederle. Di conseguenza, una semplice analisi trasversale che si limiti a confrontare la produttività di chi lavora a casa con chi lavora in ufficio rischierebbe di generare stime gravemente distorte da “errore da variabili omesse”. In altre parole, si correrebbe il rischio di attribuire al WFH un incremento di efficienza che è invece originato dalle doti personali latenti del lavoratore, dalle sue competenze digitali di base o dalla qualità dell’infrastruttura dell’azienda di appartenenza.

Questo lavoro di ricerca si propone di superare tali limiti strutturali offrendo un’analisi econometrica longitudinale applicata al contesto del mercato del lavoro del Regno Unito, un ecosistema storicamente reattivo alle innovazioni organizzative e caratterizzato da un’elevata diffusione del lavoro agile. L’indagine empirica attinge ai microdati estratti da *Understanding Society: the UK Household Longitudinal Study* (UKHLS). Si tratta di uno dei panel surveys¹ familiari più estesi e dettagliati a livello internazionale, la cui ricchezza informativa permette di seguire gli stessi individui ripetutamente nel corso del tempo, catturando l’evoluzione interna del fenomeno in una prospettiva di medio periodo e focalizzandosi sulla cruciale fase di stabilizzazione post-pandemica.

La natura longitudinale dei dati consente l’implementazione di un modello a “effetti fissi individuali” (*Fixed Effects*). Attraverso questa specifica strategia matematica, il fulcro del confronto si sposta: invece di paragonare individui diversi tra loro, ogni lavoratore viene confrontato esclusivamente con se stesso nel corso del tempo. Questo approccio permette di azzerare l’impatto di tutte le caratteristiche statiche non osservabili del soggetto (come l’abilità innata o l’ambizione), isolando la pura variazione intrapersonale. L’obiettivo è verificare se e come cambi la produttività oraria dello *stesso* individuo nel momento in cui transita dal

¹ Insieme di dati che segue le stesse unità statistiche (ad esempio individui, famiglie o imprese) nel tempo, raccogliendo osservazioni ripetute per ciascuna unità in diversi periodi, al fine di analizzarne le dinamiche temporali.

lavoro in presenza allo smart working, analizzando l'effetto sia in termini assoluti di accesso alla misura (margine estensivo) sia in base all'intensità e al numero di giorni spesi a casa durante la settimana (margine intensivo). Per garantire la massima robustezza e affidabilità delle stime, il modello introduce un set articolato di variabili di controllo socio-demografiche, familiari e di contesto, tra cui l'andamento dell'età, la presenza e il numero di figli minori nel nucleo familiare, il grado di urbanizzazione residenziale e il livello di familiarità e alfabetizzazione digitale del lavoratore, misurato attraverso la frequenza d'uso personale di internet.

Il presente lavoro è strutturato come segue. Il primo capitolo delinea i confini teorici ed empirici emersi nel dibattito scientifico globale, esaminando i canali di efficienza e le criticità organizzative evidenziate dalla letteratura recente. Il secondo capitolo è dedicato alla descrizione analitica della banca dati *Understanding Society*, illustrando i criteri di campionamento e i processi di costruzione delle variabili econometriche e della proxy² di produttività oraria. Il terzo capitolo formalizza la strategia empirica, descrivendo le ipotesi di ricerca e il passaggio metodologico dal modello preliminare *pooled* OLS alla specificazione a effetti fissi individuali. Infine, il quarto capitolo presenta nel dettaglio i risultati delle stime econometriche, discutendo criticamente l'assenza di significatività statistica dei coefficienti e interpretando il fenomeno alla luce delle attuali dinamiche economiche e manageriali.

² Variabile utilizzata come approssimazione di un fenomeno o di una grandezza non direttamente osservabile o misurabile, al fine di rappresentarne empiricamente gli effetti nell'analisi.

1. Literature review

La letteratura economica sul WFH ha conosciuto una forte espansione negli ultimi anni, soprattutto a seguito dello shock esogeno rappresentato dalla pandemia di COVID-19, che ha reso il lavoro a distanza una modalità diffusa in numerosi settori produttivi. Tuttavia, il dibattito accademico sull'impatto del WFH sulla produttività è antecedente alla pandemia e presenta risultati eterogenei, sia in termini di segno che di intensità degli effetti stimati.

Uno dei contributi più influenti nella letteratura pre-pandemica è quello di Bloom et al. (2015), che analizzano un esperimento controllato su lavoratori di un call center in Cina. Le stime ottenute dagli autori indicano un incremento della produttività individuale del 13% in seguito all'introduzione del lavoro da casa, effetto attribuito principalmente alla riduzione delle interruzioni, alla maggiore concentrazione e a una diminuzione dell'assenteismo. Inoltre, lo studio evidenzia un aumento della soddisfazione lavorativa e una riduzione del turnover, suggerendo che il WFH possa generare benefici sia per i lavoratori sia per le imprese in contesti caratterizzati da compiti altamente standardizzati e facilmente misurabili.

Tuttavia, risultati più recenti mostrano un quadro meno univoco, evidenziando come gli effetti del lavoro da remoto dipendano fortemente dalla natura delle attività svolte. Ad esempio, evidenze basate su dati amministrativi aziendali e studi su sviluppatori di software durante il periodo pandemico mostrano che il WFH può avere effetti differenziati a seconda della complessità del progetto e della struttura del lavoro. In particolare, alcune analisi evidenziano che la produttività può diminuire nei progetti più complessi o che richiedono elevati livelli di coordinamento, mentre può rimanere stabile o aumentare in attività più individualizzabili (Bao et al., 2020; Yang et al., 2020).

Un ulteriore filone della letteratura si concentra sugli effetti del WFH durante il periodo pandemico, interpretandolo come un esperimento naturale su larga scala. In questo contesto, Barrero et al. (2023) documentano un cambiamento strutturale nelle modalità lavorative, evidenziando che una quota significativa del lavoro rimane svolta da remoto anche dopo la fase acuta della pandemia. Gli autori sottolineano inoltre che la produttività percepita dai lavoratori tende ad essere stabile o in alcuni casi superiore rispetto al lavoro in presenza, ma con forti eterogeneità settoriali e individuali, in particolare legate all'età, alla presenza di figli e al tipo di occupazione.

In modo complementare, Aksoy et al. (2022) mostrano che le preferenze dei lavoratori e dei datori di lavoro divergono sistematicamente: i lavoratori tendono a preferire una maggiore intensità di lavoro da casa rispetto a quanto inizialmente previsto dalle imprese, suggerendo l'esistenza di un potenziale surplus di benessere associato alla flessibilità lavorativa. Tuttavia, anche in questo caso, gli effetti sul rendimento produttivo risultano non omogenei tra paesi, settori e categorie occupazionali.

Figure 1. Working from Home Is Now a Global Phenomenon

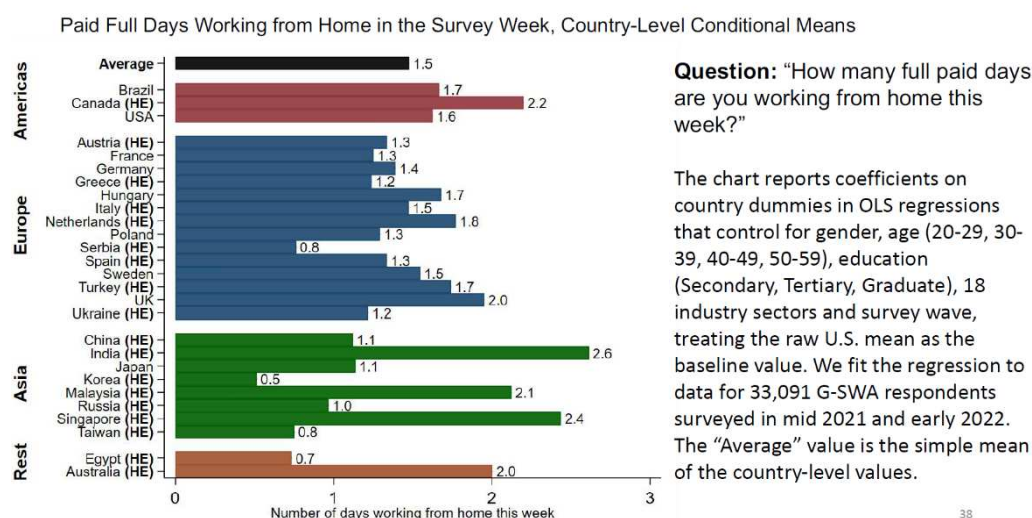


Figura 2 – Diffusione del lavoro da remoto tra paesi. La figura mostra la distribuzione del lavoro da remoto a livello internazionale, evidenziando differenze significative tra i paesi. Fonte: Aksoy et al. (2022).

Parallelamente, una parte consistente della letteratura recente ha messo in evidenza i meccanismi attraverso cui il lavoro da remoto può influenzare la produttività. Alcuni studi evidenziano che il WFH modifica la struttura delle attività lavorative, aumentando la quota di compiti non routinari e analitici e riducendo le attività manuali o strettamente coordinate, con conseguenze potenzialmente positive sulla produttività e sui salari nel medio periodo (Jiang et al., 2024). Tuttavia, tali effetti dipendono fortemente dalla capacità delle imprese di riorganizzare il lavoro e dalla natura tecnologica e organizzativa delle mansioni.

Un risultato importante emerso dalla letteratura è quindi che non esiste un effetto univoco del WFH sulla produttività. Piuttosto, l'impatto sembra essere altamente eterogeneo e condizionato da una serie di fattori, tra cui la tipologia di *job tasks*³ svolti, il grado di autonomia del lavoratore, la necessità di interazione e coordinamento, e le caratteristiche socio-demografiche degli individui. In questo senso, diversi studi recenti

³ Per *job tasks* si intendono le singole attività o mansioni che compongono un lavoro, le quali possono differire per contenuto, complessità e grado di ripetitività.

sottolineano che la produttività del lavoro da remoto è il risultato di un equilibrio tra benefici (riduzione dei costi di commuting⁴, maggiore concentrazione) e costi (isolamento, difficoltà di coordinamento, peggioramento della comunicazione informale) (Anakpo et al., 2023).

Nel complesso, la letteratura evidenzia dunque un quadro empirico non conclusivo, in cui il segno dell'effetto del WFH sulla produttività dipende dal contesto istituzionale, organizzativo e individuale. Questo rende particolarmente rilevante l'analisi empirica basata su dati microeconomici longitudinali, come quelli di *Understanding Society*, che consentono di controllare per eterogeneità individuale non osservata e di analizzare la variazione nel tempo delle modalità lavorative.

Un aspetto centrale della letteratura sul lavoro da remoto riguarda le difficoltà nell'identificazione causale dell'effetto del WFH sulla produttività. A differenza di contesti sperimentali controllati, nella maggior parte dei casi l'adozione del lavoro da remoto è una scelta endogena, influenzata sia da caratteristiche individuali sia da decisioni organizzative delle imprese. Questo genera problemi di selezione che rendono complessa l'interpretazione dei risultati empirici.

In primo luogo, un'importante fonte di distorsione è rappresentata dalla selezione dei lavoratori. Gli individui che scelgono (o sono selezionati per) il lavoro da remoto possono differire sistematicamente da quelli che lavorano in presenza lungo dimensioni non osservabili, come motivazione, capacità organizzativa o preferenze per l'autonomia. Questo implica che eventuali differenze di produttività osservate tra lavoratori remoti e non remoti potrebbero riflettere tali caratteristiche piuttosto che un effetto causale del WFH.

⁴ Per *commuting* si intende lo spostamento quotidiano e sistematico del lavoratore tra il luogo di residenza e il luogo di lavoro, che comporta un costo in termini di tempo e risorse.

Un contributo rilevante in questo ambito è fornito da Emanuel e Harrington (2024), che distinguono empiricamente tra effetti di trattamento ed effetti di selezione⁵ nel lavoro da remoto. Analizzando dati di call center statunitensi, gli autori mostrano che chi lavorava da casa presentava una produttività inferiore del 12% rispetto ai lavoratori in presenza nel periodo pre-pandemico, ma che una parte significativa di tale differenza è attribuibile alla selezione dei lavoratori meno produttivi per il lavoro da remoto, piuttosto che all'effetto diretto della modalità lavorativa.

Questo risultato evidenzia come la mancata correzione per la selezione possa portare a una sovrastima degli effetti negativi del WFH sulla produttività. Allo stesso tempo, la presenza di un effetto di trattamento negativo – seppur più contenuto – suggerisce che il lavoro da remoto può influenzare direttamente le performance, ad esempio attraverso una riduzione della supervisione o della qualità del coordinamento.

⁵ Per *effetto di trattamento* si intende la variazione causale dell'outcome (ad esempio la produttività) attribuibile all'assegnazione al trattamento, mentre per *effetto di selezione* si intende la componente della differenza osservata tra trattati e non trattati dovuta a caratteristiche preesistenti, che influenzano la probabilità di ricevere il trattamento.

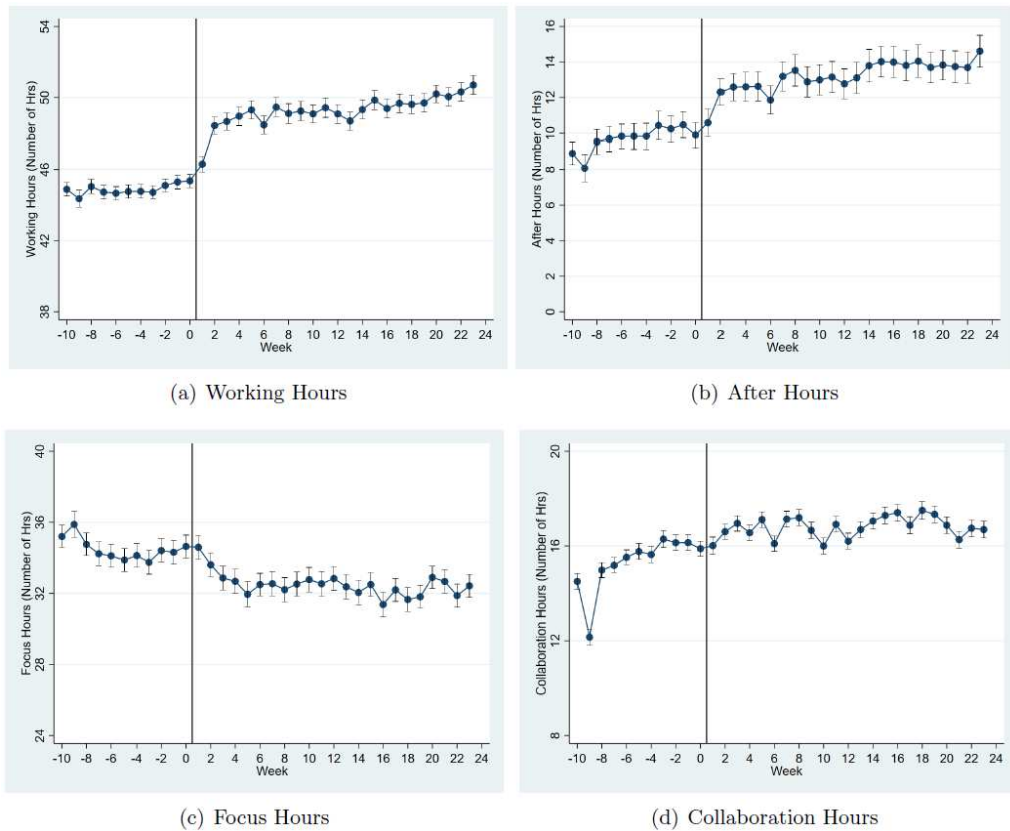


Figura 3 – Modalità di lavoro prima e dopo l'introduzione del lavoro da remoto. Pannello (a): ore medie lavorate settimanalmente. Pannello (b): ore lavorate al di fuori dell'orario di lavoro regolare. Pannello (c): numero di ore in blocchi di almeno due ore non dedicate a riunioni, chiamate o scrittura di e-mail. Pannello (d): ore trascorse in riunioni o in chiamate. La settimana 0 corrisponde al periodo 9–15 marzo 2020. Fonte: Gibbs et al. (2023).

Un secondo problema metodologico riguarda la possibilità di *reverse causality*, ovvero il fatto che la produttività stessa possa influenzare la probabilità di lavorare da remoto. Ad esempio, lavoratori più produttivi o più qualificati potrebbero avere maggiore accesso a opportunità di lavoro flessibile, oppure imprese più efficienti potrebbero essere più propense ad adottare modelli organizzativi basati sul WFH. Questo rende difficile distinguere se il lavoro da remoto aumenti la produttività o se, al contrario, siano i lavoratori più produttivi a lavorare da remoto.

Inoltre, la letteratura rileva la presenza di meccanismi di selezione dinamica nel tempo. Evidenze a livello di impresa mostrano che, nel periodo successivo alla pandemia, le imprese con livelli di produttività più bassi nel lavoro da remoto tendono ad abbandonare tale modalità, mentre quelle più efficienti continuano a utilizzarla. Questo processo di selezione porta a un miglioramento della produttività media osservata tra le imprese che adottano il WFH, ma non necessariamente riflette un effetto causale positivo del lavoro da remoto (Morikawa, 2024).

Un ulteriore elemento di complessità deriva dalla misurazione della produttività nei contesti di lavoro da remoto. Nei dati survey, come quelli utilizzati in molti studi empirici, la produttività è spesso misurata tramite indicatori soggettivi o proxy indiretti, introducendo potenziali errori di misurazione. Anche nei dati amministrativi, la produttività può essere difficile da osservare, soprattutto per lavori ad alta componente cognitiva o collaborativa. Tali problemi possono generare ulteriori distorsioni nelle stime econometriche, rendendo necessario l'uso di metodologie robuste.

Per affrontare queste criticità, la letteratura ha adottato diverse strategie di identificazione. Tra queste, l'utilizzo di esperimenti naturali – come la pandemia di COVID-19 – ha consentito di sfruttare variazioni esogene nelle modalità lavorative. Tuttavia, anche in questo caso, la natura non randomizzata dell'evento e le differenze nella capacità di adattamento tra individui e imprese richiedono cautela nell'interpretazione dei risultati.

