

No. 2225

October 2026

Which types of education raise firm productivity? Evidence for managers and workers in the UK

Tessa Hall

Abstract

This paper estimates the firm productivity returns to degree field, institution quality, and degree level, and compares them with wage returns. I link a LinkedIn panel from Revelio Labs to Orbis financial accounts for over 20,000 UK firms over 2015–2024. Within firms, a bachelor's degree raises a worker's productivity by around 20%; a top-20 institution adds a further 10% and a postgraduate qualification 6%. The returns differ by organisational role. A bachelor's degree raises productivity for managers and non-managers alike. Specialist credentials - a STEM field, an elite institution - raise productivity for non-managers but not for managers, whose premium loads instead on postgraduate qualifications, above all MBAs. Managers also contribute around 26% more than otherwise identical non-managers, a premium that education does not explain. Averaged over the workforce, wage premia are close to productivity premia for each type of degree. However, the average conceals two wedges of opposite sign: STEM-trained non-managers are paid below their relative productivity, and degree-holding managers above it. Finally, for a subset of firms matched to the World Management Survey, workforce education accounts for around a quarter the well-documented correlation between management practices and productivity.

Keywords: human capital, firm productivity, education returns, management
JEL codes: J24, J31, D24, M11

This paper was produced as part of the Labour Programme. The Centre for Economic Performance is financed by the Economic and Social Research Council.

I thank John Van Reenen and Alan Manning for supervision and comments. I am grateful to seminar participants at the Centre for Economic Performance for helpful discussion. This paper uses data from Revelio Labs and Orbis (Bureau van Dijk), and the World Management Survey. All errors are my own. Claude Code (Sonnet 4.6, Opus 4.8) was used to help with coding in R for the data analysis and cleaning.

Tessa Hall, Centre for Economic Performance at the London School of Economics.

Published by
Centre for Economic Performance
London School of Economic and Political Science
Houghton Street
London WC2A 2AE

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means without the prior permission in writing of the publisher nor be issued to the public or circulated in any form other than that in which it is published.

Requests for permission to reproduce any article or part of the Working Paper should be sent to the editor at the above address.

© T. Hall, submitted 2026.

1. Introduction

The wage returns to different types of education are well documented. In England, administrative data show that the degree premium at age 40 ranges from over 100% for medicine to negative for men in creative arts, and that attending a Russell Group university adds a further 20% for men (Britton et al., 2016, 2020). In the United States, lifetime earnings vary more across college majors than between college graduates and non-graduates (Altonji et al., 2012; Hershbein and Kearney, 2014). Whether the degrees with the highest wage returns are also those that raise firm output the most is an open question. The firm-level evidence is thin and coarse. Existing studies distinguish at best a degree from no degree (Hellerstein et al., 1999; Galindo-Rueda and Haskel, 2005; Lebedinski and Vandenberghe, 2014). None distinguish field, institution, or degree level, and none ask whether the returns depend on where in the organisation the worker sits.

This paper provides firm-level evidence on the productivity and wage returns to degree field, institution quality, and degree level, separately for managers and non-managers. I link a LinkedIn panel from Revelio Labs to Orbis financial accounts for over 20,000 UK firms over 2015–2024. By 2024 the panel covers over 55% of UK employment and records each worker’s degree subject, institution, and occupation. I embed these worker types in an efficiency-units production function (Hellerstein et al., 1999), estimate an analogous wage equation, and compare the two. For a subset of manufacturing firms matched to the World Management Survey, I then ask how much of the well-documented effect of management practices on productivity (Bloom and Van Reenen, 2007) runs through the education of the workforce.

Four findings emerge. First, the type of education matters for productivity, and institution quality matters most. Within firms, a bachelor’s degree raises a worker’s efficiency units by around 20% relative to no degree; a top-20 institution adds a further 10% and a postgraduate qualification 6%. The cross-sectional premia are two to five times larger, which reflects the sorting of educated workers into productive firms.

Second, the returns to education depend on the worker’s role, and the pattern is systematic. A bachelor’s degree raises productivity for managers and non-managers alike. What differs is the type of degree. Specialist credentials - a STEM field, an elite institution - carry large premia for non-managers that survive firm fixed effects, and no premium for managers. For managers, the premium loads instead on postgraduate qualifications, which is largely an MBA effect. A manager with no degree also contributes around 26% more efficiency units than an otherwise identical non-manager, and this premium barely moves when education is controlled. Whatever managers contribute, their education captures only part of it.

One potential reading is that human capital routes workers onto two career tracks. On the professional and technical track, specialist knowledge is an input to production throughout a career, so STEM and elite credentials keep paying. On the managerial track, the job rewards breadth - judgement across domains and the coordination of others’ expertise (Garicano, 2000)

- so specialist credentials matter on this track only at the entry margin. Firms promote on performance in the current role, so they promote their best specialists, and the promotion data confirm that a STEM degree is the strongest single predictor of reaching management. Once promoted, the specialist's depth no longer raises output. This is the Peter Principle ([Peter and Hull, 1969](#); [Lazear, 2004](#)), documented by [Benson et al. \(2019\)](#) for salespeople: the traits that earn promotion are not the traits that make a good manager.

