IRENE BRUNETTI · MICHELANGELO FILIPPI · VALENTINA GUALTIERI · **ANDREA RICCI***

EDUCATION, WAGES AND FIRM HETEROGENEITY: EVIDENCE FROM LINKED EMPLOYER-EMPLOYEE DATA

*National Institute for the Analysis of Public Policies (INAPP)

- 67th Annual Scientific Meeting of the Italian Economic Association (SIE)

University of Cagliari, 23 October 2026

OUTLINE

National Institute for Public Policy Analysis

https://www.inapp.gov.it/en/institute

- Research, monitoring and evaluation of public policies in the fields of labour, education and training, social protection and labour economics, firms' behavior and industrial organizations;
- Data surveys and Microdata, National Statistical System;
- Research Unit «Firms, labour market transitions and the Development of the Mezzogiorno;
- *Enterprises and Labor Survey* (RIL) – National Statistical Programme (7th ongoing edition)

Administrative and survey data for public knowledge and policy

- Why using administrative data and national information system; opportunities and challanges
- Linked employer-employee data: Uniemens -Inps, Cob-Mlps, Asia-Istat, Inapp Surveys etc
- Research projects on «wages, occupations, education and firms behaviour»;
- Sinappsi Special Issue:
- Brunetti I., Filippi M., Gualtieri V., Ricci A.(2026), «*Istruzione, salari ed eterogeneità delle imprese: evidenze da dati employer-employee linked*», Sinappsi, XVI, n.2,pp.91-105

WHY LINKED EMPLOYER-EMPLOYEE DATA

- **Administrative records** (INPS, COB)
  - Accurate information on wages and work histories (Uniemens-Inps) + on occupations and workers' educations (Cob-Mlps);
  - Collected for administrative purposes: no information on firm performance, internal labor market, etc;
- **Firm surveys** (RIL)
  - Detailed information on management, competitive strategies, industrial relations, technology, etc;
  - No data on individual earnings or employment trajectories
- **Micro-level integration** of the three sources (INPS-COB-RIL);
- **Enabling condition:** Inapp's membership of the National Statistical System (SISTAN; Legislative Decree 322/1989; COMSTAT Directive 9/2004) grants access to administrative microdata;



THREE SOURCES, ONE INTEGRATED DATASET

Source	Content	Use in the analysis
UniEmens (INPS)	Actual earnings, weeks worked, contribution histories; sample based on 48 birth dates	Weekly wages and employment trajectories
Mandatory Notifications – COB (Ministry of Labour)	Hirings, conversions, extensions and separations; occupation and educational attainment	Graduate / non-graduate classification
RIL (Inapp)	Survey of about 30,000 firms, 2015, 2018 and 2022 waves; panel component of about 30%	Profitability, governance, industrial relations, management

Deterministic record linkage on tax identification codes; microdata accessed through SISTAN. Result: an INPS-COB-RIL linked employer-employee panel covering 2013-2022.



BUILDING THE EMPLOYER-EMPLOYEE DATASET

- **Starting point:** firms in the RIL sample (2015, 2018 and 2022 waves)
  - Corporations and partnerships in the private non-agricultural sector; reference population from the Istat Business Register (ASIA);
- **Linkage** to INPS (UniEmens) and SISCO-COB administrative records through the firm's tax code;
- **Worker selection:** all employees with at least one employment spell in a RIL firm in 2013-2022;
- **Complete work histories** reconstructed for each worker, across all employers
  - Worker mobility between firms identifies the firm fixed effects in the AKM model
- **Key Y and X variables**
  - Full-time equivalent weekly wage from UniEmens; education from COB (highest level observed, time-invariant); sales per employee, management and other firm-level characteristics from RIL
- **Final data structure:** multilevel worker-firm-period panel (2013-16, 2017-19, 2020-22), aligned with the RIL waves



PRE-PROCESSING, RECORD LINKAGE AND VALIDATION

- **Pre-processing of data deliveries** (Inapp Statistical Service): checked, standardised and pseudonymised files
  - SISCO: administrative events aggregated into employment spells; harmonisation of occupations and education
  - INPS: validation of contribution and earnings records; longitudinal reconstruction of annual firm-worker records
  - Istat (ASIA): alignment of classifications and coverage
- **Deterministic record linkage** on firm and worker tax codes
  - Distinguishing legitimate non-matches (structural differences across archives) from identifier inconsistencies
  - Handling corporate restructuring and changes in identifiers over time
- **Validation of the integrated data**
  - Temporal consistency of employment spells and wages; cross-checks on variables common to several sources