In questo contesto, l'uso di dati panel longitudinali rappresenta una delle strategie più efficaci per mitigare i problemi di identificazione. L'inclusione di effetti fissi individuali consente di controllare per caratteristiche non osservate costanti nel tempo, mentre gli effetti temporali permettono di catturare shock comuni, come quelli legati alla pandemia. Questo approccio, adottato nel presente lavoro, permette di isolare la variazione intra-individuale nell'adozione del WFH e di

stimarne l'effetto sulla produttività in modo più credibile rispetto a semplici confronti cross-section⁶.

Un risultato ormai consolidato nella letteratura empirica sul lavoro da remoto è che l'impatto del WFH sulla produttività non è uniforme, ma presenta un'elevata eterogeneità tra individui, imprese e contesti istituzionali. Tale eterogeneità rappresenta un elemento centrale per comprendere le differenze nei risultati empirici documentati nei diversi studi.

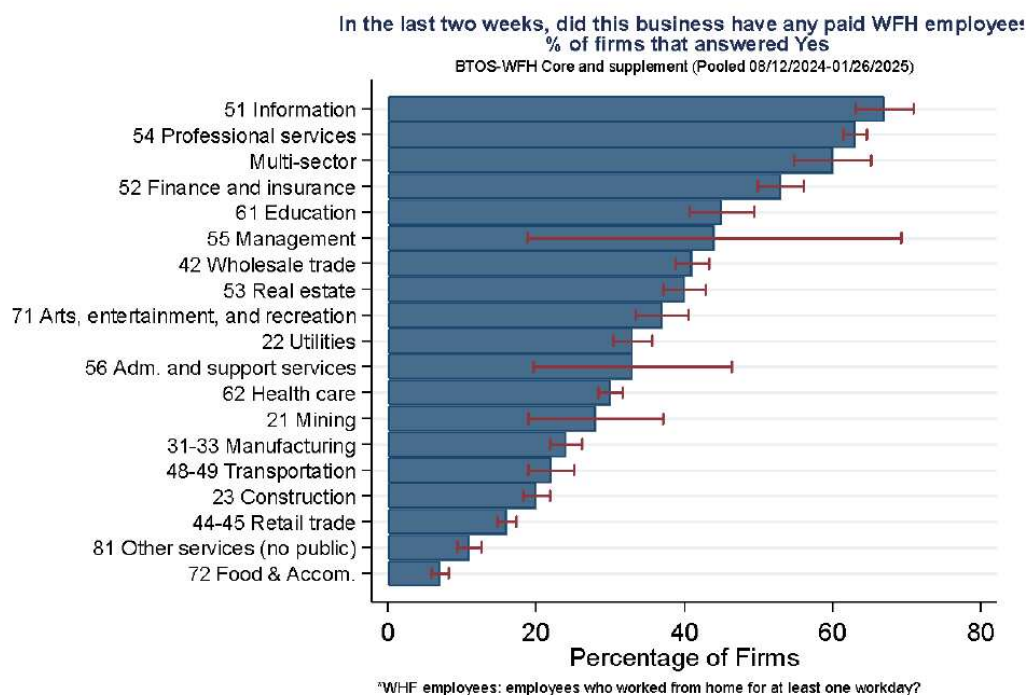
In primo luogo, numerosi contributi evidenziano che la diffusione e l'intensità del lavoro da remoto variano significativamente tra paesi e categorie di lavoratori. Utilizzando dati internazionali comparabili, Aksoy et al. (2022) mostrano che il lavoro da remoto si attesta mediamente intorno a 1,5 giorni a settimana, ma con ampie variazioni tra paesi e gruppi socio-demografici. In particolare, fattori quali livello di istruzione, presenza di figli e durata del commuting influenzano significativamente sia la probabilità di lavorare da casa sia le preferenze per tale modalità lavorativa. Questi risultati suggeriscono che l'adozione del WFH non è determinata esclusivamente da vincoli tecnologici, ma riflette anche caratteristiche individuali e preferenze eterogenee.

Un secondo elemento rilevante riguarda le differenze tra settori e occupazioni. Evidenze recenti indicano che la possibilità di svolgere il lavoro da remoto dipende fortemente dalla natura delle mansioni e dal grado di digitalizzazione delle attività. In particolare, il WFH risulta più diffuso nei settori ad alta intensità di capitale umano e nei lavori caratterizzati da task cognitivi e non routinari, mentre è molto meno presente nei settori che richiedono interazione fisica o presenza sul luogo

⁶ Per *cross-section* si intende un insieme di dati osservati in un singolo momento temporale per diverse unità statistiche (ad esempio individui o imprese), che consente di analizzare le differenze tra unità ma non la loro evoluzione nel tempo.

di lavoro (Zarate et al., 2024). Questo implica che gli effetti del WFH sulla produttività possono variare sistematicamente in funzione della composizione occupazionale.

Figure 2: Extensive Margin Sectoral



Note: 95% Confidence Intervals are represented by the horizontal red lines.

Figura 4 – Diffusione del lavoro da remoto per settore (margine estensivo). La figura mostra la percentuale di imprese che impiegano lavoratori in modalità di lavoro da casa nei diversi settori, evidenziando forti differenze settoriali: mentre nel settore dell'informazione circa il 67% delle imprese utilizza il WFH, nel settore della ristorazione e dell'alloggio tale quota si attesta intorno al 7%. Fonte: Barrero et al. (2025).

Ulteriori evidenze sottolineano come l'eterogeneità si manifesti anche all'interno delle imprese. Analizzando dati recenti a livello aziendale, Barrero et al. (2025) mostrano che l'intensità del lavoro da remoto varia considerevolmente tra settori e organizzazioni, con valori medi di circa un giorno a settimana ma con picchi molto più elevati nei settori informatici e quasi nulli nei servizi a bassa intensità tecnologica. Questo suggerisce

che le caratteristiche organizzative e manageriali delle imprese svolgono un ruolo cruciale nel determinare l'adozione e l'efficacia del WFH.

Un ulteriore aspetto riguarda le differenze tra individui all'interno dello stesso contesto lavorativo. Studi basati su dati microeconomici mostrano che l'effetto del lavoro da remoto sulla produttività può variare significativamente tra lavoratori, anche a parità di mansione. Ad esempio, analizzando dati su sviluppatori di software, Bao et al. (2022) evidenziano che il WFH ha effetti sia positivi sia negativi sulla produttività a seconda delle caratteristiche del progetto e dell'individuo, confermando la presenza di una forte eterogeneità nei risultati. In particolare, progetti più complessi o che richiedono elevato coordinamento tendono a essere penalizzati dal lavoro da remoto, mentre attività più autonome possono beneficiarne.

La letteratura più recente evidenzia inoltre che l'eterogeneità degli effetti del WFH è anche dinamica nel tempo. Evidenze empiriche suggeriscono che l'impatto del lavoro da remoto può cambiare con l'adattamento progressivo di lavoratori e imprese alle nuove modalità organizzative. In questo senso, il WFH non rappresenta semplicemente uno shock temporaneo, ma un cambiamento strutturale che evolve nel tempo, influenzato da investimenti in capitale umano e tecnologico e da processi di apprendimento organizzativo (Barrero et al., 2021).

Questi risultati suggeriscono che l'effetto del lavoro da remoto sulla produttività non può essere analizzato in termini medi senza considerare la distribuzione degli effetti. L'eterogeneità osservata implica che il WFH può generare benefici per alcuni lavoratori e contesti, ma costi per altri, rendendo fondamentale un approccio empirico che tenga conto di tali differenze. In questo contesto, l'utilizzo di dati panel come quelli di *Understanding Society* consente di esplorare in modo più approfondito le

fonti di eterogeneità, analizzando come l'impatto del lavoro da remoto vari in funzione delle caratteristiche individuali, familiari e lavorative.

Un filone più recente della letteratura si concentra su aspetti meno esplorati del lavoro da remoto, in particolare sulle implicazioni di lungo periodo per le dinamiche del mercato del lavoro e per i percorsi di carriera. In questo contesto, evidenze empiriche mostrano che il lavoro da remoto può influenzare non solo la produttività corrente, ma anche esiti dinamici quali promozioni, apprendimento e accumulazione di capitale umano. In particolare, l'analisi di dati microeconomici suggerisce che i lavoratori in modalità remota sperimentano una minore probabilità di avanzamento di carriera, probabilmente a causa di minori interazioni informali e di una ridotta visibilità nei confronti dei superiori (Emanuel e Harrington, 2024).

Un ulteriore aspetto emergente riguarda il ruolo delle frizioni informative all'interno delle organizzazioni. La letteratura evidenzia che il lavoro in presenza facilita la trasmissione di informazioni tacite e il monitoraggio informale delle performance, mentre il lavoro da remoto può ridurre la qualità delle informazioni disponibili ai manager, con conseguenze sia sulla valutazione dei lavoratori sia sull'allocazione delle risorse all'interno dell'impresa. Questo meccanismo contribuisce a spiegare perché gli effetti del WFH possano estendersi oltre la produttività individuale, influenzando l'efficienza complessiva delle organizzazioni.

Recenti contributi sottolineano inoltre come il lavoro da remoto possa generare implicazioni rilevanti anche in termini di equilibrio di mercato. In particolare, modelli teorici ed evidenze empiriche suggeriscono che l'offerta di lavoro da remoto possa essere soggetta a problemi di coordinamento tra imprese, configurando situazioni analoghe a un dilemma del prigioniero: se tutte le imprese limitassero l'offerta di lavoro da remoto potrebbero beneficiare di una forza lavoro mediamente più produttiva, ma ciascuna impresa individualmente ha incentivo a offrire

maggior flessibilità per attrarre lavoratori, anche a costo di una minore produttività media.

Un ulteriore contributo della letteratura riguarda gli effetti del WFH sulla qualità del lavoro e sulla composizione delle attività svolte. Evidenze recenti mostrano che il lavoro da remoto può modificare non solo la quantità, ma anche la natura del lavoro svolto, ad esempio riducendo la qualità delle interazioni tra colleghi o aumentando la frammentazione delle attività lavorative. In particolare, alcuni studi documentano un peggioramento della qualità dell'output e una riduzione delle opportunità di apprendimento informale, soprattutto per i lavoratori meno esperti (Emanuel e Harrington, 2024).

Alla luce di questi contributi, emerge quindi come il lavoro da remoto non possa essere interpretato unicamente come una diversa modalità di organizzazione del lavoro, ma rappresenti un cambiamento strutturale che incide su molteplici dimensioni del funzionamento del mercato del lavoro e delle imprese. In particolare, gli effetti su produttività, carriere e organizzazione interna risultano strettamente interconnessi e richiedono un'analisi integrata.

In questo contesto, il presente lavoro si propone di contribuire alla letteratura esistente focalizzandosi su un aspetto ancora relativamente poco esplorato, ovvero l'interazione tra lavoro da remoto e caratteristiche individuali e familiari all'interno di un dataset rappresentativo. L'utilizzo dei dati longitudinali di *Understanding Society* consente infatti di analizzare come l'impatto del WFH sulla produttività si inserisca nel più ampio contesto delle condizioni di vita degli individui, offrendo una prospettiva complementare rispetto agli studi basati su dati aziendali.

2. Dati e strategia empirica

Per l'analisi empirica sono stati utilizzati i file *Individual Respondent (indresp)* di *Understanding Society: the UK Household Longitudinal Study* (UKHLS), che rappresentano il principale dataset individuale dell'indagine e contiene le informazioni raccolte annualmente presso ciascun rispondente adulto. Il dataset *indresp* include un ampio insieme di variabili relative alle caratteristiche socio-demografiche degli individui, alla loro condizione occupazionale, ai livelli di reddito, alle caratteristiche dell'impiego e alle condizioni familiari, rendendolo particolarmente adatto allo studio delle trasformazioni del lavoro e degli effetti del lavoro da remoto sulle performance lavorative individuali. La struttura longitudinale del dataset permette inoltre di osservare gli stessi individui lungo più *waves*⁷ dell'indagine, consentendo la costruzione di un panel individuale utilizzabile per analisi dinamiche e modelli econometrici con effetti fissi (UK Longitudinal Linkage Collaboration, n.d.).

Il dataset costituisce uno dei nuclei principali dell'architettura dati di *Understanding Society* ed è organizzato in modo coerente tra le diverse *waves* dell'indagine, attraverso un sistema di prefissi alfabetici che identifica temporalmente ciascuna rilevazione. Ogni *wave* o set di interviste copre un periodo di circa due anni solari, a partire dalla prima (*a_indresp*), che va dall'8 gennaio 2009 al 10 marzo 2011, fino alla quindicesima (*o_indresp*), che va dal 12 gennaio 2023 al 20 maggio 2025 (Understanding Society, n.d.). Tale struttura consente di armonizzare le informazioni raccolte nel tempo e di seguire longitudinalmente gli individui attraverso un identificativo personale univoco (*pidp*), mantenuto costante nelle diverse *waves*. La disponibilità di identificativi longitudinali

⁷ Singola "*ondata*" di raccolta dati all'interno di un'indagine longitudinale, in cui gli stessi individui vengono intervistati in un determinato periodo temporale, generalmente con cadenza regolare (ad esempio annuale).

rappresenta un elemento fondamentale per la costruzione di dataset panel e per l'implementazione di strategie empiriche che richiedono il monitoraggio delle traiettorie individuali nel tempo. In particolare, l'utilizzo di dati panel consente di controllare per caratteristiche individuali non osservabili ma costanti nel tempo, riducendo il rischio di distorsione da variabile omessa (*omitted variable bias*⁸) nelle stime econometriche (UK LLC, n.d.).

Dal punto di vista del contenuto, il file indresp include un numero molto elevato di variabili relative sia alla sfera lavorativa sia alle condizioni socioeconomiche degli individui. Le sezioni dedicate all'occupazione raccolgono informazioni dettagliate sullo status lavorativo, sulle caratteristiche dell'impiego principale, sul settore economico di appartenenza, sulla tipologia contrattuale, sulle ore lavorate e sui livelli retributivi. A queste si aggiungono variabili relative al background educativo, alla composizione familiare, alle condizioni abitative e alla salute, che rappresentano importanti determinanti delle modalità di partecipazione al mercato del lavoro e delle opportunità di accesso al lavoro flessibile o da remoto. La ricchezza informativa del dataset permette pertanto di costruire un insieme articolato di variabili di controllo, necessarie per l'implementazione di modelli empirici robusti (UK LLC, n.d.).

2.1. Definizione e costruzione delle variabili

L'implementazione della strategia empirica ha richiesto un processo di estrazione, pulizia e armonizzazione dei microdati individuali estratti dalle wave dalla *a* alla *o* di indresp. Al fine di garantire la trasparenza

⁸ Distorsione sistematica nella stima di un parametro, dovuta alla presenza di errori nel modello o nei dati, che porta i risultati a discostarsi dal valore reale della relazione analizzata.

metodologica, questa sezione esamina analiticamente i criteri di selezione del campione e il processo di trasformazione dei dati.

Le variabili inserite nei modelli econometrici si dividono in due categorie quantitative: da un lato, indicatori strutturali che presentavano già un livello di codifica idoneo all'applicazione diretta (salvo per la necessaria rimozione dei codici mancanti o non applicabili); dall'altro, variabili complesse generate ex-novo o modificate a partire dai dati originari, in modo da poter rispondere coerentemente al quesito di ricerca (Understanding Society, n.d.).

2.1.1. Criteri di campionamento e selezione del sottocampione

Il dataset indresp originale raccoglie le interviste di tutti i membri adulti del nucleo familiare che abbiano compiuto almeno 16 anni. Tuttavia, per rispondere coerentemente al quesito di ricerca, si è reso necessario circoscrivere l'analisi attraverso due filtri:

1. *Limitazione anagrafica:* Il campione è stato ristretto esclusivamente agli individui di età compresa tra i 20 e i 64 anni compresi (*keep if age >= 20 & age <= 64*). Tale scelta è motivata dalla necessità di escludere le fasi di transizione iniziale nel mercato del lavoro (spesso caratterizzate da lavori part-time o stage non indicativi della reale produttività oraria) e gli anni che precedono direttamente il pensionamento, in cui i ritmi di lavoro dell'individuo potrebbero cambiare.
2. *Esclusione del lavoro autonomo:* Utilizzando la variabile sullo status occupazionale corrente dell'intervistato, definita nell'indagine come *jbstat* (*job status*), il campione è stato limitato ai soli lavoratori dipendenti. La variabile originale definisce formalmente la condizione occupazionale dell'intervistato nella

settimana precedente l'intervista, classificandola in diverse categorie tra cui *Self employed* (valore 1) e *Paid employment (full-time/part-time)* (valore 2) (Understanding Society, n.d.). Attraverso la generazione della variabile dicotomica *selfemp*, sono stati mantenuti nell'analisi solo gli individui per i quali *jbstat* == 2, escludendo categoricamente i lavoratori autonomi (*keep if selfemp* == 0). Questo passaggio è essenziale sotto il profilo econometrico: la produttività dei lavoratori autonomi obbedisce a logiche di determinazione del reddito e di elasticità delle ore lavorate strutturalmente diverse da quelle dei lavoratori dipendenti, rendendo i due gruppi non comparabili in un modello panel.

2.1.2. La variabile dipendente: costruzione della proxy di produttività oraria

La misura della produttività individuale costituisce una delle sfide metodologiche principali nell'utilizzo di survey basate su questionari familiari, data l'assenza di dati certi sul valore aggiunto per ora lavorata: anche qualora fossero presenti, non potrebbero che essere basati su percezioni puramente soggettive. In linea con la consolidata letteratura in economia del lavoro, si adotta come proxy della produttività oraria individuale il *salario orario effettivo*, calcolato come il rapporto tra la remunerazione settimanale e le ore normalmente lavorate in una settimana (Raitano, 2021).

La costruzione di questa variabile ha richiesto la combinazione di due distinti indicatori:

- *Ore lavorate (hours_worked)*: derivata dalla variabile *jbhrs* (*Number of hours normally worked per week*). Nella documentazione ufficiale di *Understanding Society*, questa

variabile è definita come il numero di ore che il rispondente dichiara di dedicare normalmente all'impiego principale in una settimana tipo, escludendo gli straordinari e le pause pranzo (Understanding Society, n.d.). Nel do-file, per eliminare anomalie o risposte non plausibili, sono state considerate valide solo le osservazioni con un numero di ore positivo e inferiore a 100 settimanali (*replace hours_worked = . if hours_worked <= 0 | hours_worked > 100*).