Third, the relative wage and productivity premia for each degree type follow each other closely on average, but the average misleads. Pooling the workforce, the within-firm wedges are small, and only the degree shows a gap, with a relative wage premium 6 log points above its productivity premium. Splitting by role shows that the small pooled wedges are sums of two wedges of opposite sign. Within firms, STEM-trained non-managers earn a relative wage premium that falls short of their relative productivity premium by 7 log points; degree-holding managers earn one that exceeds theirs by 13 log points. The STEM gap has an incentive implication. If wages understate the productivity of specialist skills, the private return to acquiring them falls short of the firm return, before counting the spillovers from technical human capital that push the social return higher still ([Moretti, 2004](#); [Bloom et al., 2013](#)).

Fourth, workforce education accounts for a substantial share of the management–productivity relationship. For the manufacturing firms matched to the World Management Survey, the education of managers and workers explains around 26% of the correlation between management score and productivity, close to the roughly 30% that [Bender et al. \(2018\)](#) attribute to workforce selection and pay premia in Germany. The explained share runs through the general education of the workforce rather than through managerial education. That is consistent with the production-function result: the education of managers captures only part of what they contribute.

The estimates are conditional correlations, not causal returns. The LinkedIn data raise two measurement concerns: sampling error in the education shares, which attenuates the estimates, and the over-representation of educated workers, which biases the shares upward in thinly covered firms. I address both with coverage weighting, occupation sampling weights, and a coverage control, and I show that the estimates rise as coverage improves. Firm fixed effects absorb time-invariant sorting but worsen attenuation for slowly varying shares; long-difference estimates in [Appendix E.9](#) partly recover the lost signal. Lagged shares argue against simple reverse causality ([Appendix E.8](#)), but ability bias and endogenous promotion remain. The perfect-substitutes assumption is relaxed in [Appendix E.6](#). The framework and the data open a route to stronger identification - for example, distance to universities of different types - that I leave to future work.

Section [2](#) sets out the framework. Section [3](#) describes the LinkedIn–Orbis panel. Section [4](#) sets out the estimating equations and identification. Section [5](#) presents results. Section [6](#) decomposes the management channel. Section [7](#) concludes.

1.1 Related Literature

A large literature documents the individual returns to schooling ([Becker, 1964](#); [Mincer, 1974](#); [Card, 1999](#)). The returns vary widely by type of education. For England, [Britton et al. \(2016\)](#) use linked tax and student loan records to show that graduate earnings differ substantially across subjects and institutions, even controlling for prior attainment and family background, and [Britton et al. \(2020\)](#) extend this to lifetime earnings: the degree premium grows sharply between ages 30 and 40 but ranges from 90–110% for medicine and economics to negative for men in creative arts. For the United States, [Altonji et al. \(2012\)](#) and [Altonji et al. \(2016\)](#) show that the earnings differences across majors are as large as the college premium itself. [Kirkeboen et al. \(2016\)](#) use admission cutoffs in Norway to show that these differences are causal. All of this work observes the worker’s pay and not the employer’s output.

A smaller literature links workforce composition to firm productivity through the efficiency-units framework. [Hellerstein et al. \(1999\)](#) pioneered the approach, recovering marginal productivity differentials between worker types in US manufacturing and comparing them with wage differentials, with a focus on gender and race. [Hægeland and Klette \(1999\)](#) apply it to Norwegian plant–worker data and find that the higher wages of educated workers broadly reflect higher productivity. For the UK, [Galindo-Rueda and Haskel \(2005\)](#) link the Annual Business Inquiry to the Employer Skills Survey and find that establishments employing more graduates are more productive, though with a small cross-sectional sample their premia are imprecise. [Dearden et al. \(2006\)](#) estimate the same production and wage system on a panel of British industries and find that training raises productivity by around twice as much as it raises wages. The most rigorous study is [Lebedinski and Vandenberghe \(2014\)](#), who combine firm fixed effects with the control function of [Akerberg et al. \(2015\)](#) on Belgian panel data and estimate that graduates are 23–42% more productive than non-graduates, with labour costs broadly aligned. Two gaps remain. These studies distinguish at best a degree from no degree, yet field, institution, and degree level are where the individual returns vary most. And they treat the workforce as a single pool, although in [Lucas \(1978\)](#) and [Garicano \(2000\)](#) managerial talent multiplies subordinates’ output in a way non-managerial skill does not. The LinkedIn–Orbis panel allows this paper, to my knowledge the first, to fill both gaps.

