

Istruzione, salari ed eterogeneità delle imprese: evidenze da dati employer-employee linked

Irene Brunetti
INAPP

Michelangelo Filippi
Laboratorio Revelli

Valentina Gualtieri
INAPP

Andrea Ricci
INAPP

Lo studio analizza come le politiche retributive delle imprese contribuiscano ai rendimenti salariali dell'istruzione, considerando eterogeneità delle imprese. L'analisi utilizza un dataset integrato che combina Uniemens (Inps), Comunicazioni Obbligatorie (MLPS) e dati RIL Inapp, applicando una strategia econometrica a due stadi basata su modelli AKM e regressioni ad effetti fissi. I risultati mostrano che le politiche salariali influenzano positivamente, seppur debolmente, i premi dei laureati; che i differenziali retributivi riflettono processi di sorting non uniformi lungo la distribuzione del capitale umano; e che le imprese trasferiscono solo parzialmente e nel medio periodo i guadagni di produttività ai lavoratori più qualificati.

Education, wages, and firm heterogeneity: evidence from linked employer-employee data

The study examines how firms' wage setting practices shape the wage returns to education, accounting for firm's heterogeneity. Using a unique dataset combining Uniemens (Inps), COB (MLPS) with RIL-Inapp firm data, a two stage AKM fixed effects strategy shows that firm pay policies weakly raise graduates' wage premiums; sorting patterns differ across human capital levels; and productivity gains are only partly and slowly passed on to highly educated workers.

DOI: 10.53223/Sinappsi_2026-02-6

Citazione	Parole chiave	Keywords
Brunetti I., Filippi M., Gualtieri V., Ricci A. (2026), Istruzione, salari ed eterogeneità delle imprese: evidenze da dati employer-employee linked, <i>Sinappsi</i> , XVI, n.2, pp.91-105	Employer-employee data Imprese Politiche salariali	<i>Employer-employee data</i> <i>Enterprises</i> <i>Wage setting</i>

Introduzione

La persistente stagnazione delle retribuzioni reali rappresenta uno dei principali temi del dibattito economico e istituzionale sul mercato del lavoro italiano (OECD 2024; Inps 2024; Banca d'Italia 2025). Parallelamente, il crescente interesse della letteratura per le determinanti della dispersione salariale ha riportato al centro dell'analisi il ruolo delle imprese nella formazione dei differenziali retributivi. Numerosi studi evidenziano come la persistente debolezza della produttività del lavoro

e della produttività totale dei fattori, unitamente ad altri fattori strutturali, abbia limitato la capacità del sistema produttivo italiano di sostenere la crescita economica e la dinamica delle retribuzioni (Andrews *et al.* 2016; Bugamelli *et al.* 2018). In questo contesto, il tema dei rendimenti salariali dell'istruzione assume particolare rilevanza, considerando il ruolo sempre più centrale del capitale umano e delle competenze nell'accompagnare le transizioni tecnologica, demografica e ambientale che stanno interessando il sistema economico.

I rendimenti dell'istruzione costituiscono da tempo uno dei principali ambiti di studio dell'economia del lavoro. Secondo la teoria del capitale umano (Becker 1964), il premio salariale associato a un maggiore livello di istruzione riflette l'incremento della produttività individuale derivante dall'investimento in formazione. La teoria del *signaling* (Spence 1973), invece, interpreta il titolo di studio come un segnale credibile delle abilità individuali osservabile dai datori di lavoro. Pur muovendo da presupposti differenti, entrambe le impostazioni attribuiscono un ruolo sostanzialmente marginale all'impresa nella determinazione dei rendimenti dell'istruzione: a parità di caratteristiche individuali, lavoratori con lo stesso livello di istruzione dovrebbero ricevere salari simili, indipendentemente dall'impresa presso cui sono occupati.

Negli ultimi due decenni questa interpretazione è stata progressivamente rivista grazie alla crescente disponibilità di basi dati amministrative longitudinali lavoratore-impresa (*linked employer-employee data*), che hanno consentito di osservare simultaneamente le caratteristiche dei lavoratori, delle imprese e delle loro relazioni nel tempo. Abowd *et al.* (1999) hanno introdotto il modello a effetti fissi di lavoratore e impresa (AKM), mostrando come una quota rilevante della variabilità salariale possa essere attribuita a componenti specifiche dell'impresa, oltre che alle caratteristiche individuali. Successivi contributi hanno evidenziato che l'eterogeneità tra imprese rappresenta una determinante fondamentale della dispersione salariale complessiva (Card *et al.* 2013; Song *et al.* 2019). Inoltre, Card *et al.* (2018) mostrano come i processi di *assortative matching* tra lavoratori e imprese contribuiscano significativamente alla disuguaglianza salariale, poiché i lavoratori con caratteristiche produttive migliori tendono a concentrarsi nelle imprese che riconoscono premi salariali più elevati. Questa letteratura suggerisce che anche i rendimenti salariali dell'istruzione possano dipendere, almeno in parte, dalle caratteristiche delle imprese presso cui i lavoratori trovano occupazione. In particolare, l'accesso dei laureati a imprese caratterizzate da premi salariali sistematicamente più elevati potrebbe spiegare una quota dei differenziali retributivi normalmente attribuiti esclusivamente all'investimento in istruzione. In questa prospettiva, i rendimenti dell'istruzione rifletterebero non soltanto differenze nella produttività individuale, ma anche il diverso accesso a imprese con differenti politiche di wage-setting. L'evidenza empirica più recente si muove in questa direzione. Song *et al.* (2019) mostrano che l'aumento della disuguaglianza salariale negli Stati Uniti è riconducibile in larga misura alla crescente dispersione salariale tra imprese e ai cambiamenti nei meccanismi di allocazione dei

lavoratori. Più recentemente, Pan (2025) evidenziano come una quota non trascurabile dei rendimenti dell'istruzione terziaria derivi dall'accesso a imprese che riconoscono premi salariali più elevati, suggerendo che parte dei ritorni dell'investimento educativo rifletta meccanismi di *firm upgrading* piuttosto che esclusivamente differenze di produttività individuale. Nonostante questi progressi, rimane ancora limitata l'evidenza empirica sui meccanismi attraverso cui specifiche caratteristiche delle imprese influenzano i rendimenti salariali dell'istruzione. La letteratura ha infatti documentato il ruolo dei premi salariali d'impresa, ma ha dedicato minore attenzione all'identificazione delle caratteristiche organizzative, istituzionali e produttive che ne sono alla base. Comprendere tali meccanismi appare particolarmente rilevante nel caso italiano, caratterizzato da un'elevata eterogeneità del tessuto produttivo e da persistenti differenze salariali tra imprese.

L'analisi di questi aspetti richiede basi dati in grado di integrare informazioni sui lavoratori, sulle imprese e sulle loro interazioni nel tempo. La crescente disponibilità di dati amministrativi ha ampliato significativamente le possibilità di analisi dei fenomeni economici e sociali. Ciò non implica, tuttavia, che tali archivi siano di per sé sufficienti a rispondere a quesiti di ricerca complessi. Le fonti amministrative sono infatti prodotte prevalentemente per finalità gestionali e regolative e, se considerate isolatamente, presentano limiti informativi che ne restringono il potenziale analitico. Analogamente, le indagini statistiche offrono informazioni dettagliate sulle strategie, sull'organizzazione e sulle relazioni industriali delle imprese, ma non consentono generalmente di osservare con la medesima accuratezza le retribuzioni individuali e le traiettorie lavorative. L'integrazione micro tra dati amministrativi e dati di indagine rappresenta pertanto una strategia fondamentale per superare i limiti delle singole fonti informative e costruire basi dati longitudinali idonee allo studio congiunto delle caratteristiche dei lavoratori, delle imprese e delle loro relazioni.

Muovendosi in questa prospettiva, il presente lavoro analizza in quale misura le caratteristiche delle imprese contribuiscano all'eterogeneità dei rendimenti salariali dell'istruzione universitaria. In particolare, ci si domanda se la diversa capacità delle imprese di riconoscere premi salariali sia influenzata da caratteristiche quali la governance societaria, le relazioni industriali e la redditività aziendale e se tali fattori incidano sui rendimenti salariali dei lavoratori laureati attraverso meccanismi di *rent sharing* o politiche retributive maggiormente favorevoli ai lavoratori ad alta qualificazione.