- *Reddito settimanale (weekly_pay)*: derivata dalla variabile *paygu_dv* (*gross monthly pay – derived*), si tratta di una variabile derivata (*derived variable*, come segnalato dal suffisso *_dv*) che standardizza i guadagni mensili lordi del lavoratore, correggendo eventuali discrepanze nei periodi di paga dichiarati e applicando tecniche di imputazione nei casi di mancata risposta (Understanding Society, n.d.)⁹. Al fine di ricondurre tale informazione alla stessa unità temporale delle ore lavorate, il reddito mensile è stato diviso per il coefficiente di conversione standard di 4,345 settimane al mese (*gen weekly_pay = `X'paygu_dv / 4.345*). Anche in questo caso, i valori nulli o negativi sono stati trattati come dati mancanti per evitare distorsioni nella stima.

Il rapporto tra queste due componenti ha permesso di generare la variabile *productivity_hour* (*gen productivity_hour = weekly_pay / hours_worked*). Nei modelli econometrici, la variabile dipendente viene inserita tramite la sua trasformazione in logaritmo naturale, *log_productivity_hour* (*gen log_productivity_hour =*

⁹ Il termine “derivata” indica che la retribuzione mensile è armonizzata matematicamente a partire dalle risposte individuali (es. su base settimanale o annuale), mentre “imputazione” fa riferimento alle tecniche statistiche utilizzate dall'ente per stimare i valori mancanti (*missing values*) sulla base delle caratteristiche socio-demografiche dell'intervistato.

$\log(\text{productivity_hour})$). L'utilizzo del logaritmo naturale risponde a una doppia esigenza pratica: da un lato, rende la distribuzione dei salari più simmetrica, correggendo la presenza di valori eccessivamente alti che comprometterebbero la validità statistica del modello. Dall'altro, facilita l'interpretazione dei risultati: i coefficienti del modello non indicheranno una variazione in termini monetari assoluti, ma esprimeranno direttamente l'impatto dei fattori considerati in termini di variazione percentuale sulla produttività.

2.2. Le variabili esplicative: margine intensivo ed estensivo del WFH

L'analisi dell'effetto dello smart working sulla produttività richiede di isolare sia l'impatto della scelta di lavorare o meno da remoto (accesso al WFH), sia l'impatto della frequenza con cui tale modalità viene esercitata (intensità del ricorso al WFH). Una complicazione tecnica rilevante nell'utilizzo delle *waves* dall'inizio dell'indagine (2009) fino al periodo recente sta nel fatto che la struttura delle domande sul lavoro da casa ha subito alcune modifiche, specialmente a seguito dello shock pandemico.

Per ovviare a ciò, nel do-file è stato implementato un comando di cattura e armonizzazione *cross-wave (capture confirm variable)*¹⁰ che attinge a tre diverse variabili originali (Understanding Society, n.d.):

- *wkhome (working at/from home frequency)*: variabile presente nelle *waves* 12 e 14, generata a partire dalla domanda: “*Quanto spesso lavori a/da casa?*” e che può assumere valori corrispondenti alla frequenza del ricorso al WFH: “mai”, “1-4 giorni al mese”, “1, 2, 3 o 4 giorni a settimana”, “sempre”, “variabile”.

¹⁰ La “cattura” identifica il recupero automatico in Stata di variabili che cambiano nome tra le diverse *waves* o sono presenti solo in alcune di esse, evitandone il blocco del codice; l’“armonizzazione” indica la loro successiva ricodifica in un’unica scala omogenea e confrontabile nel tempo.

- *jbwah (working at home)*: variabile categoriale presente nelle *waves* 13 e 14 che rileva la frequenza complessiva del lavoro da remoto nelle *waves* più recenti. Risponde alla stessa domanda di *wkhome*: “Quanto frequentemente lavori da casa?”, ma assume valori diversi e meno puntuali: “sempre”, “spesso”, “a volte”, “mai”.
- *jbwahnew (working at home)*: variabile presente solo nella *wave* 15, introdotta per mappare dettagliatamente la frequenza post-pandemica in termini di giorni precisi al mese o alla settimana. Può assumere gli stessi valori di *wkhome*: “mai”, “1-4 giorni al mese”, “1, 2, 3 o 4 giorni a settimana”, “sempre”, “variabile”.

Al fine di garantire la massima precisione ed evitare distorsioni dovuti a errori di misurazione, si è reso necessario valutare ed escludere dal processo di armonizzazione altri due indicatori attinenti al tema, ma metodologicamente inadeguati: *jbflex7* e *wkhomewil*.

2.2.1. Il ruolo e l'esclusione di *jbflex7*

La variabile *jbflex7 (flexible working arrangements: To work from home on a regular basis)* è presente nel panel dalla *wave* 2 alla 14, comparando ogni due *waves*, e rileva la disponibilità formale di accordi per il lavoro da casa su base regolare (Understanding Society, n.d.). Sotto il profilo informativo, questa variabile ha un valore teorico significativo: indica *l'eleggibilità formale* del lavoratore, ovvero se l'azienda di appartenenza offra o meno la flessibilità contrattuale a prescindere dal suo effettivo utilizzo. Inoltre, questo indicatore è presente in molte più *waves* rispetto ai tre già esaminati, consentendo di avere un quadro storico più ampio in merito all'accesso al WFH, seppur limitatamente alla sola dimensione della possibilità di farne uso e non all'utilizzo effettivo.

Tuttavia, per gli scopi di questo studio, *jbflex7* è stata volutamente esclusa dall'analisi econometrica: poiché l'obiettivo è stimare l'effetto del WFH sulla *produttività individuale*, è la pratica *effettiva* (i giorni reali passati a casa) e non il semplice accesso contrattuale teorico a impattare sui canali di efficienza del lavoratore, come l'abbattimento dei tempi di pendolarismo e la gestione del tempo. L'inclusione di *jbflex7* nel calcolo della frequenza avrebbe confuso l'offerta di flessibilità aziendale con il comportamento reale del lavoratore, generando stime inefficienti dell'impatto sulla produttività.

Ad ogni modo, per completezza, è utile riassumere le principali evidenze che è possibile ricavare dall'analisi di questa variabile nel corso delle diverse *waves*. I dati descrittivi estratti dal panel rivelano un trend di crescita costante nella concessione teorica del lavoro da casa da parte delle aziende britanniche, accelerato ulteriormente nel periodo post-pandemico.

Considerando la percentuale di rispondenti che dichiarano la presenza di accordi formali di WFH (*yes mentioned*), si osserva la seguente evoluzione temporale:

- *Il periodo pre-pandemico (Wave 2 – 10):* nella *wave 2* (2010-2012), i lavoratori con accesso formale al WFH rappresentavano appena il 5,67% del campione totale. Questa quota è rimasta sostanzialmente stabile per quasi un decennio, oscillando tra il 6,31% della *wave 4* e il 7,45% della *wave 10*.
- *La transizione e l'effetto post-Covid (Wave 12 – 14):* il vero cambio di passo strutturale si registra a partire dalla *wave 12* (2020-2022), dove la percentuale sale al 10,36%, fino a raggiungere il picco dell'11,61% nella *wave 14*.

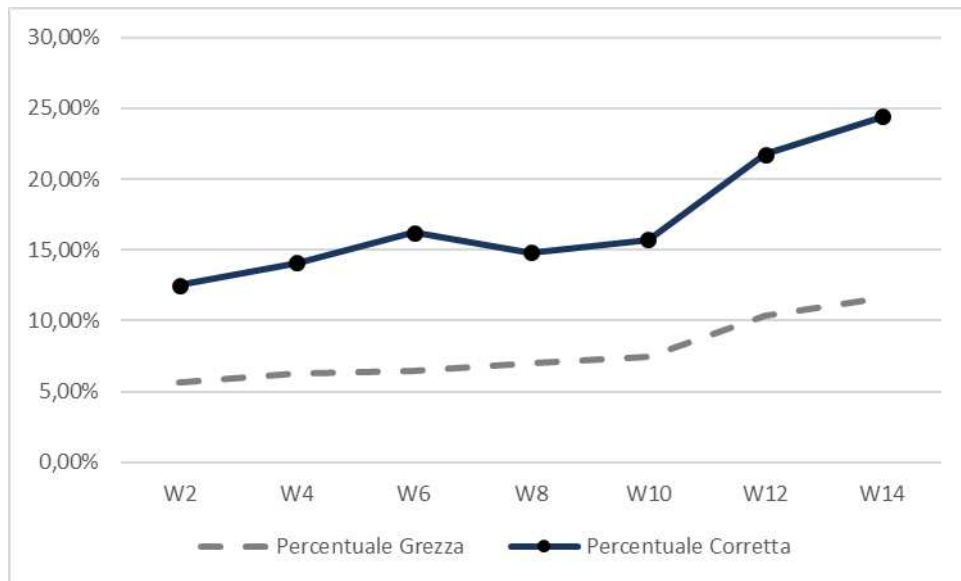


Figura 5 – Evoluzione storica dell’eleggibilità formale al WFH (Wave 2 - Wave 14).

Confronto tra dati grezzi e dati corretti sul campione attivo. Fonte: elaborazione autonoma su dati Understanding Society.

Un’analisi più approfondita, condotta al netto dei dati non applicabili (*inapplicable*) e delle interviste per interposta persona (*proxy*), evidenzia un raddoppio del fenomeno: se nella seconda *wave* solo il 12,51% dei lavoratori eleggibili dichiarava di avere a disposizione contrattuale il lavoro da remoto, nella quattordicesima *quasi il 25% del campione attivo e rispondente* segnala la presenza di questa opzione flessibile all’interno dei propri accordi aziendali.

Questi dati confermano che, sebbene *jbflex7* descriva unicamente una dimensione potenziale e non l’effettivo comportamento sul posto di lavoro, essa documenta una profonda trasformazione istituzionale e contrattuale nel mercato del lavoro del Regno Unito.

2.2.2. L’esclusione di *wkhomewil*

La variabile *wkhomewil* (*willingness to work from home*) mira a quantificare la propensione psicologica o il desiderio del rispondente di lavorare da remoto (Understanding Society, n.d.). Questo indicatore è stato

escluso per via della forte distorsione campionaria che avrebbe comportato. La documentazione ufficiale mostra infatti che la variabile è presente esclusivamente nella *wave 2* con la dicitura *b_wkhomewil*, oltretutto presentando una quota di valori non applicabili (*inapplicable*) pari al *90,73% delle osservazioni totali della wave* (49.510 casi non applicabili su un totale di 54.568).

Un'omissione di dati così massiccia, unita alla natura puramente *cross-sectional* (essendo limitata a un unico punto nel tempo), avrebbe drasticamente ridotto i gradi di libertà dei modelli panel, impedendo la stima accurata degli effetti fissi individuali e l'osservazione dei trend evolutivi nel corso degli anni.

2.2.3. Costruzione delle variabili armonizzate finali

L'armonizzazione effettuata sulle tre variabili principali (*wkhome*, *jbwahnew*, *jbwah*) ha permesso di organizzare l'analisi su due dimensioni distinte:

- *Margine intensivo (wfh_freq)*: è la variabile categoriale ordinata finale, standardizzata su una scala a 7 punti (1 = *Never*; 2 = *1-4 days/month*; 3 = *1 day/week*; 4 = *2 days/week*; 5 = *3 days/week*; 6 = *4 days/week*; 7 = *Always*). I valori non validi, identificati nei codici d'indagine con valori negativi relativi a mancata risposta, rifiuto o *inapplicable*, sono stati convertiti in dati mancanti (.). Questa variabile cattura la granularità del fenomeno, consentendo di verificare se l'effetto sulla produttività segua un trend lineare o se alcune configurazioni di WFH spicchino rispetto ad altre.
- *Margine estensivo (wfh_dummy)*: è una variabile indicatrice dicotomica costruita direttamente a partire dalla precedente (*gen wfh_dummy = 0 if wfh_freq == 1, e pari a 1 se wfh_freq >= 2 & wfh_freq <= 7*). Questa *dummy*

discrimina nettamente tra la popolazione di lavoratori che non usufruiscono in alcun modo del lavoro da casa (gruppo di controllo) e coloro che vi fanno ricorso con una frequenza qualunque (gruppo di trattamento), offrendo una misura aggregata del differenziale di produttività associato al ricorso o meno al WFH.

2.3. Variabili di controllo

2.3.1. Età (*dvage*)

Il controllo per l'età dei rispondenti all'interno del modello risponde a una duplice esigenza: in primo luogo, l'età è una proxy diretta dell'*esperienza lavorativa* e del capitale umano accumulato sul campo. Lavoratori più anziani possiedono generalmente competenze consolidate e una maggiore autonomia gestionale, fattori che influenzano positivamente la produttività individuale indipendentemente dal luogo in cui la prestazione viene fornita.

In secondo luogo, è ragionevole pensare che l'età condizioni sia la preferenza soggettiva per lo smart working, sia la capacità di adattamento tecnologico. Da questo punto di vista c'è probabilmente una differenza generazionale: i lavoratori più giovani (i cosiddetti nativi digitali) potrebbero mostrare una maggiore efficienza nell'utilizzo dei software per il lavoro da remoto e, in generale, nell'uso del computer. Al contempo, però, i lavoratori nelle prime fasi della carriera traggono un beneficio immenso dalle interazioni faccia a faccia in ufficio per via dell'apprendimento informale e del *networking*. Al contrario, i lavoratori in età avanzata beneficiano maggiormente del WFH in termini di bilanciamento vita-lavoro, ma potrebbero risentirne nel caso in cui avessero scarse capacità informatiche.

Non includere l'età genererebbe una distorsione da variabile omessa (*omitted variable bias*), poiché la struttura anagrafica è correlata sia con la

probabilità di vedersi concesso il lavoro da casa (spesso legato a ruoli di anzianità aziendale), sia con i livelli di produttività espressi.

Nel panel *Understanding Society*, l'età del rispondente viene tracciata attraverso la variabile derivata *dvage* (presente nei file *indresp* con il rispettivo prefisso di *wave*, da *a_dvage* per la *wave* 1 a *n_dvage* per la *wave* 14). Dal punto di vista della costruzione del dato, come riportato nel database dell'indagine, *dvage* esprime l'età del rispondente al momento dell'ultimo compleanno, calcolata algebricamente incrociando la data esatta dell'intervista (la compilazione dell'unità familiare) con la data di nascita dell'individuo (Understanding Society, n.d.).

Un elemento metodologico di grande rilievo per la tesi, per quanto banale possa sembrare, risiede nella stabilità del dato lungo le *waves*. Come documentato dai registri del dataset, l'istituto di ricerca effettua un controllo incrociato *cross-wave* (particolarmente robusto a partire dalla *wave* 6) per sanare i valori mancanti. Se l'informazione sulla data di nascita non è disponibile nell'intervista corrente, l'algoritmo attinge al file storico complessivo del panel (*INDALL*) o alle dichiarazioni rilasciate nelle *waves* precedenti dagli altri membri dello stesso nucleo familiare. Solo in casi residuali si ricorre a una stima approssimativa o all'età dichiarata originariamente (*AGEIF*).

L'analisi descrittiva delle diverse *waves* mostra un campione con un'età minima stabilmente agganciata a *16 anni* (soglia d'ingresso nel panel per le sezioni adulti) e massimi che superano i *100 anni*, con una media complessiva che si sposta progressivamente da 45.6 anni nella *wave* 1 a 50.9 anni nella *wave* 13 e 51.4 nella *wave* 15, riflettendo il fisiologico invecchiamento dei partecipanti all'indagine.

Nel modello econometrico qui elaborato, per non vincolare l'effetto dell'età a una forma strettamente lineare (che presupporrebbe che ogni

anno in più spostati la produttività sempre della stessa quota), la variabile viene trattata introducendo un *termine quadratico* (age^2), consentendo così di catturare la classica relazione a “U rovesciata” tra ciclo di vita, performance lavorativa e adattamento alle modalità flessibili di impiego.

2.3.2. Genere (*sex*)

L’inserimento del genere come variabile di controllo è fondamentale, poiché l’accesso allo smart working e il suo impatto sull’efficienza lavorativa non si distribuiscono in modo uguale tra uomini e donne.

Molte ricerche evidenziano come le donne tendano a richiedere il lavoro da casa principalmente come strumento di conciliazione per gestire il carico della cura familiare e del lavoro domestico non retribuito (Angelici e Profeta, 2020). Questa sovrapposizione tra spazio professionale e impegni personali può influenzare la performance: alcune evidenze mostrano che le lavoratrici in smart working possono fronteggiare un maggior numero di interruzioni dovute alle responsabilità domestiche, con un impatto asimmetrico sulla produttività oraria rispetto ai colleghi uomini (Gibbs et al., 2023). Al contempo, la letteratura mostra che lo smart working può ridurre i costi di trasferimento e aumentare la soddisfazione generale, ma espone a dinamiche diverse a seconda della composizione del nucleo familiare.

Al contrario, la componente maschile della popolazione aziendale tende a utilizzare il lavoro da remoto in contesti di maggiore autonomia gestionale o per isolarsi su compiti complessi ad alto valore aggiunto. Inoltre, lavorare da casa riduce la visibilità in ufficio e questo penalizza soprattutto le donne, dato che sono loro a usare di più lo smart working per motivi familiari e quindi finiscono per trascorrere meno tempo davanti ai manager. Questa minore presenza fisica fa sì che vengano spesso

scavalcate al momento delle promozioni, anche se producono quanto i colleghi in presenza (Bloom et al., 2015).

Inserire il genere nel modello evita quindi una distorsione da variabile omessa, impedendo che l'effetto stimato del WFH sulla produttività sia alterato da queste differenze di base nei ruoli familiari e di carriera già presenti nel Regno Unito.

All'interno di *Understanding Society*, la rilevazione del genere è affidata alla variabile *sex*, inclusa nei file individuali dei rispondenti (*indresp*) e rintracciabile lungo le diverse *waves* mediante l'applicazione dei prefissi sequenziali standard (da *a_sex* per la prima *wave*, fino a *n_sex* per la quattordicesima).