A distinct literature isolates the effect of individual top managers on firm performance. [Bertrand and Schoar \(2003\)](#) show that manager fixed effects, identified from executives who move between firms, explain a substantial share of the variation in investment, financial, and organisational policy. Subsequent work links CEO traits and behaviour to performance and documents large returns to talent at the top of the distribution ([Kaplan et al., 2012](#); [Bandiera et al., 2020](#)). That literature is largely silent on the workforce beneath the executive. Appendix E.2 bridges the two by splitting managers into frontline, middle, and executive tiers.

Comparing wage and productivity premia requires a reason for the two to diverge, and the candidate mechanisms differ in which workers they predict are paid below their marginal product. Under monopsony, search frictions give employers wage-setting power ([Manning,](#)

2003; Burdett and Mortensen, 1998); the markdown is smaller for groups with better outside options. Part of this margin runs through amenities: Card et al. (2016) find that women capture a smaller share of firm-specific rents than men. Under Nash bargaining over match surplus (Cahuc et al., 2006), a worker's wedge reflects bargaining weight and surplus, both of which plausibly rise with seniority. Efficiency wages lift pay above marginal product to elicit effort (Shapiro and Stiglitz, 1984). Wage back-loading (Lazear, 1979) produces markdowns early in the career and supra-competitive pay later. If managerial matches carry more surplus and managers capture a larger share of it, the managerial wedge is positive. If STEM-trained non-managers supply labour inelastically to their employer, because their skills are deployed at few local employers, their markdown is larger. This paper measures the wedges and leaves the adjudication between mechanisms to future work.

The policy implication follows Acemoglu and Pischke (1999): where graduates face wage markdowns relative to marginal product, the private incentive to acquire the credential falls short of its productivity return. Technical human capital also generates spillovers through R&D that reach beyond the employing firm (Moretti, 2004; Bloom et al., 2013), so the social return to STEM skills exceeds the private return even when wages are competitive. A markdown widens the gap further.

A parallel literature establishes management practices as a first-order determinant of firm performance, accounting for large and persistent productivity gaps across firms and countries (Bloom and Van Reenen, 2007; Bloom et al., 2016). The human capital of workers and managers is a natural channel through which management quality varies. Bender et al. (2018) provide the closest estimate, linking German manufacturing employer–employee data to World Management Survey scores. They find that workforce human capital and pay premia jointly account for around 30% of the productivity–management association, with the human capital of the highest-paid workers, their proxy for managers, explaining around 13%. This paper replaces their AKM-based ability measures (Abowd et al., 1999) with education-based shares constructed from LinkedIn.

The data are themselves a contribution. Tambe et al. (2020) use LinkedIn IT labour records to measure digital capital. Benzell et al. (2019) use occupational histories to decompose skill-biased technological change. Jeffers (2024) link employment histories to firm investment. Dorn et al. (2025) use self-reported skill sets of US college graduates on LinkedIn to measure multidimensional human capital. For the UK, Exadaktylos et al. (2024) link LinkedIn-derived manager histories to firm accounts and find that hiring foreign managers raises productivity, with gains driven by prior industry experience abroad – evidence that managerial value travels through embodied knowledge that education level does not capture.

2. Framework

Section 2.1 embeds heterogeneous worker types in a Cobb–Douglas production function and derives the estimating equation for the productivity premia. Section 2.2 derives a parallel wage equation and defines the wage–productivity wedge for each education–role cell.

2.1 Production Functions with Heterogeneous Human Capital

Firm f produces value added VA in period t according to

$$VA_{ft} = A_{ft} K_{ft}^{\alpha} \tilde{L}_{ft}^{\beta} \quad (1)$$

where A_{ft} is total factor productivity, K_{ft} is capital, and $\tilde{L}_{ft}$ is a human-capital-augmented labour aggregate. The modelling choice is how to specify $\tilde{L}_{ft}$. Workers differ in role, degree attainment, field, institution, demographics, and experience. The aggregate must be rich enough to capture this heterogeneity and simple enough to estimate. I build it in three steps, each resting on an explicit assumption.

Assumption 1 (Perfect substitutes across groups) *The marginal rate of technical substitution between a manager and a non-manager is constant at $(1 + \lambda)$, regardless of their mix.*

Under Assumption 1 the labour aggregate is

$$\tilde{L}_{ft} = L_{ft}^{nm} + (1 + \lambda)L_{ft}^m \quad (2)$$

where L_{ft}^m and L_{ft}^{nm} are the headcounts of managers and non-managers. This rules out complementarity between the two groups, including the Lucas (1978) span-of-control mechanism by which managers raise the productivity of those beneath them. Appendix E.6 relaxes the assumption. The restriction is rejected statistically, so the linear aggregate is a reasonable approximation.