A tal fine si utilizza una banca dati *employer-*

employee che integra le informazioni amministrative sulle storie lavorative e retributive contenute negli archivi Uniemens (Inps) e nel Sistema delle Comunicazioni Obbligatorie (COB-Mlps) con i dati campionari sulle caratteristiche delle imprese provenienti dalle diverse edizioni dell'Indagine RIL (Inapp). L'applicazione di una strategia econometrica che combina modelli alla Abowd-Kramarz-Margolis (AKM) e regressioni lineari con effetti fissi sui dati Inps-COB-RIL per il periodo 2013-2022 consente di ottenere tre risultati principali. In primo luogo, le politiche retributive delle imprese influenzano in modo positivo seppur debole l'evoluzione dei premi salariali dei laureati essendo questi ultimi riconducibili prevalentemente alla componente legata alle eterogeneità e abilità degli occupati. In secondo luogo, i differenziali retributivi legati al titolo di studio riflettono un processo di selezione (*sorting*) che non è uniforme lungo la distribuzione del capitale umano: l'*assortative matching* tra qualità delle imprese e abilità dei lavoratori è infatti debolmente positivo per i laureati e negativo per i meno qualificati. Infine, le regressioni mostrano come le aziende non trasferiscono automaticamente la performance competitiva in maggiori redditi per i lavoratori più qualificati e quando ciò accade la redistribuzione dei guadagni di produttività avviene solo parzialmente e nel medio periodo.

L'integrazione di tali fonti consente di associare alle caratteristiche organizzative e produttive delle imprese informazioni dettagliate sulle retribuzioni, sulle carriere lavorative e sulla mobilità dei lavoratori, rendendo possibile un'analisi che nessuna delle singole basi informative consentirebbe di realizzare autonomamente.

Il contributo del lavoro è triplice. In primo luogo, fornisce nuova evidenza sul ruolo delle caratteristiche d'impresa nella determinazione dei rendimenti salariali dell'istruzione universitaria, andando oltre le spiegazioni fondate esclusivamente sul capitale umano e sui processi di *sorting*. In secondo luogo, valorizza il potenziale conoscitivo derivante dall'integrazione tra dati amministrativi e dati di indagine, mostrando come tale approccio permetta di collegare gli effetti fissi d'impresa stimati mediante il modello AKM a caratteristiche osservabili delle imprese. Infine, contribuisce al dibattito sulle determinanti microeconomiche della disuguaglianza salariale, evidenziando come le politiche retributive e le caratteristiche organizzative delle imprese influenzino l'efficacia economica dell'investimento in istruzione.

L'articolo è organizzato come segue. La sezione

1 presenta il quadro teorico e la letteratura di riferimento. La sezione 2 descrive il quadro istituzionale che disciplina l'accesso ai dati, le fonti informative utilizzate e il processo di integrazione della banca dati. La sezione 3 illustra la metodologia empirica e discute i principali risultati. La sezione finale conclude.

1. Il quadro teorico e la letteratura di riferimento

Come discusso nell'introduzione, la crescente attenzione al ruolo delle imprese nella determinazione dei salari ha favorito lo sviluppo di due filoni di ricerca strettamente connessi. Il primo analizza il contributo dell'eterogeneità tra imprese alla formazione dei differenziali salariali e della disuguaglianza retributiva; il secondo esamina i meccanismi attraverso cui le caratteristiche economiche e organizzative delle imprese, in particolare la loro capacità di generare rendite, si riflettono sulle retribuzioni attraverso fenomeni di *rent sharing*. È all'intersezione di questi due filoni che si colloca il presente contributo.

Una svolta fondamentale nello studio del ruolo delle imprese nella determinazione dei salari è rappresentata dal modello con effetti fissi additivi di lavoratore e impresa proposto da Abowd *et al.* (1999), che consente di separare empiricamente l'eterogeneità permanente dei lavoratori da quella delle imprese utilizzando dati longitudinali lavoratore-impresa. In questa tradizione, Card *et al.* (2013) mostrano, per la Germania, che l'eterogeneità tra stabilimenti e la sua evoluzione nel tempo spiegano una parte rilevante dell'aumento della disuguaglianza salariale. Grazie all'utilizzo dei dati amministrativi, Song *et al.* (2019) documentano come negli Stati Uniti la crescita della varianza dei guadagni nel lungo periodo sia trainata in larga misura dall'aumento della dispersione tra imprese (oltre che da dinamiche interne alle imprese), consolidando l'idea che la struttura 'firm-level' del mercato del lavoro sia centrale per comprendere la disuguaglianza. Un risultato complementare, nello stesso spirito di decomposizione tra-entro impresa¹ emerge anche da Barth *et al.* (2016), che attribuiscono una parte significativa della crescita della dispersione dei guadagni all'aumento della dispersione tra stabilimenti. Accanto a questi contributi, Bonhomme *et al.* (2019) propongono un quadro *distributional* per dati *matched* che consente una rappresentazione più flessibile dell'eterogeneità tra imprese e del *sorting*.

Rispetto al legame tra eterogeneità tra imprese e rendimenti dell'istruzione, l'idea di fondo è che una

1 Per decomposizione tra-entro le imprese si intende la distinzione tra la componente della dispersione attribuibile alle differenze tra imprese e quella riconducibile alla variabilità entro ciascuna impresa.

parte del premio salariale associato all'istruzione possa riflettere non solo maggiore produttività individuale, ma anche un migliore accesso a imprese che offrono premi salariali più elevati. In questo senso, Pan (2025) fornisce evidenza diretta su questo meccanismo mostrando che una quota dei rendimenti dell'istruzione terziaria può essere attribuita all'accesso a imprese con premi salariali più alti. Questa prospettiva è coerente con la lettura più ampia del ruolo delle imprese nella disuguaglianza e nei salari, sintetizzata da Card *et al.* (2018), che discutono i due canali principali attraverso cui le imprese incidono sui salari tra cui il *sorting* dei lavoratori. Nel nostro contesto, ciò suggerisce che l'eterogeneità dei rendimenti della laurea può emergere sia perché i laureati accedono più spesso a imprese *'better paying'*, sia perché a parità di laurea e di lavoratore le imprese differiscono nel modo in cui remunerano le competenze.

Il secondo filone, molto rilevante per la nostra ipotesi di ricerca, studia se e quanto i salari incorporino una componente di condivisione delle rendite generate dall'impresa. Un contributo classico è Van Reenen (1996), che trova evidenza coerente con meccanismi di *rent sharing*, collegando innovazione (e rendite associate) a differenze salariali tra imprese. Kline *et al.* (2019) stimano come shock indotti da brevetti (e quindi da produttività/rendite) si trasmettano ai salari, offrendo evidenza su chi cattura i surplus generati dall'innovazione. In modo complementare, Lamadon *et al.* (2022) sviluppano un modello strutturale con dati *matched* per quantificare l'importanza di rendite e imperfezioni concorrenziali nel mercato del lavoro e discutere le implicazioni per la distribuzione dei salari. Nel loro insieme, questi lavori supportano l'idea che caratteristiche economiche e performance dell'impresa (tra cui la redditività) possano riflettersi nei salari attraverso canali di wage-setting e condivisione delle rendite, e che tali canali possano interagire con la composizione della forza lavoro e con il modo in cui le imprese valorizzano le competenze.

Il presente lavoro si inserisce in questa letteratura combinando due elementi. Da un lato, adottiamo l'impostazione *employer-employee* per stimare e scomporre i premi salariali d'impresa nel solco di Abowd *et al.* (1999) e della successiva letteratura sulla disuguaglianza tra imprese. Dall'altro, riprendiamo l'intuizione di Pan (2025) sul ruolo dei premi delle imprese nei rendimenti di istruzione, concentrando su alcune caratteristiche specifiche

dell'impresa, fra cui la redditività, come potenziali determinanti osservabili dei premi d'impresa e, quindi, dell'eterogeneità dei rendimenti della laurea.