Poiché il genere è una caratteristica che non cambia nel tempo, i ricercatori di *Understanding Society* riescono a mantenere il dato estremamente pulito. Se in una *wave* un rispondente salta questa domanda, il sistema recupera in automatico l'informazione guardando alle interviste degli anni precedenti o ai dati del suo nucleo familiare. Grazie a questo controllo incrociato, i dati mancanti sono praticamente inesistenti: i rifiuti o le mancate risposte pesano pochissimo, oscillando tra lo 0% e lo 0,03% a seconda della *wave*.

La distribuzione statistica evidenzia una presenza femminile prevalente all'interno del panel in tutte le rilevazioni. Nella *wave* 1, su circa 51.000 rispondenti, le donne rappresentano il 54,5% del campione (*female* = 2), a fronte di una quota maschile del 45,5% (*male* = 1). Tale rapporto rimane pressoché stabile lungo l'intero orizzonte temporale, assestandosi nella *wave* 15 su valori pari al 55,62% per le donne e al 44,38% per gli uomini.

Nel modello econometrico, questa informazione viene formalizzata mediante la costruzione di una variabile *dummy* di genere, ponendo il

gruppo maschile come categoria di riferimento (pari a 0) e assegnando il valore 1 alle risposte delle donne intervistate.

2.3.3. Livello di istruzione (*hiqual_dv*)

All'interno del modello econometrico, il livello di istruzione rappresenta una variabile di controllo fondamentale per isolare il vero impatto dello smart working sulla produttività. Il capitale umano, infatti, influenza simultaneamente sia la probabilità di accedere al lavoro da casa, sia l'efficienza nello svolgimento delle proprie mansioni.

La letteratura economica internazionale evidenzia che la possibilità di lavorare da remoto è fortemente sbilanciata a favore dei lavoratori altamente qualificati (Dingel e Neiman, 2020). I profili con livelli di istruzione elevati (come lauree o specializzazioni) occupano più spesso ruoli manageriali, tecnici o di concetto, che per loro natura richiedono l'uso di tecnologie digitali e presentano un'alta "remotizzabilità". Al contrario, i lavoratori con titoli di studio inferiori sono prevalentemente impiegati in mansioni manuali, operative o di front-office, che richiedono necessariamente la presenza fisica sul luogo di lavoro.

Inoltre, il livello di istruzione impatta direttamente sulla produttività individuale: un maggiore capitale umano si traduce in una superiore capacità di autogestione, *problem solving* e familiarità con gli strumenti digitali, competenze cruciali quando viene meno la supervisione diretta del manager in ufficio. Non controllare per l'istruzione creerebbe quindi una grave distorsione da variabile omessa, poiché il modello rischierebbe di attribuire allo smart working un aumento di produttività che è invece dovuto alle superiori competenze di base dei lavoratori scelti per lavorare da casa.

Per tracciare il capitale umano, l'indagine adotta la variabile derivata *hiqual_dv* (*Highest qualification ever reported*), contenuta nei file dei

questionari individuali (*indresp*). Questa variabile fotografa il titolo di studio più elevato conseguito dal rispondente fino al momento dell'intervista.

Understanding Society adotta una strategia molto logica per questa variabile: l'informazione viene raccolta al momento dell'ingresso del rispondente nel panel e, negli anni successivi, il dato viene ereditato e aggiornato solo se il soggetto dichiara di aver conseguito un nuovo titolo di studio, evitando così di dover ripetere la stessa domanda a ogni *wave*.

Le risposte mancanti o i rifiuti per questo dato sono quasi inesistenti, aggirandosi intorno allo 0,1% in quasi tutte le rilevazioni. L'unica eccezione visibile si registra nella *wave* 6 (circa il 3,7% di dati mancanti), dovuta, come riportato in una nota nel database, all'inserimento nel campione di una forte quota di popolazione immigrata, per la quale la classificazione dei titoli di studio ottenuti all'estero ha richiesto verifiche più complesse (Understanding Society, n.d.).

Dal punto di vista statistico, i dati mostrano una chiara e progressiva transizione della popolazione del Regno Unito verso livelli di istruzione più elevati nel corso dei quindici anni di rilevazione:

- *I titoli universitari (degree)* mostrano una crescita costante, passando dal 21,48% della *wave* 1 a oltre il 32% nella *wave* 15.
- *I soggetti senza alcuna qualifica (no qualification)* quasi si dimezzano, scendendo dal 17,78% iniziale fino a un ridotto 6,89% nell'ultima *wave*.
- *I titoli intermedi*, come i diplomi di scuola superiore (*A-Level*) e quelli corrispondenti alla scuola dell'obbligo (*GCSE*), mantengono una presenza robusta e stabile, oscillando costantemente attorno al 20% ciascuno per l'intero periodo.

Nel modello di stima lineare (OLS), l'istruzione viene inserita come variabile qualitativa (*i. education*). Questo permette di stimare l'impatto di ciascun titolo di studio sulla produttività oraria prendendo una categoria come base di confronto. Infine, coerentemente con la teoria econometrica, la variabile viene esclusa dalla specificazione a effetti fissi (xtreg), poiché il livello di istruzione degli individui adulti rimane stabile nel tempo e il suo effetto viene assorbito direttamente dall'effetto fisso individuale.

2.3.4. Settore industriale (*jbsic07_cc*)

Per il presente lavoro, la struttura industriale di riferimento rappresenta una covariata di controllo fondamentale. Per schematizzare l'eterogeneità settoriale, questo studio utilizza la variabile armonizzata *jbsic07_cc*, rilevata costantemente lungo le quindici *waves* del panel.

Questa variabile riflette la classificazione standard delle industrie del Regno Unito (SIC 2007) condensata a livello di codice a due cifre, riportando l'attività economica principale del rispondente attraverso 99 categorie discrete (dall'agricoltura 01 alle organizzazioni extraterritoriali 99). L'inclusione di questo controllo nei modelli di regressione è necessaria per isolare la varianza legata alla diversa "telelavorabilità" delle mansioni, strutturalmente più elevata nei servizi avanzati e ridotta nel comparto manifatturiero.

La consistenza longitudinale di questo indicatore richiede tuttavia un'attenta contestualizzazione metodologica a causa di due variazioni operative introdotte nel protocollo di rilevazione del questionario:

1. *La discontinuità della wave 13*: in questa *wave* (*m_jbsic07_cc*), la sospensione temporanea della tecnica delle *interviste dipendenti* (*dependent interviewing*) ha costretto a raccogliere ex-novo le informazioni sulle attività industriali, senza poter verificare la codifica dell'anno precedente. Questo mutamento procedurale ha

causato un prevedibile incremento delle risposte mancanti – con i valori di sistema (*missing* -9) saliti al 4,09% e i casi non applicabili (*inapplicable* -8) al 46,38% – introducendo al contempo il rischio di transizioni settoriali artificiali dovute a lievi differenze nella descrizione linguistica dell'azienda da parte dello stesso rispondente.

2. *La stabilizzazione nelle waves 14 e 15*: a partire dalla Wave 14 (*n_jbsic07_cc*), l'ISER ha introdotto un controllo basato sulla stabilità del posto di lavoro (*jobcodechk*): e il rispondente conferma di non aver cambiato datore di lavoro, il software assegna automaticamente il codice SIC della *wave* precedente, minimizzando i vizi di codifica. Nella *wave* 15 (*o_jbsic07_cc*), questo impianto stabilizzato restituisce un totale di 32.849 osservazioni lorde, in cui i casi non applicabili (46,06%) riflettono coerentemente i soggetti fuori dalle forze di lavoro, i disoccupati e, per il disegno di questa ricerca, i lavoratori autonomi.

All'interno del campione attivo dei lavoratori dipendenti, la distribuzione mostra una forte concentrazione nei settori dei servizi pubblici e del welfare, guidati dall'Istruzione (settore 85; 6,19% del campione totale), dai servizi sanitari (settore 86; 4,95%) e dal commercio al dettaglio (settore 47; 4,09%). Al contrario, molti comparti industriali specifici presentano frequenze estremamente esigue (inferiori allo 0,05%).

Per garantire la stabilità algebrica dei modelli ed evitare problemi dovuti alla frammentazione originaria in 99 codici, le categorie sono state aggregate in 8 *macro-settori industriali*:

1. *Agriculture & Primary* (SIC 01-09)
2. *Manufacturing* (SIC 10-33)
3. *Construction* (SIC 41-43)

4. *Trade & Transport* (SIC 45-56)
5. *Information & Communication* (SIC 58-63)
6. *Finance* (SIC 64-66)
7. *Public & Social Services* (SIC 84-88)
8. *Other Services* (SIC 68-79, 80-82, 90-96)

Questa suddivisione in macro-settori viene inserita nel modello attraverso un gruppo di variabili dicotomiche. In questo modo, l'analisi riesce a isolare il vero impatto dello smart working sulla produttività oraria, "ripulendolo" dalle naturali differenze di stipendio e di organizzazione che caratterizzano i diversi settori lavorativi.

2.3.5. Contesto geografico e residenziale (*country* e *urban_dv*)

Il presente studio integra poi due indicatori geografici come variabili di controllo: la nazione costitutiva di residenza (*country*) e il grado di urbanizzazione dell'area di domicilio (*urban_dv*).

La variabile armonizzata *country* mappa la scomposizione geopolitica del campione tra le quattro nazioni costitutive del Regno Unito: Inghilterra (1), Galles (2), Scozia (3) e Irlanda del Nord (4). Come specificato dalle note metodologiche del panel, l'indicatore viene generato a partire dai codici delle macro-regioni d'ufficio (*Government Office Region*, GOR) ed è strettamente vincolato alla presenza di un codice postale valido nei database gestionali della rilevazione; nei rari casi in cui tale informazione risulta mancante o errata, la variabile assume un codice di *missing* di sistema (-9) (Understanding Society n.d.).

La distribuzione dei dati tra le quattro nazioni rimane molto stabile nel tempo. Si nota un progressivo calo nel numero totale di intervistati dovuto

al naturale abbandono del panel da parte di alcuni partecipanti; questa perdita è stata però compensata dall'introduzione di nuovi campioni nelle *waves* successive, come dimostra il netto recupero registrato nella *wave* 14, dove il numero di osservazioni sale a 35.471 rispetto alle 29.271 della *wave* 12. Nella *wave* 15 (*o_country*), a fronte di un tasso di dati mancanti quasi nullo (0,02%), la distribuzione della popolazione attiva riflette la storica asimmetria demografica del dataset: l'Inghilterra raccoglie la netta maggioranza delle osservazioni (79,44%), seguita dalla Scozia (8,88%), dal Galles (5,92%) e dall'Irlanda del Nord (5,74%). L'inserimento di questo controllo nei modelli multivariati è essenziale per assorbire le differenze istituzionali e legislative in materia di tutele del lavoro, nonché i diversi mercati occupazionali regionali.

Dopodiché, per catturare la densità infrastrutturale e il divario digitale tra centri urbani e aree rurali – fattori che influenzano direttamente sia la qualità della connessione internet sia la concentrazione di lavori che è possibile svolgere al computer – lo studio si avvale dell'indicatore dicotomico derivato *urban_dv*. Questa variabile classifica gli indirizzi degli intervistati in:

- *Area urbana (1)*: insediamenti con una popolazione pari o superiore a 10.000 abitanti.
- *Area rurale (2)*: contesti sotto la soglia dei 10.000 abitanti o insediamenti sparsi.

La classificazione si basa sull'applicazione dei protocolli ufficiali dell'Office for National Statistics (*Rural and Urban Classification of Output Areas*). L'analisi longitudinale mostra uno spostamento progressivo e lineare nella composizione del campione osservato lungo le quindici *waves*: se nella *wave* 1 i residenti in aree urbane rappresentavano il 79,32% del totale, tale quota si contrae costantemente fino ad attestarsi

al 73,78% nella *wave* 15, a favore di una crescita della componente rurale, che raggiunge il 26,20%. I dati mancanti di sistema rimangono marginali per l'intera durata del panel, con un'unica fluttuazione isolata nella *wave* 11, in cui i mancati abbinamenti postali hanno temporaneamente spinto i *missing* al 2,06%.

Nei modelli di regressione, l'interazione controllata di queste due variabili geografiche impedisce che l'effetto stimato dello smart working sulla produttività venga confuso con i vantaggi agglomerativi tipici delle grandi aree metropolitane inglesi rispetto ai contesti rurali o periferici delle nazioni del Nord.

2.3.6. Numero di figli minori di 16 anni (*nchresp*)

L'impianto empirico integra inoltre una covariata relativa alla composizione familiare. Per quanto concerne il contesto domestico, la letteratura economica evidenzia come la presenza di figli minorenni o a carico eserciti una forte pressione sul lavoratore, condizionando sia l'allocazione del tempo sia l'efficienza della prestazione lavorativa. Se da un lato il *WFH* agevola la conciliazione dei tempi di vita, eliminando i tempi e i costi del pendolarismo e offrendo una maggiore flessibilità organizzativa (Bloom et al., 2015), dall'altro la sovrapposizione tra lo spazio professionale e quello domestico può generare una frammentazione dell'attività lavorativa.

Le frequenti interruzioni legate alla gestione e alla cura dei figli tendono a interrompere i flussi di concentrazione, riducendo la produttività oraria complessiva, soprattutto in contesti in cui i servizi educativi o di supporto all'infanzia sono carenti o non accessibili (Gibbs et al., 2023). Questo effetto di spillover negativo, inoltre, non si distribuisce in modo omogeneo tra i generi: la letteratura empirica dimostra che l'impatto asimmetrico tende a colpire in misura sproporzionata le lavoratrici madri, esasperando

i divari di genere preesistenti nei carichi di cura e nel lavoro domestico non retribuito (Andrew et al., 2022).

Per controllare questa potenziale fonte di eterogeneità e catturare l'intensità dei carichi di cura domestica, il modello introduce la variabile continua derivata *nchresp*, contenente il numero di figli minori di 16 anni di cui ogni intervistato è responsabile. Sotto il profilo della costruzione del dato, l'indicatore viene calcolato dai programmatori del dataset a partire dalla griglia relazionale del nucleo familiare (*household grid*), che mappa i legami di parentela e di convivenza all'interno di ciascuna unità abitativa. È importante sottolineare che, per via di come è strutturato il questionario, nei nuclei familiari in cui sono presenti entrambi i genitori la responsabilità dei figli viene quasi sempre assegnata alla madre (Understanding Society, n.d.); di conseguenza, la variabile è particolarmente efficace per catturare il carico di cura che grava sulle donne del campione.

Dal punto di vista longitudinale, l'analisi della variabile mostra dinamiche interessanti lungo le diverse *waves* del panel. Nelle prime *waves*, infatti, si nota una quota non trascurabile di dati mancanti dovuti alle *proxy interviews* (le interviste rilasciate da un altro membro della famiglia quando il rispondente principale era temporaneamente irreperibile, codificate con il valore -7). Questo fenomeno tocca il suo picco dell'8,60% nella *wave* 5, ma crolla progressivamente nelle rilevazioni successive. Nella *wave* 15 le mancate risposte si stabilizzano a un livello fisiologico dello 0,23% (appena 74 osservazioni su un totale di 32.849 individui), a dimostrazione dell'ottima capacità del panel di mantenere e tracciare il campione nel corso degli anni.

I dati descrittivi longitudinali mostrano inoltre che il numero medio di figli *under* 16 per rispondente cala in modo quasi lineare, passando dallo 0,343

della *wave* 1 allo 0,217 della *wave* 15. Questo trend discendente è legato a due fattori principali:

1. *La denatalità strutturale*: il generale calo delle nascite e la contrazione del tasso di fecondità che hanno caratterizzato il Regno Unito nel periodo considerato.
2. *L'invecchiamento del campione*: per la natura stessa dei panel longitudinali, i figli dei rispondenti intervistati nelle prime *waves* crescono e superano la soglia dei 16 anni. Di conseguenza, escono dal computo della variabile *nchresp* per entrare in quello dei membri adulti del nucleo familiare.

Sotto il profilo econometrico, nei modelli di stima lineare OLS, la variabile *nchresp* viene inserita come variabile continua per catturare l'effetto marginale di ogni figlio in più sulla produttività oraria logaritmica. Il ruolo di questo controllo diventa ancora più centrale nel modello panel a effetti fissi (*xtreg, fe*) per identificare l'effetto causale dello smart working. Poiché eventi come la nascita di un figlio o il superamento dei 16 anni sono variazioni che avvengono nel tempo all'interno della storia del singolo lavoratore (*within*), questi mutamenti non vengono assorbiti dall'effetto fisso individuale α_i . Inserire *nchresp* tra i controlli permette quindi di depurare il coefficiente dello smart working dagli shock temporanei legati ai carichi di cura e alla gestione della casa, garantendo che la stima della produttività non sia distorto dall'omissione di fattori legati al ciclo di vita familiare.

2.3.7. Frequenza di utilizzo di internet (*netpusenew*)

Parallelamente alla dimensione familiare, un fattore altrettanto critico per isolare l'impatto del lavoro da remoto sulla produttività è rappresentato dall'alfabetizzazione digitale del lavoratore. La letteratura economica evidenzia da tempo che l'introduzione di nuove tecnologie, come

l'adozione di software collaborativi, reti aziendali protette e piattaforme in cloud, richiede abilità specifiche da parte del personale per evitare perdite di efficienza e cali di performance (Acemoglu e Autor, 2011). Un lavoratore che possiede già una buona familiarità con gli strumenti informatici e telematici sperimenta minori costi di transizione, una riduzione dei tempi morti legati a problemi tecnici e una maggiore autonomia operativa quando si trova a lavorare da casa, rispetto a colleghi meno integrati digitalmente o meno abituati all'uso quotidiano della tecnologia (Frank et al., 2019).