Writing $s_{ft}^m = L_{ft}^m / L_{ft}$ for the manager share, so that $\tilde{L}_{ft} = L_{ft}[1 + \lambda s_{ft}^m]$, substituting into (1), taking logs, and applying $\ln(1 + z) \approx z$ gives

$$\ln\left(\frac{VA}{L}\right)_{ft} \approx \ln A_{ft} + \alpha \ln\left(\frac{K}{L}\right)_{ft} + (\alpha + \beta - 1) \ln L_{ft} + \beta \lambda s_{ft}^m \quad (3)$$

The coefficient on s_{ft}^m identifies $\beta\lambda$, and β is recovered from the capital and labour coefficients. So λ , the proportional efficiency-unit premium of a manager over a non-manager, is identified.

Workers also differ in human capital. A fully non-parametric treatment would need a separate labour type for every combination of characteristics. I therefore impose additivity.

Assumption 2 (Additive efficiency units) *Human capital characteristics are perfect substitutes and enter the efficiency unit premium additively.*

Under Assumption 2 alone, abstracting from role, the labour aggregate is

$$\tilde{L}_{ft} = L_{ft} \left(1 + \sum_k \mu_k \bar{x}_{k,ft} \right) \quad (4)$$

where $\bar{x}_{k,ft}$ is the share of workers at firm f in period t with characteristic k , and μ_k is the proportional efficiency-unit premium of characteristic k relative to the base category (in the main specification, a female worker with no degree and no firm experience). Additivity rules out interactions between characteristics. Appendix E.6 adds pairwise interactions among the education shares. Additivity is rejected, driven by a complementarity between degree attainment and institution quality, but within firms the interactions are small. Substituting (4) into (1) and log-linearising gives the pooled estimating equation:

$$\ln \left(\frac{VA}{L} \right)_{ft} \approx \ln A_{ft} + \alpha \ln \left(\frac{K}{L} \right)_{ft} + (\alpha + \beta - 1) \ln L_{ft} + \beta \sum_k \mu_k \bar{x}_{k,ft} \quad (5)$$

The characteristic set k is the same in the production function and the wage equation throughout: degree attainment and, conditional on holding a degree, a postgraduate qualification, a STEM field, and a top-20 institution, as defined in Section 3.2; plus the male share and experience controls of Section 4.1. The conditional construction gives the education coefficients a sequential reading: μ_{deg} is the premium of a bare bachelor's degree over no degree, while μ_{post} , μ_{stem} , and μ_{elite} are the additional premia from each characteristic, conditional on a degree. The final assumption lets these premia differ by role.

Assumption 3 (Heterogeneous returns across groups) *Within each occupational group g , human capital characteristics enter the efficiency unit premium additively as in Assumption 2, but the efficiency unit premia μ_k^g are unrestricted across groups.*

The framework extends to N mutually exclusive groups. The main text uses two, $g \in \{m, nm\}$; Appendix E.2 uses four, splitting managers into frontline, middle, and executive tiers. Combining Assumptions 1–3, the full labour aggregate is

$$\tilde{L}_{ft} = L_{ft}^{nm} \left(1 + \sum_k \mu_k^{nm} \bar{x}_{k,ft}^{nm} \right) + (1 + \lambda) L_{ft}^m \left(1 + \sum_k \mu_k^m \bar{x}_{k,ft}^m \right) \quad (6)$$

Substituting (6) into (1), factoring out L_{ft} , and rearranging gives

$$VA_{ft} = A_{ft} K_{ft}^\alpha L_{ft}^\beta \left[1 + \lambda s_{ft}^m + (1 - s_{ft}^m) \sum_k \mu_k^{nm} \bar{x}_{k,ft}^{nm} + (1 + \lambda) s_{ft}^m \sum_k \mu_k^m \bar{x}_{k,ft}^m \right]^\beta \quad (7)$$

Applying $\ln(1 + z) \approx z$ to the bracketed term and writing the result as log value added

per worker gives the baseline estimating equation:

$$\begin{aligned} \ln\left(\frac{VA}{L}\right)_{ft} &\approx \ln A_{ft} + \alpha \ln\left(\frac{K}{L}\right)_{ft} + (\alpha + \beta - 1) \ln L_{ft} \\ &+ \underbrace{\beta \lambda s_{ft}^m}_{\text{manager premium}} + \underbrace{\beta(1 - s_{ft}^m) \sum_k \mu_k^{nm} \bar{x}_{k,ft}^{nm}}_{\text{non-manager human capital}} + \underbrace{\beta(1 + \lambda) s_{ft}^m \sum_k \mu_k^m \bar{x}_{k,ft}^m}_{\text{manager human capital}} \end{aligned} \quad (8)$$

The log-linearisation requires $z_{ft} \ll 1$, where

$$z_{ft} = \lambda s_{ft}^m + (1 - s_{ft}^m) \sum_k \mu_k^{nm} \bar{x}_{k,ft}^{nm} + (1 + \lambda) s_{ft}^m \sum_k \mu_k^m \bar{x}_{k,ft}^m \quad (9)$$

and the approximation error is of order $z_{ft}^2/2$. Section 4.2 bounds the error, and Appendix E.7 drops the approximation with nonlinear least squares.