2. Il quadro istituzionale e le fonti per l'integrazione micro

L'accesso ai dati elementari e il ruolo dell'Inapp nel Sistema statistico nazionale

La possibilità di operare su dati elementari costituisce una condizione imprescindibile per la produzione di analisi quantitative avanzate e per la costruzione di basi informative integrate. Nel caso dell'Inapp, tale possibilità è resa operativa dall'appartenenza dell'Istituto al Sistema statistico nazionale (SISTAN), istituito con il decreto legislativo n. 322 del 6 settembre 1989². Il SISTAN rappresenta la rete dei soggetti pubblici e privati che concorrono alla produzione e alla diffusione dell'informazione statistica ufficiale destinata al Paese e agli organismi internazionali. Ne fanno parte, tra gli altri, l'Istituto nazionale di statistica (Istat), gli uffici di statistica delle amministrazioni centrali e territoriali, le Regioni e Province autonome, le Camere di commercio e altri enti che svolgono funzioni di interesse pubblico in ambito statistico. All'interno del Sistema, l'Istat esercita funzioni di indirizzo, coordinamento e assistenza tecnica, avvalendosi del Comitato di indirizzo e coordinamento dell'informazione statistica (COMSTAT). L'appartenenza al SISTAN non assume una valenza meramente formale, ma comporta prerogative operative rilevanti, tra cui la possibilità di accedere – mediante procedure formalizzate – ad archivi contenenti dati elementari prodotti da altri soggetti del Sistema. Le modalità di scambio dei microdati sono disciplinate dal D.Lgs. n. 322/1989 e da successivi atti di indirizzo, tra cui la Direttiva COMSTAT n. 9 del 2004 e gli aggiornamenti intervenuti nel 2019³, che hanno precisato condizioni di accesso, finalità ammissibili e garanzie a tutela della riservatezza. Questo assetto normativo e istituzionale consente all'Inapp di operare su microdati, anziché esclusivamente su dati aggregati, nel quadro delle disposizioni del Programma statistico nazionale e nel rispetto della normativa in materia di protezione dei dati personali. La disponibilità di dati elementari rappresenta la condizione abilitante per la costruzione di infrastrutture informative integrate e per la realizzazione di analisi multilivello e longitudinali. Nel caso oggetto del presente contributo, tale prerogativa ha reso possibile

2 Decreto legislativo n. 322 del 6 settembre 1989, *Norme sul Sistema statistico nazionale e sulla riorganizzazione dell'Istituto nazionale di statistica*.

3 Direttiva COMSTAT n. 9/2004; decreto-legge n. 4 del 28 gennaio 2019; decisioni COMSTAT del 2 aprile 2019.

l'integrazione di informazioni relative alle imprese con dati amministrativi riferiti ai lavoratori e ai rapporti di lavoro, permettendo la costruzione di una struttura informativa di tipo *employer-employee*, nella quale le politiche retributive adottate dalle imprese possono essere collegate ai rendimenti dell'istruzione universitaria dei lavoratori.

Le fonti statistiche e amministrative utilizzate

La base dati integrata impiegata nell'analisi empirica deriva dalla combinazione di una rilevazione campionaria condotta dall'Inapp con archivi amministrativi, convertiti a uso statistico, acquisiti tramite il canale SISTAN.

Un ruolo centrale nella costruzione della base dati integrata è svolto dagli archivi amministrativi forniti dall'Inps attraverso il Coordinamento generale statistico attuariale, anch'essi acquisiti mediante canale SISTAN. La fornitura avviene con cadenza annuale e riguarda un campione di individui selezionato secondo il criterio della data di nascita (48 date di nascita per anno), modalità che garantisce la rappresentatività dell'occupazione dipendente nel settore privato non agricolo. La base informativa messa a disposizione comprende differenti archivi amministrativi, riferiti sia a persone sia a imprese: l'anagrafica individuale; le informazioni relative al lavoro dipendente, autonomo e parasubordinato; le misure di sostegno al reddito (indennità di mobilità, disoccupazione, CIG); il casellario dei pensionati; gli estratti conto contributivi; nonché dati riepilogativi relativi alle imprese. Nel contesto del presente contributo, la componente di maggiore rilievo è costituita dai dati relativi ai rapporti di lavoro dipendente afferenti al Fondo lavoratori dipendenti (FLD) nel settore privato extra-agricolo, con esclusione del lavoro domestico. Tali informazioni derivano dalle denunce mensili UNIEmens trasmesse dai datori di lavoro e successivamente rielaborate dall'Inps in modo da produrre record annuali distinti per impresa e caratteristiche contributive del lavoratore.

L'archivio consente di osservare lo stock di occupati con un elevato livello di dettaglio sulle retribuzioni effettivamente corrisposte, sulle settimane lavorate, sulla contribuzione versata e sulle caratteristiche contrattuali. I singoli dati annuali possono essere aggregati in modo da ricostruire un 'rapporto di lavoro' comprensivo dell'intero

periodo intercorrente tra data di assunzione e data di cessazione⁴, rendendo possibile un'analisi longitudinale coerente con la struttura dei dati SISCO-COB. La rilevanza degli archivi Inps nel presente lavoro è duplice. In primo luogo, essi consentono di disporre di una misura amministrativa della retribuzione, osservata a livello individuale e non soggetta a errori di autovalutazione o dichiarazione. La variabile salariale utilizzata nell'analisi empirica della seconda parte del contributo deriva infatti da queste informazioni contributive. In secondo luogo, la struttura longitudinale degli archivi permette di ricostruire le traiettorie occupazionali dei lavoratori nel tempo, includendo eventuali transizioni tra imprese. Ciò consente di controllare per dinamiche individuali pregresse e per effetti di mobilità, elementi fondamentali per stimare correttamente la dinamica salariale individuale e per isolare una componente residua della retribuzione, successivamente utilizzabile in analisi a livello d'impresa.

La seconda banca dati amministrativa è rappresentata dal Sistema delle Comunicazioni Obbligatorie (COB) online, un archivio gestito dal Ministero del Lavoro e delle politiche sociali, che raccoglie informazioni sugli eventi che caratterizzano i rapporti di lavoro: instaurazioni, trasformazioni, proroghe e cessazioni. L'obbligo di comunicazione telematica è stato introdotto dalla legge n. 296/2006⁵. L'archivio ha natura di flusso: l'unità elementare è l'evento amministrativo, che può essere aggregato per ricostruire il 'rapporto di lavoro' intercorrente tra la comunicazione di avviamento e quella di cessazione. L'archivio copre la quasi totalità del lavoro dipendente e alcune tipologie di lavoro autonomo e parasubordinato. Occorre, tuttavia, considerare alcuni limiti strutturali: i lavoratori assunti prima dell'introduzione dell'obbligo di comunicazione e che non hanno subito variazioni nel rapporto di lavoro possono non comparire nell'archivio⁶; inoltre, i dati sono disponibili con un certo ritardo temporale rispetto al periodo di riferimento⁷. Ai fini delle analisi presentate nel presente articolo, SISCO svolge una funzione complementare rispetto agli archivi previdenziali, fornendo informazioni sulle qualifiche professionali e sui titoli di studio dei lavoratori, non presenti negli archivi Inps.

Con riferimento alla banca dati di natura campionaria, prima di descrivere nel dettaglio RIL è necessario introdurre il Registro statistico

4 I singoli record annuali possono essere aggregati in modo da ricostruire un 'rapporto di lavoro' che, analogamente a quanto avviene per SISCO, comprende tutte le informazioni intercorrenti tra la data di assunzione e quella dell'eventuale cessazione.

5 Legge n. 296/2006 (Legge Finanziaria 2007).

6 Lavoratori assunti prima del 2008 e senza successive variazioni contrattuali possono non risultare nell'archivio.

7 I dati sono disponibili con uno scarto temporale di circa due trimestri rispetto al periodo di riferimento.

delle imprese (ASIA) dell'Istat che ne costituisce la base informativa per la definizione dell'universo di riferimento e per il disegno campionario, probabilistico e complesso garantendo la coerenza tra la popolazione osservata nella survey e quella rilevata nei registri statistici ufficiali. ASIA include le unità economiche che esercitano arti e professioni nelle attività industriali, commerciali e dei servizi, fornendo informazioni anagrafiche e strutturali quali attività economica, numero di addetti, forma giuridica, data di inizio e fine attività e fatturato. Dal campo di osservazione sono escluse le attività relative all'agricoltura, alla pubblica amministrazione, alle organizzazioni associative e ad altri settori specifici⁸. Tale perimetro risulta coerente con quello adottato nella RIL e con la copertura dei principali archivi amministrativi sul lavoro dipendente nel settore privato. La funzione di ASIA nel presente contesto non è soltanto quella di frame campionario, ma anche quella di riferimento strutturale per garantire coerenza tra l'universo delle imprese osservate nella survey e quelle presenti negli archivi amministrativi.