Per inserire questa fondamentale dimensione di controllo all'interno del modello empirico, si utilizza la variabile categoriale ordinata *netpusenew*, definita come la frequenza di utilizzo di internet e costruita a partire dalle risposte fornite al quesito: *"How often do you use the internet for your personal use?"* (Con quale frequenza utilizzi internet per motivi personali?). Sotto il profilo concettuale, questo indicatore viene utilizzato come proxy – ossia come una misura indiretta e di approssimazione – dell'inclinazione tecnologica generale dell'individuo. L'assunto teorico alla base di questa scelta è che un uso pervasivo e frequente della rete nel proprio tempo libero rifletta lo sviluppo di competenze digitali intuitive e di una confidenza tecnica di fondo. Si tratta di un bagaglio di abilità che il lavoratore può riutilizzare e spendere con successo anche nello svolgimento delle proprie mansioni lavorative da remoto, facilitando l'autodisciplina digitale e l'uso dei canali di comunicazione aziendali. Questa variabile armonizzata sostituisce nei questionari di *Understanding Society* la vecchia codifica *netuse*, utilizzata nelle prime *waves*; questo aggiornamento metodologico si è reso necessario per catturare in modo più fedele l'evoluzione tecnologica dell'ultimo decennio, segnato dall'esplosione della connettività mobile tramite smartphone e tablet (*Understanding Society*, n.d.).

Dal punto di vista longitudinale, l'analisi della qualità del dato mostra una stabilità e una robustezza eccezionali nel corso del tempo. I tassi di mancata risposta complessiva – che includono i rifiuti, le risposte incerte (*don't know*) o i dati mancanti dovuti a interviste rilasciate da un altro membro della famiglia (*proxy*) – rimangono costantemente al di sotto dell'1,5% nella maggior parte delle *wave*, toccando un minimo fisiologico dello 0,41% proprio nella *wave* 15. Questa precisione nella raccolta conferma la familiarità ormai totale dei rispondenti con l'oggetto della domanda e riduce al minimo il rischio di distorsioni nelle stime dovute alla perdita di dati. I dati descrittivi evidenziano una imponente transizione del campione verso l'iper-connettività di massa nel corso degli anni: se la categoria di chi usa internet più volte al giorno (*Several times a day*) si mantiene dominante e sostanzialmente stabile attorno al 43,99% nella *wave* finale, la quota di utilizzo continuativo (*Almost all of the time*) mostra una crescita solida e costante, passando dal 32,70% della *wave* 12 al 35,82% della *wave* 15. Specularmente, i soggetti che dichiarano di non utilizzare mai la rete (*Never use*) o di non disporre di alcun tipo di accesso a internet (*No access*) scendono complessivamente dal 5,83% all'inizio del periodo osservato fino al 3,81% nella *wave* finale.

Trattandosi di un indicatore categoriale, nei modelli di stima OLS la variabile *netpusenew* non può essere inserita in forma continua. Viene quindi implementata attraverso la creazione di un set di variabili dicotomiche, una per ciascuna modalità di risposta. Da questo gruppo viene volutamente esclusa la categoria degli utilizzatori ad altissima intensità (*Almost all of the time*), che viene scelta come base di riferimento per verificare se esista un'eventuale differenza di produttività oraria legata alla familiarità tecnologica delle altre categorie rispetto ai lavoratori costantemente connessi. Nel modello panel a effetti fissi (*xtreg, fe*), l'inclusione di questo set di controlli diventa ancora più centrale per

garantire la corretta identificazione dell'effetto causale dello smart working. Poiché il mutamento individuale nell'intensità d'uso della rete nel corso del tempo (ad esempio, il passaggio da un uso saltuario a un uso continuativo dovuto all'acquisto di un nuovo dispositivo o a nuove abitudini personali) rappresenta una variazione che avviene *within* – ossia nella storia interna e nella traiettoria temporale del singolo individuo –, questi cambiamenti non vengono assorbiti dall'effetto fisso specifico α_i . Di conseguenza, l'inserimento delle dummy di *netpusenew* permette di depurare il coefficiente stimato del WFH dall'effetto legato all'apprendimento tecnologico spontaneo del lavoratore, garantendo che l'impatto stimato sulla produttività non sia confuso con la naturale evoluzione delle competenze digitali individuali.

3. Strategia empirica e analisi causale

La valutazione dell'impatto economico dello smart working sulle performance dei lavoratori richiede l'adozione di un impianto metodologico rigoroso, capace di isolare i canali causali dalle semplici correlazioni statistiche. Questo capitolo illustra in dettaglio la strategia econometrica implementata per stimare l'effetto del lavoro da casa sulla produttività oraria nel Regno Unito, sfruttando la natura longitudinale della base dati ricavata da *Understanding Society*.

Il percorso analitico si sviluppa attraverso un confronto tra diversi approcci statistici, strutturati per correggere progressivamente i problemi di endogeneità¹¹, la distorsione da variabili omesse e i fenomeni di autoselezione tipici dei dati osservazionali sul mercato del lavoro. Nei paragrafi che seguono verranno esaminati l'obiettivo della ricerca, i modelli lineari di base stimati in pool¹², e l'estensione fondamentale verso i modelli panel a effetti fissi individuali, chiarendo per ciascuna metodologia le ipotesi di partenza, la logica utilizzata per dimostrare il nesso di causa-effetto e le modalità con cui le variabili del dataset originale sono state ricodificate e strutturate nel software Stata.

3.1. Obiettivo dell'analisi e sfide metodologiche

L'obiettivo centrale di questo lavoro di ricerca consiste nell'isolare e stimare l'effetto causale dell'adozione del lavoro agile sulla produttività oraria dei lavoratori dipendenti nel Regno Unito. Come ampiamente discusso nel primo capitolo, la letteratura economica offre visioni contrastanti sul segno e sull'intensità di questo impatto. Da un lato, vi sono

¹¹ Condizione econometrica in cui una variabile esplicativa è correlata con il termine di errore del modello, determinata da variabili omesse, errori di misura o causalità inversa, che rende gli stimatori standard (come l'OLS) distorti e non coerenti.

¹² Struttura di dati ottenuta aggregando osservazioni trasversali ripetute nel tempo (in questo caso, individui diversi intervistati in anni differenti) in un unico blocco campionario, trattando le osservazioni come campionariamente indipendenti e ignorando temporaneamente la dimensione longitudinale.

studi che sostengono l'esistenza di guadagni di efficienza legati all'abbattimento dei tempi di pendolarismo e a una maggiore flessibilità organizzativa che favorisce la concentrazione (Bloom et al., 2015). Dall'altro lato, si registrano riscontri empirici di segno opposto, che evidenziano perdite di produttività dovute alla frammentazione degli spazi domestici, alle distrazioni familiari e ai costi di ordinario coordinamento (Gibbs et al., 2023; Barrero et al., 2021).

Dal punto di vista empirico, la principale sfida metodologica risiede nel fatto che la scelta di lavorare da remoto non è assegnata in modo puramente casuale all'interno del mercato del lavoro. Al contrario, essa scaturisce da un processo di selezione endogeno, guidato sia da scelte individuali e preferenze del lavoratore (legate a caratteristiche personali, attitudini tecnologiche o precise dinamiche familiari) sia da determinanti aziendali (come il settore industriale o la natura delle mansioni svolte). Di conseguenza, un'analisi empirica che non tenga conto di questi fattori correrebbe il rischio di produrre stime distorte, confondendo l'effetto puro del lavoro da casa con caratteristiche latenti degli individui o delle organizzazioni in cui operano.

Sfruttando la struttura longitudinale dei dati di *Understanding Society* presentati nel secondo capitolo, questo studio formalizza una strategia empirica che si sviluppa in modo incrementale per isolare il vero effetto del WFH. L'analisi si articola su due pilastri metodologici fondamentali. Il punto di partenza è rappresentato dal modello *pooled OLS*, applicato a una struttura di dati in cui la dimensione temporale viene temporaneamente ignorata per fissare le correlazioni di fondo nel campione. Il nucleo centrale e più rigoroso della tesi è invece costituito dal modello panel a effetti fissi individuali. Questa seconda strategia sfrutta esclusivamente la variazione nel corso del tempo interna alla storia del singolo lavoratore per eliminare interamente la distorsione derivante

da variabili omesse non osservabili che rimangono costanti, come l'abilità innata o la motivazione personale, permettendo così di verificare se l'eventuale premio di produttività associato al lavoro da remoto sia reale o se sia guidato da un processo di autoselezione dei lavoratori più performanti. Nelle sezioni che seguono vengono presentati i singoli approcci e la logica di costruzione dei blocchi di variabili inseriti nel software Stata.

3.2. La stima preliminare: *pooled OLS*

La prima fase dell'analisi empirica prevede la stima di un modello di regressione lineare classico sui dati strutturati in forma di pool. In questo primo approccio metodologico, le quindici *waves* dell'indagine (dalla *a* alla *o*) vengono temporaneamente aggregate, ignorando la dimensione temporale del panel. Ciò significa che le osservazioni ripetute nel tempo per uno stesso individuo vengono trattate dal software come se provenissero da un'unica grande indagine trasversale, accumulando tutti i dati in un unico blocco campionario.

La strategia d'analisi applicata in questa sezione si distingue per la scomposizione dell'impatto del lavoro da casa lungo due direttrici distinte, stimando modelli separati per il margine intensivo e per il margine estensivo:

- *Il margine intensivo (frequenza del WFH)*: si concentra sull'intensità dell'utilizzo dello smart working. Per stimarlo, viene inserita una variabile categoriale articolata su sette livelli di frequenza (che spaziano da "Mai" fino a "Sempre"). La categoria di chi non lavora mai da casa viene impostata come livello di riferimento, consentendo di misurare il guadagno o la perdita percentuale di produttività per ogni specifica frequenza di lavoro da remoto rispetto a chi opera esclusivamente in presenza.

- *Il margine estensivo (accesso al WFH)*: si focalizza sulla scelta binaria. Viene costruita una variabile dummy che assume valore zero per chi non lavora mai da remoto e valore uno per tutti i lavoratori che utilizzano lo smart working, a prescindere che lo facciano per pochi giorni al mese o stabilmente tutti i giorni. Questo indicatore cattura l'effetto medio dell'avere accesso a questa modalità di lavoro.

3.2.1. La strategia di stima a blocchi progressivi

Per tracciare l'impatto dei diversi fattori confondenti e osservare come cambiano i coefficienti del lavoro agile, sia i modelli del margine intensivo sia quelli del margine estensivo vengono stimati inserendo le variabili nel modello attraverso tre blocchi progressivi:

1. *Il blocco 1 (modello base)* introduce esclusivamente l'indicatore del lavoro agile, i controlli demografici fondamentali legati all'età e al suo quadrato (essenziale per catturare l'andamento non lineare e speculare della produttività lungo il ciclo della vita lavorativa), il genere e il livello massimo di istruzione conseguito. Quest'ultima variabile viene ricodificata su sei livelli quantitativi che riflettono i titoli di studio (dall'assenza di qualifiche fino alla laurea universitaria) per evitare di assumere una linearità irrealistica nei rendimenti dell'istruzione. Vengono inoltre inseriti gli effetti fissi temporali per ogni singola *wave* della rilevazione, necessari per assorbire gli shock macroeconomici comuni, i cicli economici generali e i mutamenti normativi che colpiscono simultaneamente tutti i lavoratori nel Regno Unito in un dato anno.
2. *Il blocco 2 (famiglia e geografia)* estende il modello iniziale integrando i fattori domestici e territoriali. Viene introdotto il numero di figli conviventi sotto i sedici anni di cui il rispondente è responsabile (inserito in forma lineare continua) e un indicatore

dicotomico relativo alla residenza in aree rurali rispetto alle aree urbane, oltre alle variabili relative alla specifica nazione del Regno Unito di residenza (Inghilterra, Galles, Scozia o Irlanda del Nord).

3. Il *blocco 3 (modello completo)* rappresenta la configurazione più ricca e restrittiva del modello. Al suo interno vengono aggiunti otto indicatori per i macro-settori industriali di attività (dall'agricoltura alla manifattura, fino ai servizi finanziari e sociali) e le variabili relative alla frequenza di utilizzo personale di internet, impostando la categoria degli utilizzatori più intensivi ("Quasi tutto il tempo") come base di confronto.

Sotto il profilo strettamente empirico, l'introduzione di quest'ultimo blocco comporta un importante trade-off campionario, poiché la variabile sulla frequenza di utilizzo di internet è presente in modo sistematico solo nelle *waves* più recenti dell'indagine di *Understanding Society*. Per evitare che il confronto tra i diversi blocchi sia inficiato da shock di composizione dei dati, si è scelto di adottare una strategia a campione costante (*constant sample*). Di conseguenza, tutte le specificazioni econometriche (dalla baseline al modello completo) vengono stimate esclusivamente sul medesimo sotto-insieme di osservazioni per cui sono disponibili le informazioni digitali del blocco 3. Questo approccio garantisce una perfetta confrontabilità dei coefficienti. Di conseguenza, il passaggio progressivo dal blocco 1 al blocco 3 consente di isolare analiticamente l'impatto della distorsione da variabili omesse, valutando se e quanto l'effetto originariamente attribuito allo smart working fosse in realtà guidato dai carichi familiari, dal contesto territoriale di residenza, dal macro-settore industriale di appartenenza o dalle competenze digitali individuali del lavoratore.

3.2.2. Limiti del modello *pooled OLS* e correzione degli errori

Nonostante la ricchezza dei controlli inseriti nel blocco 3, i coefficienti stimati tramite il modello *pooled OLS* non possono ancora essere interpretati in senso causale. Perché questo stimatore sia privo di distorsione, dovrebbe valere l'assunzione di esogeneità¹³ forte, secondo cui il termine di errore non deve contenere alcun fattore non osservato che sia simultaneamente correlato con la scelta di fare smart working e con la produttività del lavoratore.

Questa assunzione risulta sistematicamente violata a causa dei meccanismi di autoselezione degli individui. Il termine di errore complessivo racchiude infatti al suo interno l'effetto fisso specifico dell'individuo, ovvero l'insieme delle caratteristiche personali, psicologiche e attitudinali stabili nel tempo che l'analista non può inserire nel dataset, come l'abilità innata, la motivazione interna, l'auto-disciplina e la resistenza allo stress.

Nel modello *pooled OLS* questo effetto fisso individuale viene ignorato e incorporato nel termine di errore. Poiché è logico ipotizzare che un lavoratore intrinsecamente più motivato e disciplinato sia al contempo più produttivo e abbia una probabilità significativamente più alta di chiedere, ottenere o scegliere accordi di lavoro agile dal proprio datore di lavoro, si genera una correlazione sistematica tra la variabile indipendente del WFH e l'errore del modello. Questa violazione produce una distorsione verso l'alto dello stimatore OLS, che tenderà a sovrastimare il vero impatto del lavoro agile sulla produttività, attribuendo allo smart working un merito statistico che appartiene invece alle doti caratteriali del soggetto.

Un ulteriore limite strutturale del modello *pooled* riguarda il calcolo degli intervalli di confidenza. Poiché lo stesso individuo viene osservato ripetutamente per più anni, le osservazioni non sono indipendenti tra loro

¹³ Proprietà statistica per cui una variabile esplicativa è indipendente dal termine di errore del modello, garantendo che il coefficiente stimato rifletta esclusivamente la relazione sistematica con la variabile dipendente, senza risentire dell'influenza di fattori esterni non osservati.

a livello campionario. I residui dello stesso soggetto mostrano una forte correlazione seriale nel tempo che, se trascurata da una regressione standard, porta a una grave sottostima degli errori standard e, di conseguenza, a un'inflazione artificiale della significatività statistica. Per correggere questo secondo problema, tutte le regressioni OLS introducono una correzione della varianza tramite stimatori robusti. Tuttavia, per risolvere il problema centrale della distorsione da variabili omesse e neutralizzare l'effetto delle caratteristiche latenti individuali, si rende indispensabile il passaggio alla strategia panel a effetti fissi descritta nella sezione successiva.

3.3. Il modello panel a effetti fissi individuali

Per superare i limiti strutturali del modello OLS e neutralizzare la distorsione derivante dalle variabili omesse non osservabili, la seconda parte della strategia empirica sfrutta appieno la dimensione longitudinale del dataset attraverso la stima di un modello panel a effetti fissi individuali. Sotto il profilo operativo, questo approccio viene implementato nel software Stata utilizzando il comando specifico *xtreg* con l'opzione *fe* (*fixed effects*). La transizione metodologica segna un punto di svolta cruciale nell'impianto della ricerca: mentre il modello *pooled OLS* confrontava i livelli di produttività tra individui diversi trattandoli come un aggregato campionario omogeneo, il modello a effetti fissi opera una trasformazione matematica dei dati basata sugli scostamenti dalla media individuale, nota in letteratura come trasformazione *within*. Per ciascun lavoratore, il software calcola la media storica di ogni variabile lungo tutte le *waves* in cui è presente e la sottrae dal valore osservato nell'anno specifico. Il modello, di conseguenza, ignora completamente le differenze statiche tra i lavoratori e isola esclusivamente le variazioni che avvengono nel corso del tempo all'interno della storia professionale del singolo soggetto.

Il vantaggio fondamentale di questa strategia risiede nella scomposizione analitica del termine di errore. L'errore complessivo del modello viene scisso in due componenti distinte: un errore casuale classico, che varia in modo imprevedibile tra gli individui e nel tempo, e un effetto fisso individuale stazionario. Quest'ultimo cattura l'insieme di tutte le caratteristiche personali, psicologiche e attitudinali del lavoratore che rimangono costanti nel corso dei quindici anni della rilevazione, come l'abilità innata, l'orientamento al risultato, l'estrazione culturale di partenza, l'ambizione e i tratti stabili della personalità. Attraverso la trasformazione dei dati incentrata sulle medie individuali, l'effetto fisso costante viene matematicamente annullato e rimosso dall'equazione di stima. Di conseguenza, tutti i fattori latenti che nel modello in pool inquinavano i coefficienti dello smart working e ne gonfiavano i risultati vengono interamente eliminati, risolvendo alla radice il problema dell'autoselezione basata su caratteristiche fisse del lavoratore.

Anche per questa seconda metodologia, la strategia di stima viene applicata separatamente al margine intensivo (le frequenze d'uso del lavoro agile) e al margine estensivo (la scelta binaria di accesso), replicando la struttura a blocchi progressivi già impostata per il modello OLS:

- *Il blocco 1 (modello base ad effetti fissi)* include le variabili del lavoro da remoto, l'età e il suo termine quadratico e gli effetti fissi legati alle *waves* temporali (*i.wave*). In questa configurazione, la variabile indicatrice del genere biologico e i titoli di studio che non hanno subito variazioni nel corso del tempo vengono automaticamente esclusi dal calcolo o ridotti a costanti individuali dal software, poiché non presentano alcuna variazione interna (*within*) per lo stesso soggetto. Le dummy temporali mantengono invece un ruolo cruciale per assorbire le tendenze

macroeconomiche globali e i mutamenti del mercato del lavoro che si sviluppano parallelamente all'invecchiamento dei lavoratori.