Details in the methodological appendix of the paper



RESEARCH PROJECT: MOTIVATION I

The Italian context

- Persistent stagnation of real wages (OECD 2024, 2025; INPS 2024; Bank of Italy 2025);
- Weak labour productivity and TFP growth constraining growth and wage dynamics (Bank of Italy 2024; Inapp 2025);
- Role of human capital in the technological, demographic and green transitions (Inapp Report on Workers' transitions, forthcoming);

The "puzzling" role of (Italian) firms

- Traditional theory of Human capital assigns firms a "marginal" role in shaping wages: workers with the same education should earn similar wages regardless of the employer
- Evidence based on linked employer-employee data (and AKM literature) shows that firm-specific policy affects wage premia - also accounting for workers' heterogeneity and assortative matching
- Returns to education may therefore also depend on when, where and why graduates are employed
- Inside the «black box» of firms' decisions on wage-setting at the workplace;

REASERCH PROJECT: MOTIVATION II

To what extent do firm characteristics - governance, industrial relations and profitability etc - explain the heterogeneity in the returns to university education, through rent-sharing mechanisms or pay policies favouring high-skilled workers?

The paper makes three contributions:

1. **New evidence** on the returns to workers' education, beyond human capital and sorting
2. **Integration of administrative and survey data**, which makes it possible to link time-variant firm specific wage policy to workers' education
3. **Microeconomic determinants of firm-specific wage setting process:** how firm pay policies and organisation shape the economic return to investment in education

Some literature

- **Firm heterogeneity and wage inequality**: Card, et al 2013, 2018; Song et al. 2019; Saggio et al 2024; Devicienti et al, etc;
- **Rent sharing and firm characteristics:** Kline et al. (2019), Lamadon et al (2022), Fana et al (2026) etc;
- **Firm pay premiums and returns to education:** Pan, 2025; Bonacini et al. 2025;
- Many others



EMPIRICAL STRATEGY I

Two step econometric strategy that combines:

- First-step regressions: extended time-varying version of AKM models (Lachowska et al, 2023; Kleine, 2025; Palladino et al 2024) on COB-INPS data*;
- Second-step regressions: OLS and FE models on COB-INPS-RIL data;

* We use matched employer-employee administrative records to run:

$$\ln w_{ijt} = \alpha_i + \boldsymbol{\psi_{jt}} + \varepsilon_{ijt} \quad (1)$$

where w_{ijt} fte weekly wage, α_i is (time-invariant) worker unobserved heterogeneity, $\boldsymbol{\psi_{jt}}$ is a firm-by-year wage component - i.e. *dynamics of firm specific pay policy* - ε_{ijt} idiosyncratic terms; t=2013-2022

- *What does* $\boldsymbol{\psi_{jt}}$ *reflect?* Changes in rent sharing, productivity, bargaining, unobserved job attributes, etc
- *Why does we depart from standard AKM specification?*
- *Identification:* estimating eq.(1) on largest connected sets



EMPIRICAL STRATEGY II

- **Step 1 · AKM model, estimated separately for graduates (L) and non-graduates (NL) on INPS-COB merged data**

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

- Firm effects allowed to vary across three periods (2013-16, 2017-19, 2020-22); estimated on the largest connected set;
- we estimate the firm-specific component of the wage difference between graduate and non-graduate workers: $\psi^{L}_{jt} - \psi^{NL}_{jt}$

- **Step 2 · Firm fixed-effects regression on INPS-COB-RIL merged data**

$$\psi^{L}_{jt} - \psi^{NL}_{jt} = \beta_0 + \beta_1 \ln\Pi_{jt} + \beta_2 \ln\Pi_{jt} \times 1(2017\text{–}19) + \beta_3 \ln\Pi_{jt} \times 1(2020\text{–}22) + \delta' F_{jt} + \mu_j + \varepsilon_{jt}$$

- $\psi^{L}_{jt} - \psi^{NL}_{jt}$ is the tertiary education-related wage differential paid by firm (j) in period (t).
- Π_{jt}: sales per employee (RIL); β_2, β_3 capture how the link between firm performance and the graduate premium evolves over time
- F_{jt}: management, governance, industrial relations and workforce organisation;