La Rilevazione longitudinale su Imprese e Lavoro-RIL costituisce la fonte campionaria utilizzata nell'analisi. Si tratta di un'indagine statistica inserita nel Programma statistico nazionale (PSN) e caratterizzata da obbligo di risposta per i soggetti privati coinvolti, in quanto rilevazione di interesse pubblico. La prima edizione della RIL risale al 2005; l'indagine è stata successivamente riproposta con cadenza periodica, fino all'edizione più recente (la VII) attualmente in fase di realizzazione. La continuità nel tempo della rilevazione consente di disporre di una base informativa strutturata e coerente, idonea ad analisi diacroniche sui comportamenti delle imprese. L'indagine è finalizzata allo studio del comportamento delle imprese nei confronti del lavoro, con particolare riferimento alle strategie organizzative, alle politiche di gestione delle risorse umane, alle relazioni industriali e all'introduzione di innovazioni tecnologiche e organizzative. Le informazioni raccolte consentono di analizzare le relazioni tra le scelte aziendali in materia di utilizzo del lavoro e i risultati economici dell'impresa. L'unità di analisi è l'impresa. Il campione, composto da circa 30.000 unità, è rappresentativo delle società di capitali e delle società di persone attive sul territorio nazionale nei settori

privati non agricoli. Sono pertanto escluse dal campo di osservazione le imprese individuali e le altre forme giuridiche non rientranti in tali categorie, oltre ai settori economici non compresi nel perimetro definito dal Registro ASIA⁹. Un elemento metodologicamente rilevante della RIL è la presenza di una componente longitudinale: una quota, circa il 30%, del campione è costituita da imprese già intervistate nelle edizioni precedenti, consentendo di seguire nel tempo l'evoluzione delle strategie aziendali. Nel contesto dell'analisi applicativa sviluppata nella seconda parte del contributo, la RIL fornisce l'insieme di variabili relative a redditività, relazioni industriali, investimenti, innovazione e competitività. L'integrazione con gli archivi Inps e SISCO consente di collegare tali informazioni alle retribuzioni effettivamente percepite dai lavoratori impiegati nelle imprese osservate, costruendo una base dati multilivello nella quale il livello decisionale dell'impresa può essere messo in relazione con gli esiti salariali individuali nel tempo costruendo una struttura *employer-employee*. Le edizioni utilizzate nell'articolo fanno riferimento alle rilevazioni condotte nel 2015, 2018 e 2022.

Costruzione della struttura employer-employee panel

L'obiettivo dell'integrazione tra RIL, archivi Inps e SISCO è la costruzione di una struttura informativa di tipo *employer-employee*, nella quale le caratteristiche e le scelte dell'impresa possano essere collegate agli esiti individuali dei lavoratori nel tempo. A tal fine, il primo passaggio consiste nell'identificazione delle imprese incluse nella RIL e nel collegamento di tali unità con gli archivi amministrativi mediante il codice fiscale. Una volta individuate le imprese della survey negli archivi Inps e SISCO, vengono selezionati tutti i lavoratori che risultano aver intrattenuto un rapporto di lavoro con tali imprese nel periodo di osservazione. In un secondo momento, per ciascun lavoratore così identificato, viene ricostruita l'intera storia lavorativa disponibile negli archivi amministrativi, indipendentemente dal datore di lavoro. La struttura risultante è dunque un panel multilivello nel quale l'impresa costituisce il livello decisionale e il lavoratore rappresenta l'unità di outcome.

La costruzione della base dati integrata richiede un'articolata attività di verifica e validazione, finalizzata a garantire la coerenza interna degli

8 Sono escluse, tra le altre, le attività relative ad agricoltura (sezione A), amministrazione pubblica (sezione O), organizzazioni associative (divisione 94), famiglie come datori di lavoro (sezione T), organismi extraterritoriali (sezione U).

9 Sono escluse dal campo di osservazione le attività economiche relative a: agricoltura, silvicoltura e pesca (sezione A); amministrazione pubblica e difesa; assicurazione sociale obbligatoria (sezione O); attività di organizzazioni associative (divisione 94); attività di famiglie e convivenze come datori di lavoro per personale domestico; produzione di beni e servizi indifferenziati per uso proprio da parte di famiglie e convivenze (sezione T); organizzazioni ed organismi extraterritoriali (sezione U); unità classificate come istituzioni pubbliche e istituzioni private non profit.

archivi e l'affidabilità delle informazioni risultanti dall'integrazione. Le verifiche riguardano, in primo luogo, la coerenza temporale dei rapporti di lavoro e delle retribuzioni osservate, con particolare attenzione alla continuità delle carriere individuali e alla corretta attribuzione dei lavoratori alle imprese nel tempo. In secondo luogo, vengono effettuati controlli incrociati sulle variabili comuni tra le diverse fonti, al fine di identificare eventuali discrepanze e valutarne l'impatto sull'analisi. È inoltre necessario tenere conto dei limiti derivanti dalla parziale sovrapposizione degli universi osservati, dalla natura campionaria della fornitura Inps e dalle possibili lacune informative presenti negli archivi amministrativi. Tali elementi vengono considerati nella definizione dell'universo analitico e nell'interpretazione dei risultati. Nonostante tali limiti, l'integrazione tra le fonti consente di ampliare significativamente il contenuto informativo disponibile rispetto all'utilizzo isolato delle singole basi dati, permettendo di costruire una struttura *employer-employee* panel coerente e metodologicamente solida¹⁰.

3. Analisi empirica

La strategia econometrica

In questo paragrafo illustriamo la strategia econometrica adottata per analizzare in quale misura le caratteristiche delle imprese contribuiscano all'eterogeneità dei rendimenti salariali dell'istruzione universitaria. L'analisi si articola in due fasi sequenziali e complementari. Nella prima vengono stimati i premi salariali d'impresa attraverso un modello a effetti fissi lavoratore-impresa (AKM), distinguendo tra lavoratori laureati e non laureati; nella seconda tali premi salariali sono messi in relazione alle caratteristiche economiche, organizzative e istituzionali delle imprese rilevate dall'indagine RIL. La prima fase dell'analisi si colloca nel solco della letteratura inaugurata da Abowd *et al.* (1999), secondo cui la determinazione dei salari riflette congiuntamente l'eterogeneità osservata e non osservata dei lavoratori e quella delle imprese presso cui essi sono occupati (Card *et al.* 2013; Bonhomme *et al.* 2019; Casarico e Lattanzio 2024). Utilizziamo i dati amministrativi integrati Inps-COB per stimare una versione estesa del modello AKM nella quale gli effetti fissi individuali sono invariati nel tempo, mentre la componente di impresa è consentita variare tra differenti intervalli temporali, così da ricostruire l'evoluzione delle politiche

retributive aziendali (Babet *et al.* 2015; Lachowska *et al.* 2023).

Coerentemente con le indicazioni della letteratura più recente, il modello viene stimato separatamente per due sottocampioni di lavoratori: quelli in possesso di un titolo di istruzione terziaria e quelli con un titolo di studio non universitario. La classificazione è definita in modo costante nel tempo. In particolare, il livello di istruzione è identificato mediante una variabile dummy che assume valore pari a uno per i lavoratori laureati e zero altrimenti, costruita sulla base del massimo livello di istruzione osservato per ciascun individuo nell'intero periodo disponibile. Tale scelta evita che eventuali variazioni nella registrazione del titolo di studio interferiscano con l'identificazione degli effetti fissi individuali e consente un confronto omogeneo tra lavoratori laureati e non laureati lungo tutto l'orizzonte temporale di osservazione. Formalmente, per ciascun gruppo di istruzione (*e*) (laureati e non laureati) viene stimata la seguente equazione:

$$[1] \quad \ln w_{ijt,e} = \alpha_{i,e} + \psi_{jt,e} + X_{it,e}\beta + \varepsilon_{it,e}$$

dove $\ln w_{ijt,e}$ rappresenta il logaritmo del salario settimanale full-time equivalent del lavoratore (*i*), occupato nell'impresa (*j*) nel periodo $t = [2013-16, 2017-19, 2020-22]$. Il vettore $X_{it,e}$ include le principali caratteristiche osservabili del lavoratore, mentre rappresenta l'effetto fisso individuale e cattura l'eterogeneità non osservata e invariante nel tempo. Il termine $\psi_{jt,e}$ identifica l'effetto fisso impresa-periodo e misura il premio salariale riconosciuto dall'impresa (*j*) ai lavoratori appartenenti al gruppo di istruzione (*e*) nel periodo (*t*). Il termine di errore $\varepsilon_{it,e}$ è un disturbo idiosincratico a media nulla e varianza finita. La differenza tra i premi salariali stimati per laureati e non laureati, $(\psi_{jt}^L - \psi_{jt}^{NL})$, rappresenta la componente *firm-specific* dei rendimenti salariali dell'istruzione universitaria e costituisce la variabile dipendente dell'analisi svolta nel secondo stadio.