- *Il blocco 2 (famiglia e geografia ad effetti fissi)* introduce nell'equazione longitudinale il numero di figli conviventi minori di sedici anni e i mutamenti legati alla localizzazione territoriale (il passaggio da un'area urbana a una rurale o il trasferimento tra le diverse nazioni del Regno Unito). I coefficienti di questo blocco misurano come varia la produttività oraria dello stesso individuo nei periodi in cui si trova a gestire un carico familiare mutato o un cambiamento di contesto residenziale, permettendo di capire se la flessibilità del lavoro da casa aiuti a bilanciare questi shock esogeni.
- *Il blocco 3 (modello completo ad effetti fissi)* integra infine le variabili relative ai macro-settori industriali di attività e alla frequenza di utilizzo della rete internet. Anche nel contesto dei *fixed effects*, l'inclusione di quest'ultimo blocco determina una contrazione delle osservazioni disponibili, limitando l'analisi alle *waves* in cui l'indicatore sull'intensità d'uso di internet è presente. Il modello permette tuttavia di verificare se la stabilità del nesso causale risenta dei passaggi di carriera da un settore economico all'altro o delle variazioni nelle abitudini digitali del soggetto.

Sotto il profilo del rigore statistico, la pulizia operata dal modello a effetti fissi comporta un preciso trade-off metodologico in termini di varianza. Poiché lo stimatore cancella le differenze tra gli individui e attinge solo dalle variazioni storiche interne al singolo lavoratore, i coefficienti vengono identificati esclusivamente grazie a quel sotto-insieme di soggetti che, nel corso degli anni osservati, modificano la propria condizione lavorativa, passando dal lavoro in presenza allo smart working o variando la frequenza dei giorni spesi a casa. Se i comportamenti individuali

rimangono fortemente stabili nel tempo, la variabilità utile ai fini della stima si riduce sensibilmente, aumentando gli errori standard dei coefficienti.

Inoltre, sebbene il modello elimini ogni distorsione derivante da fattori omessi *costanti*, esso rimane potenzialmente vulnerabile a shock non osservati che variano nel tempo e che potrebbero influenzare simultaneamente la produttività e le modalità di lavoro (come una promozione improvvisa o un cambiamento radicale delle mansioni). Per garantire la validità dei test statistici in presenza di osservazioni ripetute nel tempo per lo stesso individuo, in tutti i blocchi a effetti fissi vengono utilizzati errori standard robusti, raggruppati (*clustered*) a livello del singolo lavoratore. Questa correzione è fondamentale per tenere conto del fatto che i comportamenti e i fattori non osservati di uno stesso soggetto sono inevitabilmente correlati tra una *wave* e l'altra, evitando così di sovrastimare la precisione dei nostri coefficienti.

3.4. Risultati attesi e ipotesi di ricerca

Prima di passare all'analisi dei dati empirici, è opportuno formalizzare le aspettative teoriche circa il segno, l'ampiezza e la significatività dei coefficienti, delineate sulla base della letteratura economica esistente e della natura degli stimatori econometrici utilizzati. Le aspettative si articolano attorno a due ipotesi fondamentali, speculari ai due margini di analisi.

3.4.1. Ipotesi 1: il premio di produttività nel modello *pooled OLS*

Per quanto riguarda il modello di stima preliminare, ci si attende di osservare coefficienti positivi, statisticamente significativi e di intensità rilevante associati allo smart working, sia sul margine estensivo sia su quello intensivo. In linea con la teoria del capitale umano e i modelli di

selezione sul mercato del lavoro, questa aspettativa non riflette necessariamente un impatto causale della tecnologia sulla performance. Si ipotizza, piuttosto, che il modello risenta pesantemente della distorsione da variabili omesse descritta in precedenza. I lavoratori dotati di maggiore autodisciplina, capacità organizzativa e motivazione intrinseca hanno maggiori probabilità di rinegoziare i propri contratti a favore del lavoro flessibile e, simultaneamente, registrano performance e salari orari più elevati.

Pertanto, nel modello *pooled OLS* ci si aspetta che lo smart working agisca da “segnale” (proxy) di queste abilità latenti, mostrando un forte legame positivo con la produttività che tenderà a ridursi progressivamente, senza tuttavia scomparire, man mano che si inseriscono i controlli per il settore industriale (blocco 3) e il capitale umano.

3.4.2. Ipotesi 2: la riduzione dell’effetto e il ruolo della non-significatività nel modello Panel (*Fixed Effects*)

La seconda ipotesi rappresenta il fulcro dell’analisi empirica e prevede un ridimensionamento dei coefficienti, unito a una potenziale perdita di significatività statistica, nel passaggio dai modelli OLS a quello a effetti fissi. Si ipotizza che, una volta che la trasformazione *within* ha rimosso l’impatto di fattori personali stabili nel tempo (come l’abilità innata, la motivazione individuale e l’autodisciplina), l’effetto puro dello smart working sulla produttività oraria possa rivelarsi più contenuto o statisticamente nullo.

Dal punto di vista teorico, questa ipotesi si fonda su due motivazioni principali:

1. *L’eliminazione dell’autoselezione*: gli studi empirici più recenti (Emanuel e Harrington, 2024; Gibbs et al., 2023) evidenziano che i lavoratori intrinsecamente più produttivi o più orientati

all'autogestione tendono a scegliere lo smart working rispetto agli altri. Se il premio di produttività osservato nei modelli *pooled* fosse guidato prevalentemente da questo meccanismo di selezione, una volta isolata la storia del singolo lavoratore nel tempo, il semplice cambio di modalità organizzativa potrebbe non generare grandi variazioni sulla sua efficienza oraria.

2. *La frequenza del WFH e i limiti della varianza panel*: la letteratura internazionale (Aksoy et al., 2022; Barrero, Bloom e Davis, 2021, 2023) suggerisce che gli effetti positivi sulla produttività non siano lineari, ma si concentrino nelle formule ibride (2-3 giorni a settimana), capaci di bilanciare il risparmio sui tempi di pendolarismo con il lavoro di squadra in ufficio. Al contrario, frequenze estreme possono generare costi di coordinamento o isolamento professionale (Yang et al., 2020). Nel modello a effetti fissi applicato ai dati di *Understanding Society*, la forte stabilità nel tempo delle scelte dei lavoratori – legata alla rigidità strutturale degli accordi contrattuali e delle routine organizzative – fa sì che i casi di effettivo cambiamento da una modalità all'altra siano limitati. Di conseguenza, è probabile che molti dei coefficienti legati alle singole frequenze perdano significatività statistica, svelando i reali limiti di un impatto puramente meccanico del lavoro da remoto sulla performance, al netto delle doti personali del lavoratore (Anakpo et al., 2023; Morikawa, 2024).

4. Analisi empirica e risultati del modello

Questo capitolo presenta e discute in dettaglio i risultati emersi dall'analisi empirica condotta sul campione di dati panel. La trattazione si concentra sulla lettura e sull'interpretazione dei coefficienti emersi dalle stime, con l'obiettivo di valutare la relazione tra il WFH e le performance dei lavoratori. L'analisi si sviluppa a partire dai risultati del modello di benchmark¹⁴ ad effetti fissi, che rappresenta il nucleo centrale dell'indagine, per poi approfondire il ruolo delle diverse componenti di controllo inserite nelle regressioni.

4.1. Il modello di benchmark: stime ad effetti fissi (FE) individuali

Il modello di benchmark rappresenta la formulazione centrale dell'indagine empirica, selezionata in quanto considerata teoricamente più solida e idonea a isolare la relazione causale di interesse. La scelta di adottare lo stimatore ad effetti fissi (FE) come riferimento principale risiede nella sua capacità di risolvere il problema dell'eterogeneità individuale non osservata. Sfruttando la dimensione longitudinale del panel, questo approccio controlla tutte le caratteristiche dei lavoratori che rimangono costanti nel tempo – come le abilità innate, i tratti caratteriali, la motivazione di fondo o il posizionamento iniziale sul mercato del lavoro –, eliminando il *bias* da variabili omesse che inficerebbe una stima trasversale e garantendo così una valutazione molto più pulita dell'impatto del WFH.

A livello statistico, gli errori standard sono clusterizzati a livello individuale (sulla variabile *pidp*), garantendo la robustezza delle stime contro fenomeni di eteroschedasticità e la naturale correlazione dei dati

¹⁴ Formulazione principale e teoricamente più solida dell'analisi empirica, utilizzata come termine di paragone per valutare la stabilità e la coerenza dei coefficienti ottenuti in configurazioni alternative o stime preliminari.

raccolti sullo stesso individuo anno dopo anno. Al fine di rispondere puntualmente alle indicazioni metodologiche emerse in letteratura e catturare le diverse sfumature del fenomeno, l'analisi strutturata ad effetti fissi viene qui scissa in due macro-identità: l'analisi del margine estensivo (la scelta dicotomica di accedere o meno al WFH) e l'analisi del margine intensivo (la frequenza, espressa in giorni a settimana, del lavoro da remoto).

Le tabelle sottostanti raccolgono e ordinano i coefficienti emersi dalle stime effettuate per entrambi i margini, applicando una strategia a blocchi progressivi che parte da una configurazione iniziale con i soli controlli temporali (colonna 1), poi integra i fattori familiari e geografici variabili (colonna 2) e infine giunge al modello completo (colonna 3), comprensivo di caratteristiche settoriali e intensità d'uso di Internet.

4.1.1. Analisi del margine estensivo: l'impatto dell'accesso al WFH

Il margine estensivo esamina l'effetto netto derivante dallo *status* di lavoratore in smart working, indipendentemente dai giorni effettivi trascorsi a distanza. L'obiettivo econometrico è verificare se la transizione da un regime di presenza esclusiva a un regime di flessibilità generi uno shock positivo o negativo sulla performance oraria del lavoratore, una volta depurato il campione dalle determinanti fisse personali.

Tabella 1 – WFH e produttività oraria: stime FE sul margine estensivo
*Note: errori standard robusti riportati tra parentesi sotto i rispettivi coefficienti. Variabile dipendente: log_productivity_hour. I livelli di significatività sono indicati come segue: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. Elaborazione autonoma su dati Understanding Society.*

Variabile	Blocco 1: Base	Blocco 2: Geo & Famiglia	Blocco 3: Completo
Usa WFH (Rif: Never)	0.009 (0.012)	0.009 (0.012)	0.010 (0.012)
Area Geografica & Famiglia	//	//	//
Area rurale	–	-0.025 (0.040)	-0.025 (0.040)

<i>(Rif: Urban area)</i>			
Galles (<i>Rif: Inghilterra</i>)	–	-0.094 (0.058)	-0.097 (0.059)
Scozia	–	-0.435 (0.506)	-0.436 (0.506)
Irlanda del Nord	–	0.133 (0.148)	0.132 (0.147)
Numero di figli	–	-0.006 (0.018)	-0.007 (0.018)
Settore Industriale <i>(Rif: Agriculture & Primary)</i>	//	//	//
Manifattura	–	–	0.064 (0.095)
Edilizia	–	–	0.060 (0.099)
Commercio e Trasporti	–	–	0.050 (0.093)
Informazione e Comunicazione	–	–	-0.011 (0.105)
Finanza	–	–	0.048 (0.105)
Servizi Pubblici e Sociali	–	–	0.020 (0.096)
Altri Servizi	–	–	0.018 (0.105)
Uso Internet <i>(Rif: Almost all of the time)</i>	//	//	//
Diverse volte al giorno	–	–	-0.004 (0.009)
Una o due volte al giorno	–	–	0.028* (0.016)
Diverse volte a settimana	–	–	0.039* (0.021)
Diverse volte al mese	–	–	0.038 (0.026)
Una volta al mese	–	–	0.045 (0.034)
Meno di una volta al mese	–	–	0.091 (0.101)
Mai	–	–	0.104* (0.056)
Nessun accesso a internet	–	–	-0.030 (0.154)
Effetti Fissi Temporal <i>(Rif: Wave 12)</i>	//	//	//
Wave 13	0.039*** (0.008)	0.039*** (0.008)	0.039*** (0.008)
Wave 14	0.108*** (0.007)	0.108*** (0.007)	0.109*** (0.007)
Wave 15	0.170*** (0.008)	0.170*** (0.008)	0.172*** (0.008)
Costante	2.667*** (0.006)	2.713*** (0.050)	2.680*** (0.107)
<i>Osservazioni</i>	29,624	29,624	29,624
<i>R-quadrato (within)</i>	0.0346	0.0352	0.0360

I risultati empirici legati a questa prima specificazione offrono un quadro interpretativo di assoluta neutralità. Come evidenziato dai log di stima, la variabile dicotomica *Uses WFH* (avente come *baseline* di riferimento la condizione *Never WFH*) mostra un coefficiente pari a 0.0091 nel modello base (colonna 1), con un errore standard robusto di 0.0121. Il relativo test

t statistico si attesta a un valore di appena 0.75, generando un p -value pari a 0.451. Tale parametro risulta quindi ampiamente non significativo a qualunque livello di confidenza convenzionale.

L'invarianza del parametro viene pienamente confermata dall'introduzione dei successivi blocchi di controllo: nel secondo modello, che integra le variabili sul numero di figli e i fattori geografici, il coefficiente rimane infatti 0.0091 (errore standard 0.0121, $p = 0.451$). Nella specificazione *Full* (colonna 3), che include anche i diversi settori di impiego e l'intensità d'uso di internet, il coefficiente finale si attesta a 0.0097 (errore standard 0.0121, $p = 0.422$).

Dal punto di vista dell'interpretazione economica, poiché la variabile dipendente è espressa in logaritmi (*log_productivity_hour*), il coefficiente indicherebbe che il passaggio all'utilizzo del WFH determina una micro-variazione positiva della produttività oraria compresa appena tra lo 0.91% e lo 0.97% rispetto a chi lavora sempre in presenza. Tuttavia, l'assoluta assenza di significatività statistica impone di considerare questo millimetrico incremento come una mera oscillazione campionaria priva di validità scientifica: l'impatto reale dell'accesso allo smart working sulla produttività oraria è econometricamente pari a zero.

Sotto il profilo della teoria economica, questa evidenza suggerisce che l'adozione dello smart working non comporta alcun "costo di efficienza" per il sistema produttivo, smentendo le tesi tradizionali che associano il lavoro a distanza a un calo del rendimento orario dovuto a minori stimoli o a una riduzione del controllo manageriale diretto. D'altro canto, il risultato nullo indica che l'introduzione di modalità lavorative ibride non agisce nemmeno da shock tecnologico intrinsecamente espansivo per la produttività del singolo lavoratore. In sintesi, i benefici individuali derivanti dall'autonomia organizzativa e dal risparmio di tempo e stress legati al pendolarismo si bilanciano perfettamente con le frizioni

istituzionali e comunicative derivanti dalla distanza fisica dai locali aziendali.

Un ulteriore elemento di robustezza è la straordinaria stabilità del coefficiente nel corso delle tre stime sequenziali ($0.0091 \rightarrow 0.0091 \rightarrow 0.0097$). La costanza del parametro di interesse a fronte dell'inserimento di covariate fortemente correlate con la produttività (come il settore industriale e i canali tecnologici) rappresenta la prova del fatto che lo stimatore ad effetti fissi sta operando correttamente, assorbendo a monte l'eterogeneità non osservata e immunizzando la stima dal rischio di *omitted variable bias*.

4.1.2. Il ruolo dei controlli: shock temporali e impatto di internet

Un dettaglio importante che emerge dal modello completo riguarda il comportamento delle variabili di controllo. La quasi totalità di esse – come il numero di figli, la regione o il settore industriale – non risulta statisticamente significativa. Questo accade perché lo stimatore ad effetti fissi pulisce già il modello da gran parte delle differenze strutturali tra i lavoratori. Le uniche due eccezioni significative sono i macro-shock temporali (*wave*) e l'intensità d'uso di internet (*net_use*).

I controlli temporali (*wave*) mostrano un impatto positivo, forte e altamente significativo che cresce di anno in anno: rispetto alla *wave* 12, la produttività oraria sale del 3,9% nella *wave* 13, del 10,9% nella *wave* 14 e ben del 17,1% nella *wave* 15 ($p < 0.01$ per tutte). Questo trend lineare cattura l'effetto dell'inflazione, dei rinnovi contrattuali e del progresso tecnologico generale che ha coinvolto tutti i lavoratori del campione.

Per quanto riguarda l'uso di Internet (*net_use*), i risultati mostrano una significatività solo marginale (attorno al livello del 10%). Rispetto a chi è

costantemente connesso (*Almost all of the time*), chi usa la rete in modo più mirato – ad esempio solo una o due volte al giorno (+2,8%, $p = 0.081$) o qualche volta a settimana (+3,9%, $p = 0.068$) – mostra una produttività oraria leggermente superiore. Questo suggerisce che una minore esposizione alle distrazioni digitali e all’iper-connessione possa favorire la concentrazione, traducendosi in un lieve guadagno di efficienza oraria.

4.1.3. Analisi del margine intensivo: l’impatto della frequenza di WFH

Mentre l’analisi del margine estensivo si limita a verificare l’esistenza di un effetto medio legato all’accesso al WFH, lo studio del margine intensivo permette di indagare la presenza di eventuali non-linearità nella relazione tra lavoro da remoto e produttività oraria. È infatti plausibile che l’efficienza del lavoratore non risponda in modo dicotomico alla scelta di fare o meno smart working, bensì dipenda in modo cruciale dall’intensità della formula organizzativa adottata.

Per approfondire questa dimensione, la variabile principale di interesse viene disaggregata in una struttura categoriale esaustiva (*wfh_freq*). La categoria di riferimento è rappresentata dal gruppo *Never WFH* (lavoratori interamente in presenza). L’intensità del lavoro a distanza viene quindi frammentata in sei livelli crescenti:

1. 1-4 giorni a settimana (smart working occasionale);
2. 1 giorno a settimana;
3. 2 giorni a settimana;
4. 3 giorni a settimana (modelli ibridi a intensità variabile);
5. 4 giorni a settimana;
6. Sempre WFH (lavoro da remoto totale o strutturale).