Equation (8) yields two production-side hypotheses. The first asks whether human capital is rewarded differently by role. The natural null compares the within-group premia, $H_0^\mu: \mu_k^m = \mu_k^{nm} \forall k$. But a within-manager premium operates on the higher base $(1 + \lambda)$, so the economically sharper question is whether a characteristic makes the same total marginal product contribution in either role:

$$H_0^{MP}: (1 + \lambda) \mu_k^m = \mu_k^{nm} \quad \forall k \quad (10)$$

The second concerns λ itself. Conditional on the human capital of both groups, λ is the residual premium of a manager with no degree over an otherwise identical non-manager, and $H_0^\lambda: \lambda = 0$ tests whether it is zero. Rejection implies a managerial premium not explained by observed human capital, which may reflect the role itself or the sorting of unobservably able workers into it.

2.2 The Wage–Productivity Comparison

Under competitive labour markets, workers are paid their marginal products. Any departure – monopsony, search frictions, bargaining, efficiency wages, discrimination, or dynamic incentive contracts – shows up as a wedge between the wage premium and the productivity premium of a worker type. I recover the wage premia from a regression of log average wages on the same human capital shares and compare them with the productivity premia.

Under Assumption 1, the wage bill is

$$W_{ft} = w_{ft}^m L_{ft}^m + w_{ft}^{nm} L_{ft}^{nm} \quad (11)$$

where w_{ft}^{nm} is the average non-manager wage and $w_{ft}^m = (1 + d^\lambda) w_{ft}^{nm}$, so that d^λ is the

proportional wage advantage of managers. Dividing by L_{ft} and applying $\ln(1+z) \approx z$:

$$\ln\left(\frac{W}{L}\right)_{ft} \approx \ln w_{ft}^{nm} + d^\lambda s_{ft}^m \quad (12)$$

Under Assumption 2 alone, with w_{ft}^1 the wage of a base-category worker and d_k the proportional wage advantage of characteristic k , the wage bill is $W_{ft} = w_{ft}^1 L_{ft} [1 + \sum_k d_k \bar{x}_{k,ft}]$, so that

$$\ln\left(\frac{W}{L}\right)_{ft} \approx \ln w_{ft}^1 + \sum_k d_k \bar{x}_{k,ft} \quad (13)$$

Under Assumptions 1–3, the full wage bill is

$$W_{ft} = w_{ft}^{nm,1} L_{ft} \left[1 + d^\lambda s_{ft}^m + (1 - s_{ft}^m) \sum_k d_k^{nm} \bar{x}_{k,ft}^{nm} + (1 + d^\lambda) s_{ft}^m \sum_k d_k^m \bar{x}_{k,ft}^m \right] \quad (14)$$

where $w_{ft}^{nm,1}$ is the wage of a base-category non-manager, absorbed into fixed effects in estimation; d^λ is the wage advantage of a base-category manager over a base-category non-manager; and d_k^{nm} and d_k^m are the within-group wage advantages of characteristic k . Taking logs and applying $\ln(1+z) \approx z$:

$$\ln\left(\frac{W}{L}\right)_{ft} \approx \ln w_{ft}^{nm,1} + d^\lambda s_{ft}^m + (1 - s_{ft}^m) \sum_k d_k^{nm} \bar{x}_{k,ft}^{nm} + (1 + d^\lambda) s_{ft}^m \sum_k d_k^m \bar{x}_{k,ft}^m \quad (15)$$

The wage–productivity wedge. For each education–role cell, the wedge is the gap between the relative wage premium and the relative productivity premium:

$$\phi^\lambda \equiv d^\lambda - \beta\lambda \quad (\text{manager–non-manager relative wage gap minus marginal product})$$

$$\phi_k^{nm} \equiv d_k^{nm} - \beta\mu_k^{nm} \quad (\text{within-non-manager wage premium for } k \text{ minus marginal product})$$

$$\phi_k^m \equiv (1 + d^\lambda) d_k^m - \beta(1 + \lambda)\mu_k^m \quad (\text{within-manager wage premium for } k \text{ minus marginal product})$$

Each ϕ is zero under perfect competition. Since the productivity parameters come from (8) and the wage parameters from (15), testing $H_0^\phi: \phi_k^g = 0$ is a joint test across the two equations. I estimate them as a seemingly unrelated regression (SUR) system, as detailed in Section 4.1, so that the standard errors account for the correlation between the two residuals that common firm-level shocks induce. The same system tests whether the wedges differ by role, $H_0^{\phi,eq}: \phi_k^m = \phi_k^{nm}$.