Nella seconda fase le stime ottenute dal modello AKM sono integrate con le informazioni raccolte da RIL, al fine di verificare se e in quale misura la dinamica dei premi salariali riconosciuti ai lavoratori laureati sia associata alle caratteristiche economiche, organizzative e istituzionali delle imprese. A tale scopo viene stimato il seguente modello lineare:

$$[2] \quad \psi_{jt}^{L-NL} = \beta_0 + \beta_1 \ln(\Pi_{jt}) + \beta_2 \ln(\Pi_{jt}) \times 1(2017-19) + \beta_3 \ln(\Pi_{jt}) \times 1(2020-22) + \delta' F_{jt} + \mu_j + \varepsilon_{jt}$$

10 Si rimanda all'Appendice metodologica per una descrizione approfondita delle procedure di pretrattamento e standardizzazione delle forniture, *record linkage* e integrazione micro tra archivi.

dove ψ_{jt}^{L-NL} rappresenta il differenziale tra i premi salariali stimati per laureati e non laureati nell'impresa (j) e nel periodo (t). Tra le variabili di controllo di fonte RIL focalizziamo l'attenzione sul (log del) fatturato annuo per dipendente Π_{jt} e sui termini di interazione della redditività con le variabili indicatrici degli intervalli di tempo che – a loro volta – possono essere interpretati come una sorta di elasticità 'dinamica tra performance e politiche salariali per l'istruzione'¹¹. Il vettore F_{jt} include le caratteristiche del management, della specializzazione produttiva, l'organizzazione del personale, relazioni industriali, altro. Il parametro μ_j identifica l'eterogeneità non osservata e invariante nel tempo, mentre ε_{jt} è un disturbo idiosincratico. È opportuno sottolineare come la strategia econometrica viene declinata nella struttura statistico informativa del nostro dataset – anche per offrire un ordito interpretativo dei risultati delle regressioni. I premi salariali d'impresa vengono infatti stimati separatamente nei tre sottoperiodi 2013-2016, 2017-2019 e 2020-2022, costruiti in corrispondenza delle tre edizioni della RIL (2015, 2018 e 2021). Tale scelta consente di associare a ciascun premio salariale le caratteristiche dell'impresa osservate nella corrispondente rilevazione RIL, garantendo la coerenza temporale tra le variabili dipendenti stimate nel primo stadio e le covariate utilizzate nel secondo.

Le stime del modello AKM sono effettuate sulla massima componente connessa lavoratore-impresa (*largest connected set*¹²), così da garantire l'identificazione degli effetti fissi d'impresa e la confrontabilità dei premi salariali stimati all'interno di ciascun sottocampione di istruzione. L'utilizzo della massima componente connessa contribuisce inoltre a ridurre i problemi di identificazione derivanti dalla limitata mobilità dei lavoratori tra imprese (*limited mobility bias*), pur senza eliminarli completamente; di conseguenza, i risultati della decomposizione della varianza devono essere interpretati soprattutto in termini comparativi tra lavoratori laureati e non laureati.

Le elaborazioni econometriche fanno, infine, riferimento a un campione finale di lavoratori dipendenti di età compresa tra 20 e 60 anni occupati in imprese private extra-agricole con almeno dieci dipendenti e per i quali è disponibile almeno un rapporto di lavoro nelle COB, fonte dalla quale viene ricavata l'informazione sul titolo di studio. Tale selezione consente di concentrare l'analisi su imprese caratterizzate da una minima organizzazione dei

mercati interni del lavoro e, quindi, potenzialmente in grado di adottare politiche retributive differenziate tra gruppi di lavoratori. L'analisi empirica è condotta su un campione di oltre 10 milioni di osservazioni, corrispondenti a circa 2 milioni di lavoratori, dei quali circa 400 mila in possesso di un titolo universitario.

I risultati principali

La tabella 1 riporta le principali statistiche descrittive del campione, distinguendo tra lavoratori laureati e non laureati, sia con riferimento all'intero campione sia alla restrizione campionaria corrispondente al *largest connected set* lavoratore-impresa (utilizzata per la stima del modello AKM).

Le statistiche descrittive evidenziano differenze sistematiche tra i due gruppi di lavoratori. Nel campione completo, i laureati percepiscono una retribuzione settimanale mediamente più elevata rispetto ai non laureati (pari rispettivamente a 6,4 e 6,2), risultano mediamente più giovani (37 contro 41 anni circa) e presentano una maggiore incidenza della componente femminile (56% contro 39%). Differenze pressoché identiche si osservano nel *largest connected set*, indicando che la restrizione del campione necessaria all'identificazione degli effetti fissi lavoratore-impresa non altera in misura significativa le caratteristiche osservabili dei lavoratori né il differenziale salariale tra i due gruppi. Nel complesso, tali evidenze risultano coerenti con la letteratura sul mercato del lavoro italiano, che documenta l'esistenza di un premio salariale associato al conseguimento di un titolo universitario, sebbene di entità relativamente contenuta e in progressiva riduzione rispetto a quanto osservato nella maggior parte dei Paesi OCSE (Naticchioni et al. 2010).

Il semplice confronto delle retribuzioni medie, tuttavia, non consente di stabilire in quale misura tale differenziale rifletta la maggior dotazione di capitale umano di chi è in possesso di un titolo universitario oppure sia riconducibile alle specifiche caratteristiche e comportamenti delle aziende presso cui sono occupati. In particolare, una quota del differenziale retributivo associato alla laurea potrebbe essere correlata a meccanismi di *sorting* e processo di *assortative matching*: gli individui con maggiori abilità non osservate tendono ad essere concentrati in tipologie di imprese che adottano politiche più favorevoli alla forza lavoro qualificata. E il processo di *assortative matching* potrebbe essere intrinsecamente differente, nel segno e nel valore

11 Nel dettaglio, il coefficiente β_1 misura il legame tra redditività e premio per laureati nel periodo iniziale (2013-2016), mentre i coefficienti β_2 e β_3 indicano come evolve tale nesso nel 2017-2019 e nel 2020-2022, rispettivamente.

12 La massima componente connessa è identificata come il più ampio sotto grafo del network lavoratore-impresa nel quale ciascuna impresa risulta collegata alle altre attraverso almeno una transizione lavorativa osservata.

Tabella 1. Statistiche descrittive

	Campione totale		Largest connected set	
	Mean	Std. dev.	Mean	Std. dev.
Laureati				
Log (salario)	6,415	0,541	6,421	0,542
Età (in anni)	37,18	9,148	37,13	9,111
Donne	0,563	0,496	0,567	0,496
N. osservazioni	2.102.115		1.967.173	
Non laureati				
Log (salario)	6,195	0,426	6,194	0,426
Età (in anni)	40,63	11,22	40,54	11,21
Donne	0,387	0,487	0,386	0,486
N. osservazioni	8.442.235		8.039.660	

Fonte: elaborazioni degli Autori su dati Inps-COB 2013-2022

assoluto, lungo la distribuzione del capitale umano.

L'analisi econometrica sviluppata nei paragrafi successivi è finalizzata proprio a distinguere queste due componenti. Attraverso la stima del modello AKM separatamente per lavoratori laureati e non laureati è infatti possibile isolare il contributo della componente *firm-specific* alla determinazione dei salari e verificare successivamente se le differenze nei premi salariali riconosciuti ai lavoratori con istruzione universitaria risultino sistematicamente associate alle caratteristiche economiche, organizzative e istituzionali delle imprese. In questa prospettiva, la tabella 2 riporta le stime del modello AKM ottenute separatamente per i due gruppi di lavoratori. La specificazione consente di isolare il contributo dell'eterogeneità individuale e di quella d'impresa alla determinazione dei salari all'interno dei due gruppi di istruzione, fornendo la base per analizzare il ruolo delle caratteristiche aziendali nella formazione dei rendimenti salariali dell'istruzione. Per entrambi i sottocampioni, i coefficienti associati all'età e al suo quadrato presentano rispettivamente segno positivo e negativo, delineando un profilo salariale concavo lungo il ciclo di vita lavorativa, coerentemente con le previsioni della teoria del capitale umano. L'effetto risulta più pronunciato tra i laureati, suggerendo una maggiore crescita delle retribuzioni nelle fasi iniziali e centrali della carriera, in linea con percorsi professionali caratterizzati da una più elevata accumulazione di capitale umano specifico e da maggiori opportunità di progressione salariale.

L'inclusione degli effetti fissi individuali (α_i) e degli effetti fissi impresa-periodo (ψ_{jt}) consente di spiegare una quota molto elevata della variabilità salariale osservata (R^2 pari a 0,87 per i laureati e 0,83 per i non laureati), confermando l'elevata capacità esplicativa del modello AKM applicato a dati amministrativi lavoratore-impresa. Le stime degli effetti fissi ottenute in questa fase costituiscono la base per analizzare il contributo relativo delle componenti individuali e aziendali alla dispersione salariale.