L'approccio metodologico segue la stessa strategia a blocchi progressivi utilizzata in precedenza, stimando tre specificazioni a effetti fissi individuali con errori standard robusti e *clusterizzati* a livello di singolo rispondente.

Tabella 2 – WFH e produttività oraria: stime FE sul margine intensivo. Elaborazione autonoma su dati Understanding Society.

Variabile	Blocco 1: Base	Blocco 2: Geo & Famiglia	Blocco 3: Completo
Frequenza WFH (Rif: Never)	//	//	//
1-4 gg/mese	0.015 (0.018)	0.015 (0.018)	0.0147 (0.018)
1 g/settimana	0.008 (0.013)	0.008 (0.013)	0.010 (0.013)
2 gg/settimana	0.009 (0.020)	0.009 (0.020)	0.009 (0.020)
3 gg/settimana	0.006 (0.017)	0.005 (0.017)	0.006 (0.016)
4 gg/settimana	0.018 (0.026)	0.016 (0.024)	0.017 (0.024)
Sempre	-0.005 (0.018)	-0.004 (0.018)	-0.003 (0.018)
Area Geografica & Famiglia	//	//	//
Area rurale (Rif: Urban area)	–	-0.024 (0.040)	-0.025 (0.040)
Galles (Rif: Inghilterra)	–	-0.094 (0.058)	-0.098* (0.059)
Scozia	–	-0.432 (0.505)	-0.433 (0.504)
Irlanda del Nord	–	0.129 (0.146)	0.128 (0.145)
Numero di figli	–	-0.006 (0.018)	-0.007 (0.018)
Settore Industriale (Rif: Agriculture & Primary)	//	//	//
Manifattura	–	–	0.063 (0.095)
Edilizia	–	–	0.058 (0.099)
Commercio e Trasporti	–	–	0.049 (0.093)
Informazione e Comunicazione	–	–	-0.011 (0.105)
Finanza	–	–	0.048 (0.105)
Servizi Pubblici e Sociali	–	–	0.020 (0.096)
Altri Servizi	–	–	0.018 (0.105)
Uso Internet (Rif: Almost all of the time)	//	//	//
Diverse volte al giorno	–	–	-0.004 (0.009)
Una o due volte al giorno	–	–	0.028* (0.016)
Diverse volte a settimana	–	–	0.039* (0.021)
Diverse volte al mese	–	–	0.038 (0.026)
Una volta al mese	–	–	0.045 (0.034)

Meno di una volta al mese	–	–	0.091 (0.101)
Mai	–	–	0.103* (0.056)
Nessun accesso a internet	–	–	-0.031 (0.154)
Effetti Fissi Temporal (Rif: Wave 12)	//	//	//
Wave 13	0.039*** (0.009)	0.039*** (0.009)	0.039*** (0.009)
Wave 14	0.107*** (0.007)	0.107*** (0.007)	0.108*** (0.007)
Wave 15	0.170*** (0.008)	0.170*** (0.008)	0.171*** (0.008)
Costante	2.668*** (0.006)	2.714*** (0.049)	2.681*** (0.107)
<i>Osservazioni</i>	29,624	29,624	29,624
<i>R-quadrato (within)</i>	0.0346	0.0353	0.0361

4.1.4. Presentazione dei risultati empirici e assenza di significatività statistica

I risultati derivanti dalle stime longitudinali restituiscono un verdetto analogo a quello già visto per il margine estensivo: nessuna frequenza di WFH, dalla più occasionale alla più intensa, esercita un impatto statisticamente significativo sulla produttività oraria del lavoratore, in nessuno dei tre blocchi. Osservando i coefficienti dettagliati nel terzo modello, quello più completo e rigoroso, emergono le seguenti evidenze econometriche:

- Per lo smart working occasionale (1-4 giorni al mese), il coefficiente si attesta a 0.0147 con un errore standard di 0.0176 (p -value = 0.405).
- Nelle modalità ibride i parametri rimangono estremamente vicini allo zero: 0.0096 per “1 giorno a settimana” (p -value = 0.472), 0.0091 per “2 giorni a settimana” (p -value = 0.641) e 0.0059 per “3 giorni a settimana” (p -value = 0.720).
- Nella formule ad alta intensità da remoto (4 giorni a settimana), si osserva una leggera oscillazione positiva del coefficiente, pari a 0.0168, ma l’errore standard è comunque di 0.0244, mantenendo il parametro ampiamente non significativo (p -value = 0.492).

- Infine, per i lavoratori che operano interamente a distanza (Sempre), il coefficiente inverte il segno diventando negativo e pari a -0.0029 con un errore standard di 0.0182 ($p\text{-value} = 0.872$).

Sotto il profilo statistico, i test t associati a ciascun coefficiente non riescono mai a superare la soglia critica di rifiuto dell'ipotesi nulla. Di conseguenza, i risultati non possiedono alcuna rilevanza statistica e vanno interpretati come semplici fluttuazioni campionarie.

In perfetta coerenza con quanto riscontrato sul margine estensivo, anche nella stima del margine intensivo la quasi totalità dei coefficienti delle variabili di controllo (famiglia, area geografica, settore) rimane statisticamente non significativa, a eccezione di quelle temporali, due relative all'intensità d'uso di internet e una, quella del Galles, geografica.

4.1.5. Ragioni teoriche della non significatività diffusa

Di fronte a due tabelle in cui tutti i coefficienti di interesse risultano statisticamente non significativi, il primo impulso metodologico potrebbe essere quello di sospettare un errore nell'impostazione del modello o un'incapacità dei dati di stimare il fenomeno. In realtà, nel contesto dei modelli panel ad effetti fissi individuali, questa diffusa non significatività rappresenta un risultato econometrico perfettamente normale, solido e dotato di una precisa motivazione scientifica. Essa non indica un fallimento della stima, ma dimostra empiricamente che, una volta ripulito il modello dall'eterogeneità individuale non osservata (come l'abilità innata o la motivazione del lavoratore), nessun livello di intensità da remoto risulta intrinsecamente superiore o inferiore al lavoro interamente in presenza. Le variazioni nel tempo della produttività oraria non sono guidate dal “dosaggio” settimanale o mensile di lavoro a distanza, bensì dai macro-shock temporali (*wave*) e da fattori legati alla familiarità con il digitale (*net_use*).

Sotto il profilo strettamente interpretativo, questo scenario evidenzia come il WFH non agisca come un fattore di shock tecnologico capace di spostare la frontiera di efficienza del singolo individuo, ma si configuri piuttosto come una modalità organizzativa a rendimenti costanti. Il fatto che nessuna frequenza modifichi in modo significativo i coefficienti dimostra che lo svolgimento delle mansioni quotidiane segue logiche e ritmi su cui la distanza fisica dalla sede aziendale non esercita particolari influenze. Di conseguenza, le differenze di produttività osservate nel mercato del lavoro non dipendono dallo spostamento fisico del lavoratore o dai giorni trascorsi a casa, ma rimangono saldamente ancorate alle caratteristiche strutturali dell'individuo – che gli effetti fissi assorbono – e alle dinamiche congiunturali esterne che colpiscono in modo omogeneo l'intero campione.

4.2. Stima dei modelli OLS (margine estensivo e intensivo)

A seguito dell'analisi condotta sul nostro benchmark a effetti fissi, risulta di fondamentale interesse metodologico osservare come variano le stime quando si decide di allentare i vincoli econometrici legati all'identità del singolo lavoratore. A questo scopo, in questa sezione si introduce l'analisi condotta tramite stimatore *pooled* OLS sul blocco completo dei controlli, operando una sintesi analitica tra il margine estensivo e quello intensivo.

Spostando il focus dal benchmark a questi modelli, abbandoniamo lo studio della variazione intrapersonale temporale per analizzare a pieno la variazione interpersonale. L'obiettivo non è più osservare cosa accade alla produttività dello stesso soggetto quando modifica il proprio status lavorativo nel tempo, ma confrontare soggetti diversi tra loro all'interno del campione (sebbene questa eterogeneità risulti in parte “simulata” dallo stimatore *pooled*, il quale tratta come indipendenti e distinti anche i medesimi individui che si ripetono longitudinalmente nelle diverse *waves*

del panel), scattando una fotografia strutturale e cumulativa delle differenze cross-section. La tabella 3 presentata di seguito raccoglie gli esiti di questa seconda strategia di stima.

Tabella 3 – WFH e produttività oraria: stime OLS sul margine estensivo.
Elaborazione autonoma su dati Understanding Society.

Variabile	Blocco 1: Base	Blocco 2: Geo & Famiglia	Blocco 3: Completo
Usa WFH (Rif: <i>Never</i>)	0.266*** (0.007)	0.263*** (0.007)	0.235*** (0.007)
Controlli Demografici	//	//	//
Età	0.045*** (0.002)	0.047*** (0.002)	0.045*** (0.002)
Età ²	-0.000*** (0.000)	-0.000*** (0.000)	-0.000*** (0.000)
Femmina (Rif: <i>Male</i>)	-0.172*** (0.006)	-0.165*** (0.007)	-0.138*** (0.007)
Livello Istruzione (Rif: <i>No qualification</i>)	//	//	//
Altre qualifiche	0.051* (0.029)	0.047 (0.029)	0.039 (0.029)
GCSE (scuola dell'obbligo)	0.079*** (0.027)	0.073*** (0.027)	0.060** (0.028)
A-level (maturità)	0.171*** (0.027)	0.164*** (0.027)	0.148*** (0.028)
Teaching/HND (qualifiche tecniche)	0.255*** (0.028)	0.248*** (0.028)	0.230*** (0.028)
Laurea	0.444*** (0.027)	0.438*** (0.027)	0.425*** (0.028)
Area Geografica & Famiglia	//	//	//
Area rurale (Rif: <i>Urban area</i>)	–	0.013* (0.007)	0.014** (0.007)
Galles (Rif: <i>Inghilterra</i>)	–	-0.053*** (0.012)	-0.050*** (0.012)
Scozia	–	0.019** (0.009)	0.020** (0.009)
Irlanda del Nord	–	-0.056*** (0.011)	-0.056*** (0.012)
Numero di figli	–	-0.013*** (0.005)	-0.012** (0.005)
Settore Industriale (Rif: <i>Agriculture & Primary</i>)	//	//	//
Manifattura	–	–	0.189*** (0.045)
Edilizia	–	–	0.217*** (0.046)
Commercio e Trasporti	–	–	0.051 (0.045)
Informazione e Comunicazione	–	–	0.231*** (0.048)
Finanza	–	–	0.273*** (0.047)
Servizi Pubblici e Sociali	–	–	0.103** (0.045)
Altri Servizi	–	–	-0.014 (0.046)
Uso Internet (Rif: <i>Almost all of the time</i>)	//	//	//
Diverse volte al giorno	–	–	0.006 (0.007)

Una o due volte al giorno	–	–	-0.023** (0.011)
Diverse volte a settimana	–	–	-0.047*** (0.018)
Diverse volte al mese	–	–	-0.075*** (0.028)
Una volta al mese	–	–	-0.098** (0.044)
Meno di una volta al mese	–	–	0.018 (0.078)
Mai	–	–	-0.062 (0.041)
Nessun accesso a internet	–	–	-0.335*** (0.090)
Effetti Fissi Temporal (Rif: Wave 12)	//	//	//
Wave 13	0.002 (0.010)	0.002 (0.010)	0.007 (0.009)
Wave 14	0.075*** (0.008)	0.074*** (0.008)	0.077*** (0.008)
Wave 15	0.134*** (0.009)	0.133*** (0.009)	0.137*** (0.009)
Costante	1.405*** (0.049)	1.380*** (0.050)	1.300*** (0.066)
<i>Osservazioni</i>	29,624	29,624	29,624
<i>R-quadrato</i>	0.2187	0.2201	0.2316

4.2.1. Il premio di produttività nel margine estensivo

Come riportato nella tabella 3, il coefficiente associato alla dummy *Uses WFH* nel blocco 3 è pari a 0.235, fortemente significativo dal punto di vista statistico ($p < 0.01$). Trattandosi di un modello log-lineare, l'effetto percentuale netto sulla produttività oraria si ottiene calcolando $e^{0.2351} - 1$, che si traduce in un imponente +26,5% a favore di chi dichiara di utilizzare il lavoro da casa almeno un giorno al mese rispetto a chi lavora esclusivamente in presenza (*Never WFH*).

Dal punto di vista dell'efficienza dei lavoratori, questo importante premio iniziale nei modelli OLS sembra catturare un duplice effetto. Da un lato, riflette i potenziali vantaggi reali dello smart working in termini di abbattimento dei costi di transizione e di pendolarismo, che si traducono in una maggiore disponibilità di energie cognitive e in una riduzione dei tempi morti durante la giornata lavorativa. Dall'altro lato, un divario così elevato nel confronto cross-section suggerisce immediatamente la presenza di un forte sbilanciamento nella composizione del campione: chi

fa smart working appartiene a segmenti occupazionali strutturalmente più avanzati.

La persistenza del coefficiente a livelli così alti, anche dopo aver inserito l'intero blocco dei controlli settoriali, geografici e tecnologici, dimostra a prima vista che la flessibilità spaziale non è un fenomeno neutrale, ma si inserisce come una determinante fondamentale della produttività individuale all'interno del Regno Unito.

4.2.2. La non-linearità nel margine intensivo

Tabella 4 – WFH e produttività oraria: stime OLS sul margine intensivo.
Elaborazione autonoma su dati Understanding Society.

Variabile	Blocco 1: Base	Blocco 2: Geo & Famiglia	Blocco 3: Completo
Frequenza WFH (Ref: Never)	//	//	//
1-4 gg/mese	0.260*** (0.013)	0.258*** (0.013)	0.241*** (0.012)
1 gg/settimana	0.244*** (0.012)	0.242*** (0.012)	0.224*** (0.012)
2 gg/settimana	0.282*** (0.014)	0.279*** (0.014)	0.245*** (0.014)
3 gg/settimana	0.300*** (0.013)	0.298*** (0.013)	0.265*** (0.013)
4 gg/settimana	0.304*** (0.018)	0.301*** (0.018)	0.265*** (0.018)
Sempre	0.248*** (0.013)	0.245*** (0.013)	0.199*** (0.013)
Controlli Demografici	//	//	//
Età	0.045*** (0.002)	0.047*** (0.002)	0.045*** (0.002)
Età ²	-0.000*** (0.000)	-0.000*** (0.000)	-0.000*** (0.000)
Femmina (Ref: Maschio)	-0.172*** (0.006)	-0.165*** (0.007)	-0.138*** (0.007)
Livello Istruzione (Ref: No qualification)	//	//	//
Other qualification	0.051* (0.029)	0.047 (0.029)	0.039 (0.029)
GCSE (scuola dell'obbligo)	0.079*** (0.027)	0.073*** (0.027)	0.060*** (0.028)
A-level (maturità)	0.171*** (0.027)	0.164*** (0.027)	0.148*** (0.028)
Teaching/HND (qualifiche tecniche)	0.255*** (0.028)	0.248*** (0.028)	0.230*** (0.028)
University degree	0.444*** (0.027)	0.438*** (0.027)	0.425*** (0.028)
Area Geografica & Famiglia	//	//	//
Area rurale (Ref: Urban area)	—	0.013* (0.007)	0.014** (0.007)
Galles (Ref: Inghilterra)	—	-0.052*** (0.012)	-0.050*** (0.012)
Scozia	—	0.020** (0.009)	0.020** (0.009)
Irlanda del Nord	—	-0.056*** (0.012)	-0.056*** (0.012)

Numero di figli	—	-0.013*** (0.005)	-0.012** (0.005)
Settore Industriale (Ref: Agriculture & Primary)	//	//	//
Manufacturing	—	—	0.187*** (0.045)
Construction	—	—	0.214*** (0.046)
Trade & Transport	—	—	0.049 (0.045)
Information & Communication	—	—	0.235*** (0.048)
Finance	—	—	0.272*** (0.047)
Public & Social Services	—	—	0.101*** (0.045)
Other Services	—	—	-0.016 (0.046)
Uso Internet (Ref: Almost all of the time)	//	//	//
Several times a day	—	—	0.006 (0.007)
Once or twice a day	—	—	-0.023** (0.011)
Several times a week	—	—	-0.047*** (0.018)
Several times a month	—	—	-0.075*** (0.028)
Once a month	—	—	-0.099** (0.044)
Less than once a month	—	—	0.019 (0.078)
Never use	—	—	-0.063 (0.041)
Categoria 9	—	—	-0.337*** (0.090)
Effetti Fissi Temporal (Ref: Wave 12)	//	//	//
Wave 13	0.002 (0.010)	0.002 (0.010)	0.005 (0.010)
Wave 14	0.075*** (0.008)	0.074*** (0.008)	0.073*** (0.008)
Wave 15	0.134*** (0.009)	0.133*** (0.009)	0.133*** (0.009)
Costante	1.406*** (0.049)	1.380*** (0.050)	1.304*** (0.066)
<i>Osservazioni</i>	29,624	29,624	29,624
<i>R-quadrato</i>	0.2187	0.2201	0.2316

Il modello del margine intensivo arricchisce e suddivide il risultato del margine estensivo, spaccettando la dummy generica nella frequenza effettiva di fruizione del WFH. Questo passaggio rappresenta un avanzamento analitico cruciale rispetto ai blocchi precedenti, poiché permette di testare se l'associazione positiva tra flessibilità spaziale e produttività oraria sia legata a una specifica soglia di giorni spesi da remoto o se si configuri come un fenomeno diffuso lungo l'intero spettro delle frequenze di WFH. I risultati mostrano che tutti i coefficienti legati alle diverse intensità di smart working mantengono la massima

significatività statistica ($p < 0.01$), rivelando un quadro cross-section estremamente coerente.