DESCRIPTIVE STATISTICS

Table 1 shows descriptive statistics for graduates and non-graduates, distinguishing the full sample and the largest connected set – pooled over the period 2013-2022

	Full sample		Largest connected set	
	Mean	**Std. dev.**	**Mean**	**Std. dev.**
	Graduates			
Log (wage)	6.415	0.541	6.421	0.542
Age (years)	37.18	9.148	37.13	9.111
Female	0.563	0.496	0.567	0.496
Observations	2,102,115		1,967,173	
	Non-graduates			
Log (wage)	6.195	0.426	6.194	0.426
Age (years)	40.63	11.22	40.54	11.21
Female	0.387	0.487	0.386	0.486
Observations	8,442,235		8,039,660	

Source: authors' calculations on INPS-COB data, 2013-2022. Employees aged 20-60 in private non-agricultural firms with at least 10 employees.

- Over 10 million observations on about 2 million workers, of whom about 400,000 are graduates
- Graduates earn more, are younger and have a higher incidence of females than non-graduates; restricting to the largest connected set leaves observables unchanged



AKM ESTIMATES: VARIANCE DECOMPOSITION AND SORTING

Table 2: Variance-covariance matrix of AKM components – pooled over the period 2013-2022

	α_i	ψ_{jt}	$X_{it}\beta$	Ln(wage)
		Graduates		
α_i	0.146			
ψ_{jt}	0.006	0.078		
$X_{it}\beta$	−0.019	0.003	0.062	
Ln(wage)	0.133	0.087	0.047	0.294
		Non-graduates		
α_i	0.099			
ψ_{jt}	−0.007	0.070		
$X_{it}\beta$	−0.015	0.002	0.028	
Ln(wage)	0.077	0.065	0.016	0.181

Source: INPS-COB data, 2013-2022

First step results:

- Concave age-earnings profile, steeper for graduates (0.083 vs 0.041): faster wage growth in early and mid-career
- Worker and firm-period fixed effects explain most of the wage variance (R^2 = 0.87 and 0.83)

- Wage dispersion is driven mainly by worker heterogeneity: Var(α) = 0.146 vs Var(ψ) = 0.078 for graduates;
- Weakly positive assortative matching for graduates (Cov(α, ψ) = 0.006), slightly negative for non-graduates (−0.007);
- Given limited mobility bias, covariances are best read in comparative terms across the two groups



FIRM PERFORMANCE AND THE GRADUATE PREMIUM

Table 3: Estimation results. Dep var.: graduate vs non-graduate firm pay premium ($\psi^{L} - \psi^{NL}$) by firm and period

	(1) OLS	(2) Firm FE
$\ln(\Pi_{jt}) \times 1(2020–22)$	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]
Other controls	Yes	Yes
Firm fixed effects	No	Yes
Constant	0.020	0.038
	[0.054]	[0.040]
Firms	1,247	1,247
Observations	3,502	3,502
R^2	0.027	0.752

*Sources: authors' calculations on INSP-COB-RIL data. Controls: management education and gender, family governance, firm size and age; region and 2-digit sector FE. Firm-clustered SE. *** p<0.01, ** p<0.05, * p<0.10*

Second step results:

- No significant link with OLS; with firm FE, negative in 2013-16 (−0.9%) and positive and significant in 2020-22
- "Productivity" gains are passed on to graduates only partially and gradually, consistent with collective bargaining (?)



MAIN FINDINGS

1. **Firm pay policies matter, but only weakly:** their effect on graduate wage premiums is positive but small, as premiums are driven mainly by individual ability and unobserved heterogeneity.

2. **Sorting is not uniform across the human capital distribution:** assortative matching is weakly positive for graduates and negative for less-skilled workers.

3. **Rent sharing towards graduates is slow and partial:** Italian firms do not automatically translate competitive performance into higher wages for high-skilled workers.

POLICY AND FUTURE RESEARCH

Education and labour market policies

- Returns to education are not merely a reflection of individual productivity;
- The Italian production structure contributes to compressing the graduate wage premium;
- Policy should also foster graduates' access to high-quality firms and work environments that value their skills;

The value of linked employer-employee data

- INPS-COB-RIL integration enables analyses that no single source allows;
- SISTAN turns administrative microdata into a research infrastructure;
- Next step: see Inapp research programs;

Thank you

an.ricci@inapp.gov.it

INAPP
ISTITUTO NAZIONALE PER L'ANALISI
DELLE POLITICHE PUBBLICHE

www.inapp.gov.it