A questo punto è possibile esaminare i fattori sottostanti la dispersione salariale e la natura del processo di *matching/sorting* tra individui e caratteristiche delle imprese. La tabella 3 riporta la decomposizione della varianza e la matrice varianza-covarianza delle componenti stimate dal modello AKM. In linea con la letteratura sui dati *matched employer-employee* (Card *et al.* 2013), i risultati evidenziano che la dispersione salariale è spiegata prevalentemente dall'eterogeneità non osservata dei lavoratori, mentre gli effetti fissi di impresa, pur significativi contribuiscono in misura più contenuta alla variabilità complessiva delle retribuzioni. Questo risultato emerge per entrambi i sottocampioni confermando che le caratteristiche permanenti dei lavoratori rappresentano il principale fattore di differenziazione delle retribuzioni, mentre le politiche salariali delle imprese contribuiscono a spiegare una quota più contenuta della dispersione osservata¹³.

13 Per i lavoratori laureati la varianza associata agli effetti fissi individuali ($\text{Var}(\alpha_i) = 0,146$) risulta sensibilmente superiore a quella attribuibile agli effetti fissi d'impresa ($\text{Var}(\psi_{jt}) = 0,078$); una configurazione analoga si osserva anche tra i non laureati, per i quali la componente individuale continua a rappresentare la principale fonte di eterogeneità salariale ($\text{Var}(\alpha_i) = 0,099$ contro $\text{Var}(\psi_{jt}) = 0,070$).

Tabella 2. Stime AKM. Variabile dipendente (log) salario settimanale

	Laureati	Non laureati
Età	0,083***	0,041***
	[0,003]	[0,001]
Età^2	[0,001***	-0,000***
	[0,000]	[0,000]
$\alpha_{i,e}$	Sì	Sì
$\psi_{jt,e}$	Sì	Sì
Costante	4.377***	5.108***
	[0,036]	[0,021]
N. lavoratori	373.726	1.575.682
N. imprese*periodo	160.933	567.912
N. osservazioni	1.967.173	8.039.660
R2	0,87	0,83

Note: le regressioni non includono variabili di controllo per le caratteristiche demografiche che risultano invarianti nel tempo. Errori standard robusti clusterizzati a livello di impresa sono riportati in parentesi. Significatività statistica: *** al 1%, ** al 5%, al 10%.

Fonte: elaborazione degli Autori su dati campione Inps-COB

Tabella 3. Scomposizione varianza-covarianza stime AKM

	α_i	ψ_{jt}	$X_{it} \beta$	Ln(salario)
<i>Laureati</i>				
α_i	0,146			
ψ_{jt}	0,006	0,078		
$X_{it} \beta$	-0,019	0,003	0,062	
Ln(salario)	0,133	0,087	0,047	0,294
<i>Non laureati</i>				
α_i	0,099			
ψ_{jt}	-0,007	0,070		
$X_{it} \beta$	-0,015	0,002	0,028	
Ln(salario)	0,077	0,065	0,016	0,181

Fonte: elaborazioni degli Autori su dati Inps-COB 2013-2022

La matrice varianza-covarianza consente, inoltre, di approfondire il processo di allocazione dei lavoratori tra imprese. Per i lavoratori con istruzione universitaria emerge una covarianza positiva, seppure di entità molto contenuta, tra effetto fisso individuale ed effetto fisso d'impresa ($\text{Cov}(\alpha_i, \psi_{jt}) = 0,006$). Tale evidenza suggerisce la presenza di *assortative matching* positivo, ossia una moderata tendenza dei lavoratori caratterizzati da una maggiore componente di abilità non osservata a essere occupati presso imprese che riconoscono premi salariali relativamente più elevati. Il risultato è coerente sia con la letteratura internazionale sui dati *matched*

employer-employee, che documenta l'esistenza di processi di allocazione tra lavoratori e imprese in grado di influenzare la dispersione salariale (Card *et al.* 2013), sia con la più recente evidenza riferita al mercato del lavoro italiano, secondo cui una parte dei differenziali retributivi dei laureati riflette il processo di *sorting* verso imprese caratterizzate da differenti politiche salariali (Bonacini *et al.* 2025). La ridotta entità della covarianza osservata suggerisce, tuttavia, che tale meccanismo, pur presente, non rappresenti l'unico fattore alla base dei rendimenti salariali dell'istruzione universitaria. Per i lavoratori non laureati si osserva invece una covarianza lievemente

negativa tra effetto fisso individuale ed effetto fisso d'impresa ($\text{Cov}(\alpha_i, \psi_{it}) = -0,007$), che segnala un processo di allocazione di segno opposto rispetto a quello osservato per i laureati. Sebbene anche in questo caso la magnitudine della covarianza sia estremamente contenuta, la diversa configurazione dei due sottocampioni suggerisce che i meccanismi di matching tra lavoratori e imprese operino in modo differenziato a seconda del livello di istruzione¹⁴. In sintesi, quanto emerge nella tabella 3 conferma l'idea che il ruolo delle politiche retributive a livello aziendale non costituiscono la principale fonte della dispersione salariale complessiva. Ciò non implica che esse siano economicamente irrilevanti. Al contrario, la circostanza che la componente *firm-specific* presenti un peso relativamente contenuto nella varianza aggregata rende particolarmente interessante comprendere quali caratteristiche delle imprese siano associate alla sua eterogeneità. In altre parole, se i differenziali salariali riconosciuti ai lavoratori laureati non dipendono esclusivamente dalla composizione della forza lavoro diventa naturale chiedersi se vi siano fattori microeconomici interni al ciclo di vita aziendale (le relazioni industriali, la redditività e la capacità di generare rendite) che contribuiscano a spiegare le differenze nei premi salariali riconosciuti ai lavoratori più qualificati.

È proprio questa l'ipotesi che viene verificata nel secondo stadio dell'analisi. La tabella 4 riporta i risultati delle stime dell'equazione [2], ottenute sulla componente longitudinale della banca dati integrata Inps-COB-RIL, con l'obiettivo di verificare se le caratteristiche economiche e organizzative delle imprese contribuiscano a spiegare l'eterogeneità dei premi salariali riconosciuti ai lavoratori laureati. Nello specifico, la colonna (1) presenta le stime ottenute mediante regressione OLS. I coefficienti associati alla performance economica dell'impresa, misurata attraverso il logaritmo del fatturato per dipendente, non risultano statisticamente significativi né nel periodo iniziale né nei successivi intervalli temporali. Considerata isolatamente, tale evidenza sembrerebbe suggerire l'assenza di una relazione sistematica tra performance aziendale e premi salariali riconosciuti ai lavoratori con istruzione universitaria. Tale interpretazione deve essere considerata con cautela, poiché le stime OLS non tengono conto dell'eterogeneità non osservata tra imprese, che può riflettere caratteristiche

permanenti quali modelli di governance, pratiche manageriali, assetti organizzativi e altre determinanti strutturali delle politiche retributive. La colonna (2), che prevede l'introduzione nel modello degli effetti fissi di impresa, restituisce un quadro sensibilmente diverso: una volta controllato per l'eterogeneità aziendale invariante nel tempo emerge infatti una relazione dinamica tra performance economica e premi salariali riconosciuti ai lavoratori laureati.

Nel periodo iniziale (2013-2016), un incremento del fatturato per dipendente risulta associato a una riduzione del differenziale salariale a favore dei laureati (-0,9%). Tale risultato non implica che i laureati percepiscano retribuzioni inferiori rispetto ai lavoratori non laureati; suggerisce piuttosto che, nelle imprese caratterizzate da una migliore performance economica, il premio salariale associato all'istruzione universitaria risulti relativamente più contenuto. Nei periodi successivi il quadro empirico evolve progressivamente. L'interazione tra performance economica e periodo 2017-2019 assume segno positivo, pur rimanendo statisticamente non significativa, mentre nel periodo 2020-2022 il coefficiente diviene positivo e statisticamente significativo. Nell'ultima fase del periodo analizzato, pertanto, una migliore performance economica dell'impresa risulta associata a un incremento dei premi salariali riconosciuti ai lavoratori laureati, suggerendo che una quota crescente dei benefici economici generati dall'impresa venga destinata alla componente più qualificata della forza lavoro. I risultati indicano, quindi, che la relazione tra performance aziendale e rendimenti salariali dell'istruzione non è stabile nel tempo, ma evolve nel corso del periodo osservato. Il progressivo rafforzamento del legame nell'ultima fase suggerisce che gli effetti della maggiore performance economica sulle politiche retributive possano manifestarsi gradualmente, nell'ambito dei meccanismi di determinazione dei salari propri del sistema italiano di relazioni industriali. Tale interpretazione appare coerente con la letteratura che evidenzia come le istituzioni della contrattazione collettiva contribuiscano a modellare la trasmissione dei differenziali di produttività ai salari e a contenere la dispersione delle politiche retributive tra imprese (Devicienti *et al.* 2019).