Two caveats apply throughout. First, because the base wage $\ln w_{ft}^{nm,1}$ is absorbed into fixed effects, every wedge is about *relative* wages. A finding of $\phi^\lambda = 0$ is consistent with both groups being paid their marginal products, and with a uniform markdown applied to both. Second, each wedge is a group average, so a zero wedge is consistent with offsetting heterogeneity within the group.

3. Data

The estimating equations need two things that standard sources do not jointly provide: worker-level education records with field, institution, and degree level, and matched firm accounts. I link a LinkedIn panel to Orbis financials. Section 3.1 describes the LinkedIn panel and its coverage. Section 3.2 describes the matched estimation panel and the firm-level variables. Descriptive statistics on managers and non-managers are in Appendix A.

3.1 The LinkedIn Panel

The LinkedIn panel, from Revelio Labs, covers UK workers with an active profile between 2010 and 2025. For each individual I observe: (i) a full job history, with each position classified to an ONET 6-digit occupation code, a seniority level, a role title, an imputed salary, and start and end dates; (ii) education level, degree subject, and institution, which I map to university rankings in the year of graduation; and (iii) full name, from which Revelio infers gender and ethnicity probabilities. I drop job-year observations with a missing start date or role title.

Figure 1 plots LinkedIn coverage against Annual Population Survey (APS) employment. Coverage has grown over the sample period and has been above 55% and stable since 2021. It is not uniform across occupations. Table 1 compares occupation group shares in LinkedIn and the APS for 2025: managers and professionals are over-represented, and manual and industrial occupations are under-represented. This is a known feature of professional networking platforms and motivates the individual sampling weights, described below, that rebalance the LinkedIn population toward labour force totals by occupation.¹

Manager definition. SOC major group 11 spans a wide range of roles, from team leaders to chief executives. I classify workers in this group into three tiers, applied sequentially, using ONET codes and Revelio seniority scores:²

1. **Executive managers:** workers in top-executive ONET codes³ with a seniority score of 5 or above, corresponding to Director level and higher. 3% of workers in the 2025 cross-section.
2. **Middle managers:** workers in functional, operational, and professional-services management codes⁴ with a seniority score of 4 or above, corresponding to Manager level. 10% of workers.

¹Revelio maps each role title to a 6-digit occupation code and each profile to a country, then computes inverse-probability weights by comparing the occupation-by-country composition of the LinkedIn sample to International Labour Organization workforce totals; see <https://www.data-dictionary.reveliolabs.com/methodology.html>.

²Seniority scores are constructed by Revelio from an ensemble model combining job title, company, industry, and employment history.

³Executive management ONET codes: 11-1000, 11-1010, 11-1011, 11-1020, 11-1021, 11-1030, 11-1031.

⁴Middle management ONET codes: 11-2010, 11-2011, 11-2021, 11-2030, 11-2031, 11-2032, 11-2033, 11-3012, 11-3020, 11-3021, 11-3030, 11-3031, 11-3050, 11-3051, 11-3060, 11-3061, 11-3070, 11-3071, 11-3110, 11-3111, 11-3121, 11-3130, 11-3131, 11-9030, 11-9032, 11-9033, 11-9039, 11-9040, 11-9041, 11-9110, 11-9111, 11-9120, 11-9121, 11-9160, 11-9161, 11-1030, 11-1031.

3. **Frontline managers:** all remaining SOC major group 11 workers, mainly first-line supervisors and team leaders. 12% of workers.

Throughout, “managers” means executive and middle managers (13% of workers) and “non-managers” means everyone else, including frontline managers. Appendix E.2 estimates separate parameters for all three tiers.

Appendix Tables 10 and 11 describe the 2025 cross-section. Three facts matter for what follows. Managers hold degrees at a similar rate to the average worker (58% against 54%), so management is reached by several routes. Conditional on a degree, managers hold MBAs at three times the non-manager rate (9% against 3%); this is the sharpest educational difference between the two groups, with a standardised gap of 0.30 against 0.04–0.15 for the other credentials. And managers are promoted at substantially higher rates (40% against 24%).

3.2 The LinkedIn–Orbis Estimation Panel

The firm accounts come from Orbis (Bureau van Dijk) for UK companies over 2015–2024. Orbis coverage deteriorates in 2025 and LinkedIn coverage thins before 2015. The key variables are value added, employment, staff costs, and total assets, which I use as the capital measure. Value added is reported directly for 58% of firms. For the rest I construct it as EBITDA plus staff costs, following Kalemli-Ozcan et al. (2015); where both are available, their log correlation is 0.99, and Appendix E.5 shows the results are unchanged on the directly reported subsample.