L'analisi dettagliata delle singole classi di frequenza del blocco 3 evidenzia i seguenti premi percentuali netti sulla produttività oraria:

- *Frequenze occasionali (1-4 giorni al mese)*: si associano a un incremento del +27,3% ($e^{0.2413} - 1$).
- *Frequenze intermedie (1-2 giorni a settimana)*: registrano rispettivamente un premio del +25,1% e del +29,3%.
- *Frequenze elevate (3-4 giorni a settimana)*: mostrano incrementi del +30,4% (per i 3 giorni) e del +30,3% (per i 4 giorni).
- *Il regime di full remote (Always)*: si attesta su un premio del +22,0% ($e^{0.1994} - 1$).

Sotto il profilo strettamente econometrico, un'ispezione rigorosa dei risultati impone una cautela fondamentale: le pur visibili oscillazioni dei coefficienti puntuali – che sembrerebbero a prima vista suggerire un picco di efficienza nelle modalità ibride ad alta intensità – non risultano quasi mai statisticamente significative tra loro. Gli intervalli di confidenza delle stime si sovrappongono ampiamente lungo quasi tutte le classi di frequenza; di conseguenza, dal punto di vista statistico, non è possibile rifiutare l'ipotesi nulla di uguaglianza tra i coefficienti del lavoro ibrido.

Questo implica che, nel confronto cross-section tra individui diversi, il vantaggio competitivo in termini di produttività oraria non è strettamente vincolato a un preciso “dosaggio” dei giorni trascorsi a casa. L'effetto positivo associato al WFH si manifesta in modo solido e sostanzialmente uniforme, segnalando che il mero sblocco della flessibilità – sia esso fruito in modo sporadico o come asse portante della settimana lavorativa – sia sufficiente a migliorare la produttività individuale.

4.2.3. Il ruolo delle caratteristiche demografiche e del capitale umano

L'utilità metodologica del modello OLS emerge con altrettanta forza dall'analisi delle variabili di controllo fisse nel tempo o a bassa variabilità intra-personale, che lo stimatore ad effetti fissi tendeva a neutralizzare. I coefficienti di queste variabili mostrano un comportamento altamente coerente tra i due modelli (estensivo e intensivo), confermando la stabilità strutturale delle stime complessive.

- *Il capitale umano (University Degree)*: il possesso di una laurea mostra il coefficiente più grande e dominante dell'intera regressione (0.4250 nel blocco 3 dell'intensivo, con oscillazioni irrilevanti nell'estensivo, $p < 0.01$). Questo dato si traduce in un premio produttivo del +52,9% rispetto a chi non possiede qualifiche formali. Il risultato è in perfetta linea con la teoria: l'istruzione eleva drasticamente la produttività marginale del lavoro, fungendo da principale determinante dei divari di rendimento nel mercato del lavoro britannico.
- *Il genere (Female)*: emerge un marcato e penalizzante *gender productivity gap* pari al -12,9% ($e^{-0.1384} - 1$, $p < 0.01$), che rimane rigidamente identico in entrambe le specificazioni OLS. Nel confronto tra individui diversi, a parità di istruzione, settore occupazionale, *wave* temporale e intensità d'uso di internet, le donne registrano una produttività oraria sistematicamente inferiore. Questo coefficiente, ripulito dall'effetto del WFH, denuncia la persistenza di profonde asimmetrie strutturali che i modelli OLS riescono a quantificare con precisione.
- *Il ciclo della vita (Age e Age²)*: i coefficienti confermano la classica concavità della produttività rispetto all'età, mostrando una perfetta

coerenza tra le due colonne di stima. Il segno positivo su *Age* (+4,4%) e quello negativo, seppur millesimale, sul termine quadratico *Age*² indicano che la produttività cresce con l'esperienza e l'età biologica, ma a tassi decrescenti. Il profilo segue la tipica traiettoria a parabola rovesciata: il rendimento orario aumenta nelle fasi giovanili e di maturità grazie all'accumulazione di *learning-by-doing*, per poi raggiungere un punto di inversione e flettere fisiologicamente nella parte finale della carriera lavorativa.

4.2.4. Interpretazione critica del quadro OLS: l'effetto selezione

Se ci fermassimo ai soli risultati OLS, la conclusione economico-politica sarebbe immediata: lo smart working è una tecnologia organizzativa straordinaria che aumenta la produttività di chiunque la usi, specialmente se concessa per 3 o 4 giorni a settimana.

Tuttavia, applicando il nostro spirito critico, dobbiamo chiederci se questi modelli stanno davvero stimando un impatto *causale* e la risposta è no. La variazione interpersonale qui analizzata è affetta da una profonda *distorsione da selezione positiva (eterogeneità non osservata)*.

Nel mercato del lavoro reale, l'accesso allo smart working non viene assegnato casualmente: le mansioni che permettono il WFH sono strutturalmente più complesse, digitalizzate e ad alto valore aggiunto. Di conseguenza, i lavoratori che “scelgono” o ottengono il WFH possiedono abilità non osservate nei dati – come maggiore autodisciplina, *soft skills* avanzate, spiccate capacità di *problem solving* o semplicemente una posizione contrattuale più forte – che li renderebbero enormemente più produttivi della media *a prescindere* dal luogo fisico in cui lavorano. L'OLS, accumulando queste differenze tra persone diverse, scatta una

fotografia di questa élite lavorativa, confondendo la selezione all'ingresso con il beneficio reale della modalità di lavoro.

4.3. OLS e FE: il confronto metodologico

Il vero fulcro analitico della tesi emerge dalla comparazione sistematica tra i risultati derivanti dal *pooled* OLS e quelli generati dal benchmark metodologico a effetti fissi (FE). Mettendo a confronto le due strategie di stima, si assiste a una radicale divergenza sia nella significatività statistica sia nell'entità dei coefficienti legati al WFH. Questa discrepanza evidenzia come la scelta dello stimatore possa condurre a conclusioni diametralmente opposte circa l'impatto reale della flessibilità organizzativa sulla produttività.

Come analizzato in precedenza, nel modello FE i coefficienti sia del margine estensivo sia del margine intensivo subiscono una drastica contrazione, azzerandosi e perdendo completamente significatività statistica. Al contrario, l'approccio OLS restituisce stime fortemente significative ($p < 0.01$), con impatti positivi che sfiorano il +30% di produttività oraria per i regimi ibridi.

Perché si verifica questa apparente contraddizione statistica? La spiegazione risiede nella diversa componente di varianza catturata dai due stimatori, che possiamo riassumere in tre pilastri analitici:

1. *La natura della varianza (confronto cross-section vs Variazione nel tempo)*. Lo stimatore *pooled* OLS sfrutta l'intera variabilità del campione, basando le proprie stime principalmente sulla varianza *between* (ossia sulle differenze tra individui diversi). Quando l'OLS confronta un lavoratore che usufruisce del WFH con uno che lavora esclusivamente in presenza, tende ad attribuire l'intero scarto di produttività alla modalità di lavoro. Al contrario, lo stimatore FE isola esclusivamente la varianza *within* (la variazione

intra-individuale), confrontando il medesimo soggetto con se stesso prima e dopo aver cambiato regime lavorativo. Il crollo dei coefficienti nel modello FE dimostra che, a parità di individuo, lo shock organizzativo dello smart working non genera di per sé alcun guadagno automatico di efficienza oraria.

2. *L'effetto selezione e l'eterogeneità non osservata.* La discrepanza tra i due stimatori rileva con precisione la presenza di una forte distorsione da selezione positiva nei modelli OLS. Come spiegato nel presente capitolo, nel mercato del lavoro, l'accesso allo smart working non è assegnato casualmente, ma è correlato a caratteristiche individuali non osservabili nei dati, come l'abilità innata, l'autodisciplina o una maggiore forza contrattuale. L'OLS cattura queste doti latenti, confondendo lo status di "smart worker" con l'elevato profilo di partenza di chi riesce a ottenerlo. Il modello FE, pulendo le stime da tutte le caratteristiche fisse nel tempo, svela come il premio osservato nell'OLS fosse in gran parte un effetto di composizione del campione: non è il WFH a rendere più produttivi i lavoratori, ma sono i lavoratori strutturalmente più produttivi ad avere accesso al WFH.
3. *La quota di variabilità interna al panel.* L'azzeramento della significatività nel modello FE solleva anche una questione legata alla struttura informativa dei dati. Per stimare un coefficiente preciso, il modello a effetti fissi ha bisogno di un numero sufficiente di "transizioni", ovvero di individui che modificano effettivamente il proprio status lavorativo nel corso degli anni osservati. Se la maggior parte dei lavoratori nel campione mantiene una posizione rigida nel tempo (rimanendo stabilmente in presenza o stabilmente da remoto), la varianza *within* si riduce drasticamente. Di conseguenza, gli errori standard si ampliano e il risultato perde

significatività statistica, suggerendo che il legame tra WFH e produttività sia più un fenomeno allocativo e strutturale che uno immediato di efficienza marginale legato alla prestazione del singolo.

Conclusioni

L'analisi empirica condotta in questo lavoro di tesi ha voluto offrire un contributo alla valutazione, nel tempo e per singoli giorni della settimana, dell'impatto dello smart working sulla produttività oraria del lavoro nel Regno Unito, sfruttando la ricchezza di dati del panel *Understanding Society* (UKHLS). Il ricorso alla specificazione a effetti fissi individuali, oltre all'uso dei modelli esplorativi OLS, ha rappresentato il fulcro dello studio. Questa strategia ha permesso infatti di ridurre notevolmente gli errori dovuti all'autoselezione e di tenere sotto controllo le caratteristiche personali dei lavoratori che non cambiano nel tempo – come l'ambizione o l'abilità innata – che altrimenti avrebbero distorto i risultati del confronto. Bisogna comunque riconoscere che questo approccio economico, per quanto solido, ha dei limiti: non può eliminare del tutto il rischio di escludere fattori importanti che cambiano nel tempo, né i potenziali errori legati al fatto che alcuni dati derivano da risposte fornite direttamente dai lavoratori nei questionari.

I risultati emersi dalle stime descrivono una generale neutralità economica. Né l'accesso generico al lavoro da remoto (margine estensivo), né il numero di giorni spesi a casa durante la settimana (margine intensivo) mostrano un impatto statistico significativo sulla produttività oraria dei dipendenti. Questa assenza di variazioni, rimasta costante anche cambiando le configurazioni dei modelli e inserendo diverse variabili di controllo, richiede una riflessione dal punto di vista della teoria economica. Il fatto che non emerga un effetto netto non significa che il WFH sia irrilevante per l'efficienza, ma suggerisce piuttosto che benefici e svantaggi tendano a bilanciarsi e ad annullarsi a vicenda quando si guarda al quadro generale.

Da un lato, infatti, gli studi internazionali documentano ampiamente i vantaggi dello smart working: eliminare i tempi e i costi degli spostamenti

quotidiani riduce lo stress e la stanchezza, migliorando il benessere di chi lavora (Bloom et al., 2015; Barrero et al., 2023). Inoltre, poter gestire il tempo con più flessibilità e avere meno interruzioni rispetto all'ufficio aiuta a concentrarsi meglio e a sfruttare al massimo i propri ritmi di lavoro.

Dall'altro lato, però, emergono diverse difficoltà organizzative dovute alla distanza fisica (Gibbs et al., 2023; Emanuel e Harrington, 2024). L'isolamento logistico può rallentare la comunicazione interna, ad esempio trasformando un rapido confronto verbale in una catena di e-mail. Inoltre, la mancanza di scambi informali tra colleghi rischia di impoverire la condivisione spontanea delle conoscenze nel lungo periodo, mentre la minore supervisione costringe i manager a passare da un controllo visivo a una gestione per obiettivi, un passaggio complesso che richiede strumenti adeguati. I risultati della tesi indicano che, nel campione analizzato, questi due piatti della bilancia si equivalgono. Questo ridimensiona sia l'idea eccessivamente ottimista per cui il WFH aumenti senza dubbio la produttività, sia i timori opposti di chi temeva cali drastici o cali dell'impegno per la mancanza di controllo.

In conclusione, pur con la cautela necessaria quando si analizzano dati di questo tipo, la ricerca suggerisce che il lavoro da remoto nel Regno Unito si sia consolidato come una trasformazione contrattuale sostenibile, in grado di mantenere i livelli di efficienza preesistenti pur cambiando le abitudini e le giornate dei lavoratori. Questo impatto nullo, in linea con gli studi internazionali più recenti (Anakpo et al., 2023; Morikawa, 2024), offre spunti utili per il futuro. Le scelte dei manager e le politiche pubbliche non dovrebbero concentrarsi su regole rigide basate solo sul “conteggio” dei giorni a casa o in ufficio. L'attenzione dovrebbe invece spostarsi su come riorganizzare le mansioni – distinguendo i compiti solitari che richiedono concentrazione da quelli collaborativi (Jiang et al., 2024) – e sul potenziamento delle infrastrutture digitali e delle competenze

tecnologiche dei dipendenti, che è l'unico fattore ad aver mostrato una validità statistica concreta – seppur parziale – all'interno dei modelli presentati.

Bibliografia

- Acemoglu D. e Autor D. (2011). [Skills, Tasks and Technologies: Implications for Employment and Earnings](#). *Handbook of Labor Economics*, 4, 1045-1070.
- Aksoy C. G., Barrero J. M., Bloom N., Davis S. J., Dolls M. e Zarate P. (2022). [Working from Home Around the World](#). *NBER Working Paper Series*, 30446, National Bureau of Economic Research.
- Anakpo G., Nqwayibana Z. e Mishi S. (2023). [The Impact of Work-from-Home on Employee Performance and Productivity: A Systematic Review](#). *Sustainability*, 15(5), 4529.
- Andrew A., Cattan S., Costa Dias M., Farquharson C., Kraftman L., Krutikova S., Phimister A. e Sevilla A. (2022). [The gendered division of paid and domestic work under lockdown](#). *Fiscal Studies*, 43(4), 325–340.
- Angelici M. e Profeta P. (2020). [Smart-working: Work Flexibility Without Constraints](#). *Dondena Working Paper Series*, (137), 1-8.
- Bao L., Li T., Xia X., Zhu K., Li H. e Yang X. (2022). [How does working from home affect developer productivity? – A case study of Baidu during the COVID-19 pandemic](#). *Science China Information Sciences*, 65(4), 142102.
- Barrero J. M., Bloom N. e Davis S. J. (2021). [Why Working from Home Will Stick](#). *NBER Working Paper Series*, 28731, National Bureau of Economic Research.
- Barrero J. M., Bloom N. e Davis S. J. (2023). [The Evolution of Work from Home](#). *Journal of Economic Perspectives*, 37(4), 23-50.
- Barrero J. M., Bloom N., Bonney K., Breaux C., Buffington C., Davis S. J., Foster L., McKenzie B., Savage K. e Tello-Trillo C. (2025). [Tapping Business and Household Surveys to Sharpen Our](#)

[View of Work from Home](#). *Center for Economic Studies (CES) Working Paper Series*, CES-25-36, U.S. Census Bureau.

- Bloom N., Liang J., Roberts J. e Ying Z. J. (2015). [Does Working from Home Work? Evidence from a Chinese Experiment](#). *Quarterly Journal of Economics*, 130(1), 165-218.
- Buck N. e McFall S. (2012). [Understanding Society: design overview](#). *Longitudinal and Life Course Studies*, 3(1), 5–17.
- Dingel J. I. e Neiman B. (2020). [How Many Jobs Can be Done at Home?](#) *NBER Working Paper Series*, (26948).
- Emanuel N. e Harrington E. (2024). [Working Remotely? Selection, Treatment, and the Market for Remote Work](#). *American Economic Journal: Applied Economics*, 16(4), 528–559.
- Frank M. R., Autor D., Bessen J. E., Brynjolfsson E., Deming D. J., Feldman M., Groh M., Lobo J., Moro E., Wang D., Youn H. e Rahwan I. (2019). [Toward understanding the impact of artificial intelligence on labor](#). *Proceedings of the National Academy of Sciences*, 116(14), 6531–6539.
- Gibbs M., Mengel F. e Siemroth C. (2023). [Work from Home and Productivity: Evidence from Personnel and Analytics Data on Information Technology Professionals](#). *Journal of Political Economy Microeconomics*, 1(1), 7–41.
- Hackney A., Yung M., Somasundram K. G., Nowrouzi-Kia B., Oakman J. e Yazdani A. (2022). [Working in the Digital Economy: A Systematic Review of the Impact of Work-from-Home Arrangements on Personal and Organizational Performance and Productivity](#). *PLoS One*, 17(10), e0274728.
- Jiang M., Yasui K. e Yugami K. (2024). [Working from home, job tasks, and productivity](#). *Telecommunications Policy*, 48(8), 102806.

- Marzec P., Nandi A. e Patel R. (2021). [Understanding Society COVID-19 Survey briefing note: working at home](#). *Understanding Society Briefing Notes*.
- Morikawa M. (2024). [Productivity dynamics of work from home: Firm-level evidence from Japan](#). *Journal of Evolutionary Economics*, 34(2), 465–487.
- Raitano, M. (2021). [Distribuzione e disuguaglianza retributiva in Italia: dinamiche e implicazioni per il sistema pensionistico](#). *Fondazione Giuseppe Di Vittorio – SPI CGIL*, 1–6.
- Wooldridge J. M. (2010). [Econometric Analysis of Cross Section and Panel Data](#). MIT Press, Cambridge, MA, 247-297.
- Yang L., Jaffe S., Holtz D., Suri S., Sinha S., Weston J., Joyce C., Shah N., Sherman K., Lee C. J., Hecht B. e Teevan J. (2020). [How Work From Home Affects Collaboration Among Information Workers: A Natural Experiment During COVID-19](#). arXiv preprint arXiv:2007.15584.
- Zarate P., Dolls M., Davis S. J., Bloom N., Barrero J. M. e Aksoy C. G. (2024). [Why Does Working from Home Vary Across Countries and People?](#) *NBER Working Paper Series*, 32374, National Bureau of Economic Research.

Sitografia

- Atlas of Longitudinal Datasets (n.d.). [Understanding Society, the UK Household Longitudinal Study \(UKHLS\)](#).
- UK Longitudinal Linkage Collaboration (UK LLC) (n.d.). [Understanding Society, the UK Household Longitudinal Survey & British Household Panel Survey \(UKHLS\) Datasets](#). *UK LLC Guidebook*.

- Understanding Society (n.d.). [Understanding Society: Mainstage Survey Timeline](#); [Mainstage Variables Online Documentation](#).
Institute for Social and Economic Research, University of Essex.