14 Nell'interpretare il segno e la magnitudine della covarianza tra effetti fissi individuali ed effetti fissi d'impresa è opportuno considerare le possibili implicazioni del cosiddetto *limited mobility bias*, ampiamente discusso nella letteratura sui modelli AKM (Saggio *et al.*, 2023). Tale fenomeno deriva dalla limitata mobilità dei lavoratori tra imprese, dalla quale dipende l'identificazione degli effetti fissi d'impresa, e può influenzare la stima della componente *firm-specific* e, conseguentemente, della covarianza tra lavoratori e imprese. Pertanto, i risultati relativi al processo di assortative matching devono essere interpretati principalmente in termini comparativi tra i due gruppi di lavoratori, piuttosto che come una misura puntuale dell'intensità del matching.

Tabella 4. Stime regressioni lineari. Variabile dipendente: premi salariali laureati (vs non laureati) per impresa e periodo

	[1]	[2]
$\ln(\Pi_{jt}) \times 1(2020-22)$	0,007	0,008*
	[0,006]	[0,005]
$\ln(\Pi_{jt}) \times 1(2017-19)$	0,002	0,003
	[0,006]	[0,004]
$\ln(\Pi_{jt})$	-0,008	-0,009**
	[0,005]	[0,004]
Vacancy ratio	0,029	0,056*
	[0,020]	[0,034]
Associazione datoriale	0,029*	0,012
	[0,015]	[0,010]
Altri controlli	Sì	Sì
Effetti fissi impresa	No	Sì
Costante	0,020	0,038
	[0,054]	[0,040]
N. imprese	1.247	1.247
N. osservazioni	3.502	3.502
R2	0,027	0,752

Note: altri controlli includono livello di istruzione e genere del management, governance familiare, numero di dipendenti (in classi dimensionali), età (in anni) dell'impresa. Tutte le regressioni includono effetti fissi per regione geografica e settore di attività (ATECO 2 digit). Errori standard clusterizzati a livello di impresa riportati in parentesi. Significatività statistica: *** al 1%, ** al 5%, al 10%.

Fonte: elaborazioni degli Autori su dati Inps-COB-RIL

Questi risultati sembrano essere coerenti con l'ipotesi che le caratteristiche delle imprese contribuiscano, almeno in parte, a spiegare l'eterogeneità dei rendimenti salariali dell'istruzione universitaria. Sebbene la componente *firm-specific* rappresenti una quota relativamente contenuta della dispersione salariale complessiva, i premi salariali riconosciuti ai lavoratori laureati risultano sistematicamente associati ad alcune caratteristiche economiche delle imprese nelle quali essi sono occupati quali ad esempio un diverso contesto organizzativo e retributivo. Tale evidenza è in linea con la recente letteratura che attribuisce alle imprese un ruolo non trascurabile nella determinazione della disuguaglianza salariale e dei rendimenti dell'istruzione, evidenziando come le politiche di *wage-setting* e l'eterogeneità tra imprese possano contribuire ai differenziali retributivi osservati (Card *et al.* 2018; Bonacini *et al.* 2025; Pan 2025).

Conclusioni

Le analisi presentate nelle pagine precedenti hanno esaminato la relazione tra politiche salariali delle imprese e rendimento dell'istruzione

universitaria, considerando congiuntamente l'eterogeneità degli individui, le caratteristiche delle aziende in cui sono occupati e il processo di matching tra domanda e offerta di lavoro qualificato.

A tal fine è stato sviluppato un sistema informativo integrato che combina le informazioni sulle storie retributive e lavorative contenute negli archivi Uniemens (Inps) e nel Sistema delle Comunicazioni Obbligatorie con i dati campionari sulle imprese raccolti dall'Indagine RIL (Inapp).

Le stime econometriche restituiscono evidenze relativamente nuove per il mercato del lavoro italiano. In primo luogo, emerge un ruolo positivo, seppur contenuto, delle politiche retributive aziendali nel condizionare l'andamento dei premi salariali dei laureati, che restano comunque influenzati soprattutto dalla componente legata alle abilità e all'eterogeneità individuale. Inoltre, i differenziali retributivi associati al titolo di studio riflettono un processo di selezione non uniforme lungo la distribuzione del capitale umano: l'*assortative matching* tra qualità delle imprese e abilità dei lavoratori risulta debolmente positivo per i laureati e negativo per gli occupati meno qualificati.

Le regressioni basate sui dati d'impresa mostrano poi che le aziende italiane non trasformano automaticamente la propria performance competitiva in salari più elevati per i lavoratori altamente qualificati. Quando la redistribuzione dei guadagni di produttività si verifica, essa è parziale e si manifesta solo nel medio periodo. Tale risultato suggerisce che il tessuto imprenditoriale contribuisce alla compressione dei rendimenti salariali dell'istruzione, indipendentemente dall'azione di eventuali meccanismi compensativi legati agli assetti delle relazioni industriali o alla regolamentazione dei contratti collettivi.

Nel complesso, i premi dell'istruzione non possono essere interpretati come un mero riflesso della produttività individuale: dipendono anche dalla collocazione dei lavoratori in imprese che risultano eterogenee per strategie, istituzioni e capacità di generare premi salariali. Ne deriva che le politiche educative e del lavoro dovrebbero considerare non solo l'accumulazione di capitale umano, ma anche le condizioni che favoriscono l'accesso dei laureati alle imprese di qualità e a contesti capaci di valorizzarne le competenze.

In conclusione, il nostro studio contribuisce alla letteratura su capitale umano, disuguaglianza salariale ed eterogeneità d'impresa, mostrando che una piena comprensione dei rendimenti dell'istruzione richiede un'analisi congiunta delle strategie aziendali e delle traiettorie lavorative individuali.

Appendice metodologica

Pretrattamento e standardizzazione delle forniture

L'attività di acquisizione delle banche dati contenenti informazioni elementari è svolta, per Inapp, dal Servizio statistico, che provvede, una volta ricevute le forniture, all'implementazione delle necessarie procedure di pretrattamento e trattamento dei dati. Tali attività sono finalizzate alla predisposizione di dataset controllati, coerenti e pseudonimizzati, idonei all'elaborazione statistica da parte del personale autorizzato. Le procedure di trattamento si differenziano in funzione della natura della fonte, delle modalità di raccolta originaria delle informazioni e del grado di strutturazione delle basi dati fornite.

Nel caso delle forniture derivanti dal Sistema delle Comunicazioni Obbligatorie (SISCO), i dati trasmessi dal Ministero del Lavoro sono già oggetto di una prima rielaborazione e vengono forniti in forma strutturata come 'rapporti di lavoro', ossia come aggregazioni degli eventi amministrativi intercorsi tra l'avvio e l'eventuale cessazione del

rapporto. Tuttavia, nonostante tale trattamento preliminare, la base dati mantiene un elevato grado di complessità. La natura amministrativa dell'archivio, la numerosità dei record e la possibile frammentazione o sovrapposizione dei rapporti nel tempo rendono necessario un ulteriore intervento di controllo e armonizzazione. In particolare, è indispensabile verificare la coerenza temporale dei rapporti ricostruiti, garantire la congruenza tra identificativi di impresa e lavoratore, armonizzare le classificazioni professionali e i titoli di studio e assicurare l'allineamento con le variabili presenti negli archivi Inps e nella RIL.

Per quanto riguarda le forniture Inps, il livello di strutturazione è più elevato dato un primo trattamento statistico realizzato a monte dal Coordinamento generale statistico attuariale dell'Istituto. Ciò non elimina la necessità di ulteriori verifiche. L'attività di Inapp si concentra sulla validazione della coerenza interna delle informazioni contributive e retributive, sull'armonizzazione delle variabili con le altre fonti e sulla ricostruzione coerente delle traiettorie lavorative individuali nel tempo. La presenza di record annuali distinti per impresa e caratteristiche contributive del lavoratore richiede infatti un attento lavoro di ricomposizione longitudinale, volto a garantire continuità e correttezza nell'osservazione dei percorsi occupazionali. Le basi dati Istat, caratterizzate da un elevato grado di standardizzazione, richiedono prevalentemente operazioni di allineamento classificatorio e di coerenza rispetto al perimetro di osservazione adottato nelle altre fonti. Al termine di tali operazioni vengono prodotti archivi standardizzati, controllati e pseudonimizzati, idonei all'integrazione micro tra le diverse fonti informative. Questa fase preliminare è condizione necessaria per la successiva attività di *record linkage*, dalla quale deriva la costruzione della base dati *employer-employee* utilizzata nell'analisi empirica di questo studio.