Table 2 documents six sequential sample restrictions:

1. **Employment size.** Firms with ≥ 50 employees, so that workforce shares are estimated with reasonable within-firm precision.
2. **Consolidation.** Following Kalemli-Ozcan et al. (2015), I keep the lowest-level entity in each ownership chain. A firm is dropped if its BvD identifier appears as the Domestic Ultimate Owner of another firm in the sample. The Global Ultimate Owner identifier flags foreign-owned firms for the multinational check in Appendix E.5.
3. **Non-missing financials.** Non-missing and non-zero employment, value added, staff costs, and total assets.
4. **LinkedIn match.** A two-stage match: exact on Legal Entity Identifier and firm name (79% of the target sample), then fuzzy name matching (maximum distance 0.08) within groups sharing a postcode, phone number, or website (a further 13%). In total 92% of eligible Orbis firms are matched. Appendix Table 9 shows that matched firms do not differ from the Orbis universe on observables.
5. **Outlier removal.** I drop firm-years above the 99th percentile of value added per worker, and firms with year-on-year changes in $\ln(\text{VA per worker})$, $\ln(\text{assets per worker})$, or $\ln(\text{employment})$ exceeding three standard deviations, which most likely reflect consolidation events or reporting errors.

6. **Coverage floor.** For each matched firm-year I construct

$$\text{Coverage}_{ft} = \frac{\text{LinkedIn workers}_{ft}}{\text{Orbis employment}_{ft}} \quad (16)$$

13% of firm-years have coverage above one, most plausibly from timing differences between LinkedIn job records and Orbis snapshot employment; I top-code these at one. I then keep firm-years with coverage ≥ 0.1 , since shares computed from very thin coverage are too noisy to be informative. Figure 2 shows the coverage distribution; Figure 3 plots mean coverage and the match rate over time. Appendix E.4 varies the threshold.

The final panel has 20,168 firms and 130,577 firm-years. The sample firms employ around 5 million workers in an average year, roughly a sixth of UK employment.

Coverage varies across firms (Table 3). Finance and ICT (84%) and professional services (81%) are much better covered than manufacturing (44%) and trade and hospitality (47%), and London firms are better covered (72%). Thinly covered firms have noisier workforce shares, so all regressions are weighted by coverage and include a quadratic coverage control; Section 4.2 sets out the reasoning.

Firm-level human capital variables. For each firm-year I aggregate the LinkedIn records to the workforce composition variables of Section 2: $\bar{x}_{\text{deg},ft}$ is the share of workers with a degree⁵; $\bar{x}_{\text{post|deg},ft}$, $\bar{x}_{\text{stem|deg},ft}$, and $\bar{x}_{\text{elite|deg},ft}$ are the shares of degree-holders with a postgraduate qualification, a STEM degree, and a top-20 institution;⁶ $\overline{\text{exper}}_{ft}$ and $\overline{\text{exper}}_{ft}^2$ are the mean of years of firm experience and of its square; and $\bar{x}_{ft}^{\text{male}}$ is the male share. Each is constructed separately for managers and non-managers. Shares are weighted means across the LinkedIn profiles matched to the firm-year, with each profile carrying its Revelio sampling weight:

$$\bar{x}_{k,ft}^g = \frac{\sum_{i \in \mathcal{L}_{ft}^g} w_i \cdot \mathbf{1}[x_{ik} = 1]}{\sum_{i \in \mathcal{L}_{ft}^g} w_i} \quad (17)$$

where $\mathcal{L}_{ft}^g$ is the set of profiles matched to firm f in year t within group g . The weights remove the between-occupation component of LinkedIn selection: because credentials correlate with occupation, raw shares are biased upward, and more so in low-coverage firms. Within-occupation selection remains and is discussed in Section 4.2. Appendix E.4 shows that raw and reweighted shares give similar coefficients.

Table 4 reports summary statistics for the estimation sample. The degree, STEM, and

⁵Postgraduate qualifications include Masters degrees, PhDs, and MBAs. STEM degrees include Biology, Chemistry, Physics, Mathematics, Statistics, Computer Science, Engineering, Medicine, and Economics-related degrees (Accounting, Business, Finance, and Economics); hereafter I use “STEM” for this combined share. Elite institution status is a top-20 ranking in the Complete University Guide in the year of graduation; for non-UK graduates I assign the ranking of the nearest UK equivalent using the QS World University Rankings. Workers without a degree either report a school-level or associate qualification or have no education record; imputed earnings are similar across these sub-groups, and I do not distinguish them.

⁶Appendix E.1 uses a continuous log rank measure instead.

elite shares are close to those of the individual cross-section; the postgraduate share is lower (25% against 33%), most likely because universities, a major employer of postgraduates, are largely absent from Orbis. The interquartile range of log value added per worker is 0.77 log points, the familiar dispersion of UK firm productivity (Andrews et al., 2016).

4. Empirical Implementation

This section sets out the estimating equations and discusses identification. The primary objects are the full equations (8) and (15), which let the returns differ between managers and non-managers. The pooled equations (5) and (13) are estimated in the same way and are the first results reported in Section 5.