Record linkage e integrazione micro tra archivi

La fase di integrazione tra le diverse fonti informative non si esaurisce in un semplice processo tecnico di accostamento tra archivi, ma costituisce un passaggio metodologicamente complesso che richiede la definizione preliminare dell'universo di riferimento, delle chiavi di collegamento e dei criteri di validazione degli abbinamenti effettuati. Nel caso in esame, l'integrazione avviene prevalentemente mediante *linkage* deterministico basato sugli identificativi fiscali delle imprese e dei lavoratori. L'utilizzo del codice fiscale come chiave di collegamento garantisce, in linea teorica, un'elevata affidabilità dell'abbinamento, in quanto si tratta di

un identificativo univoco e trasversale alle diverse basi amministrative. Tuttavia, la presenza di universi solo parzialmente coincidenti impone un'attenta valutazione degli esiti del *linkage*. In primo luogo, occorre distinguere tra mancati abbinamenti 'leciti', dovuti a differenze strutturali tra gli archivi (ad esempio, imprese presenti in ASIA ma non soggette a Comunicazioni Obbligatorie in un determinato periodo, oppure lavoratori non inclusi nel campione Inps), e mancati abbinamenti imputabili a incongruenze negli identificativi o a disallineamenti temporali. La valutazione della qualità del *linkage* richiede pertanto l'analisi sistematica delle unità abbinate e non abbinate, nonché il confronto tra variabili comuni alle diverse fonti al fine di verificarne

la coerenza. Nel processo di integrazione assumono particolare rilievo anche le trasformazioni societarie e le modifiche degli identificativi nel tempo, che possono incidere sia sulla ricostruzione delle carriere lavorative sia sulla corretta attribuzione dei lavoratori alle imprese osservate nella RIL. Tali aspetti vengono affrontati mediante procedure di controllo incrociato e verifiche longitudinali sugli identificativi, al fine di preservare la continuità informativa delle unità osservate. L'integrazione micro tra archivi costituisce dunque una fase critica dell'intero processo, nella quale la dimensione tecnica del record linkage si intreccia con scelte metodologiche relative alla definizione dell'universo analitico e alla qualità delle informazioni utilizzate.

Bibliografia

- Abowd J.M., Kramarz F., Margolis D.N. (1999), High wage workers and high wage firms, *Econometrica*, 67, n.2, pp.251-333
- Andrews D., Criscuolo C., Gal P.N. (2016), *The Best versus the Rest: The Global Productivity Slowdown, Divergence across Firms and the Role of Public Policy*, OECD Productivity Working Papers n.5, Paris, OECD Publishing
- Babet D., Godechot O., Palladino M.G. (2025), In the Land of AKM: Explaining the Dynamics of Wage Inequality in France, *Journal of Labor Economics*, <DOI: 10.1086/740187>
- Banca d'Italia (2025), *Relazione annuale sul 2024*, Roma, Banca d'Italia
- Barth E., Bryson A., Davis J.C., Freeman R. (2016), It's where you work: Increases in earnings dispersion across establishments and individuals in the United States, *Journal of Labor Economics*, 34, n.52, pp.S67-S97
- Becker G.S. (1964), *Human capital: A theoretical and empirical analysis, with special reference to education*, New York, National Bureau of Economic Research
- Bonacini L., Patriarca F., Santoni E. (2025), Background wage premia, beyond education: Firm sorting and unobserved abilities of graduates, *Economics of Education Review*, 109, Article 102733
- Bonhomme S., Lamadon T., Manresa E. (2019), A distributional framework for matched employer-employee data, *Econometrica*, 87, n.3, pp.699-739
- Bugamelli M., Lotti F., Amici M., Ciapanna E., Colonna F., D'Amuri F., Giacomelli S., Linarello A., Manaresi F., Palumbo G., Scoccianti F., Sette E. (2018), *Productivity Growth in Italy: A Tale of a Slow-Motion Change*, Questioni di Economia e Finanza (Occasional Papers) n.422, Roma, Banca d'Italia
- Card D., Cardoso A.R., Heining J., Kline P. (2018), Firms and Labor Market Inequality: Evidence and Some Theory, *Journal of Labor Economics*, 36, n.51, pp.S13-S70
- Card D., Heining J., Kline P. (2013), Workplace heterogeneity and the rise of West German wage inequality, *Quarterly Journal of Economics*, 128, n.3, pp.967-1015
- Casarico A., Lattanzio S. (2024), What firms do: Gender inequality in linked employer-employee data, *Journal of Labor Economics*, 42, n.2, pp.325-355
- Devicienti F., Fanfani B., Maida A. (2019), Collective bargaining and the evolution of wage inequality in Italy, *British Journal of Industrial Relations*, 57, n.2, pp.377-407
- Inps (2024), *XXIII Rapporto annuale*, Roma, Inps
- Kline P., Petkova N., Williams H., Zidar O. (2019), Who profits from patents? Rent-sharing at innovative firms, *Quarterly Journal of Economics*, 134, n.3, pp.1343-1404
- Lachowska M., Mas A., Saggio R., Woodbury S. (2023), Do firm effects drift? Evidence from Washington administrative data, *Journal of Econometrics*, 233, n.2, pp.375-395

- Lamadon T., Mogstad M., Setzler B. (2022), Imperfect competition, compensating differentials, and rent sharing in the U.S. labor market, *American Economic Review*, 112, n.1, pp.169-212
- Naticchioni P., Ricci A., Rustichelli E. (2010), Far away from a skill-biased change: falling educational wage premia in Italy, *Applied Economics*, 42, n.26, pp.3383-3400
- OECD (2024), *OECD Employment Outlook 2024*, Paris, OECD Publishing
- Pan W. (2025), Firm-specific pay premia and the returns to higher education: Evidence from community colleges, *Labour Economics*, 84, Article 102490
- Song J., Price D.J., Guvenen F., Bloom N., von Wachter T. (2019), Firming up inequality, *Quarterly Journal of Economics*, 134, n.1, pp.1-50
- Pence M. (1973), Job market signaling, *Quarterly Journal of Economics*, 87, n.3, pp.355-374
- Van Reenen J. (1996), The creation and capture of rents: Wages and innovation in a panel of U.K. companies, *Quarterly Journal of Economics*, 111, n.1, pp.195-226

Irene Brunetti

i.brunetti@inapp.gov.it

Ricercatrice in economia applicata presso Inapp e responsabile scientifica di un accordo di collaborazione tra la Presidenza del Consiglio dei ministri e Inapp. I suoi interessi di ricerca riguardano principalmente le dinamiche del mercato del lavoro e la valutazione delle politiche pubbliche. Fra le pubblicazioni recenti si segnalano: Work-from-Home Job Creation as a Health-Risk Mitigation Strategy: Firm-Level Evidence from the COVID-19 Pandemic, *American Journal of Health Economics*, forthcoming, 2026; Peer Interactions, Local Markets, and Within Occupation Wages: Evidence from Italy, *Journal of Regional Science*, 2025.

Michelangelo Filippi

m.filippi@repnet.it

Ricercatore esperto della trasformazione a fini statistici dei dati amministrativi, collabora con università ed enti di ricerca pubblici e privati. È specializzato nell'elaborazione e analisi di dati micro. La sua attività di ricerca si concentra prevalentemente su vari aspetti del mercato del lavoro e della salute.

Valentina Gualtieri

v.gualtieri@inapp.gov.it

Appassionata di metodologia statistica è attualmente responsabile del Servizio statistico dell'Inapp e National Project manager del Programma OCSE-PIAAC (Programme for the International Assessment of Adult Competencies). È specializzata nella pianificazione di indagini statistiche in ambito sociale e nell'elaborazione e analisi di dati micro. La sua attività di ricerca si concentra prevalentemente sull'analisi di fenomeni legati all'investimento in istruzione, al capitale umano, al mercato del lavoro, alla disuguaglianza salariale e alla qualità del lavoro.

Andrea Ricci

an.ricci@inapp.gov.it

Dirigente di ricerca in economia ed economia applicata presso l'Inapp, dove coordina la Struttura di ricerca Imprese, transizioni e sviluppo del Mezzogiorno. I suoi principali interessi di ricerca hanno per oggetto la valutazione delle politiche per il lavoro, l'economia del mercato del lavoro, l'organizzazione industriale e il comportamento delle imprese. È autore di numerose ricerche pubblicate su riviste nazionali e internazionali.