4.1 Baseline Estimating Equations

Absorbing $\ln A_{ft}$ into 4-digit NACE industry-by-year fixed effects ϕ_{jt} and firm fixed effects ϕ_f , and writing $s_{ft}^{nm} \equiv 1 - s_{ft}^m$, the regression derived from (8) is

$$\begin{aligned} \ln\left(\frac{VA}{L}\right)_{ft} &= \gamma_K \ln\left(\frac{K}{L}\right)_{ft} + \gamma_L \ln L_{ft} + \gamma_{man} s_{ft}^m \\ &+ \sum_{g \in \{nm, m\}} s_{ft}^g \left(\sum_{k \in \mathcal{K}} \gamma_k^g \bar{x}_{k,ft}^g + \gamma_{exp}^g \overline{\exp}_{ft}^g + \gamma_{exp^2}^g \overline{\exp^2}_{ft}^g \right) \\ &+ \gamma_{male} \bar{x}_{ft}^{male} + \gamma_c \text{cover}_{ft} + \gamma_{c^2} \text{cover}_{ft}^2 + \phi_{jt} + \phi_f + \varepsilon_{ft}^{pf} \end{aligned} \quad (18)$$

where $\mathcal{K} = \{\text{deg}, \text{post|deg}, \text{stem|deg}, \text{elite|deg}\}$, and $\overline{\exp}_{ft}^g$ and $\overline{\exp^2}_{ft}^g$ are the group- g means of firm experience and its square.

The wage equation derived from (15) is

$$\begin{aligned} \ln\left(\frac{W}{L}\right)_{ft} &= \delta_{man} s_{ft}^m + \sum_{g \in \{nm, m\}} s_{ft}^g \left(\sum_{k \in \mathcal{K}} \delta_k^g \bar{x}_{k,ft}^g + \delta_{exp}^g \overline{\exp}_{ft}^g + \delta_{exp^2}^g \overline{\exp^2}_{ft}^g \right) \\ &+ \delta_{male} \bar{x}_{ft}^{male} + \delta_c \text{cover}_{ft} + \delta_{c^2} \text{cover}_{ft}^2 + \phi_{jt} + \phi_f + \varepsilon_{ft}^w \end{aligned} \quad (19)$$

All production function parameters are identified from (18) in a single OLS regression. The factor elasticities are $\hat{\alpha} = \hat{\gamma}_K$ and $\hat{\beta} = \hat{\gamma}_L - \hat{\gamma}_K + 1$. The efficiency-unit premia, with delta-method standard errors, are

$$\hat{\lambda} = \frac{\hat{\gamma}_{man}}{\hat{\beta}}, \quad \hat{\mu}_k^{nm} = \frac{\hat{\gamma}_k^{nm}}{\hat{\beta}}, \quad \hat{\mu}_k^m = \frac{\hat{\gamma}_k^m}{\hat{\beta}(1 + \hat{\lambda})}$$

The wage premia are recovered from (19):

$$\hat{d}^\lambda = \hat{\delta}_{man}, \quad \hat{d}_k^{nm} = \hat{\delta}_k^{nm}, \quad \hat{d}_k^m = \frac{\hat{\delta}_k^m}{1 + \hat{\delta}_{man}}$$

and the wedges combine the two:

$$\begin{aligned} \hat{\phi}^\lambda &= \hat{d}^\lambda - \hat{\beta}\hat{\lambda} = \hat{\delta}_{man} - \hat{\gamma}_{man} \\ \hat{\phi}_k^{nm} &= \hat{d}_k^{nm} - \hat{\beta}\hat{\mu}_k^{nm} = \hat{\delta}_k^{nm} - \hat{\gamma}_k^{nm} \\ \hat{\phi}_k^m &= (1 + \hat{d}^\lambda)\hat{d}_k^m - \hat{\beta}(1 + \hat{\lambda})\hat{\mu}_k^m = \hat{\delta}_k^m - \hat{\gamma}_k^m \end{aligned}$$

So each wedge is the difference between a raw wage coefficient and the corresponding raw production coefficient. In the text I refer to $\hat{\gamma}$ as the productivity premium in log points, equal to $\hat{\beta}\hat{\mu}$, and to $\hat{\mu}$ as the premium in efficiency units.

I estimate the two equations jointly as a SUR system. The residuals are correlated across equations because common firm-level shocks move productivity and wages together, and exploiting this correlation improves inference on the wedges. Standard errors are clustered at the firm level. Full regression output is in Appendix C.

The experience quadratic and the male share enter as controls rather than as labour types of structural interest.⁷ The male share controls mainly for hours composition, since men are more likely to work full-time. Appendix E.10 treats gender as the object of interest and replicates Hellerstein et al. (1999): the male wage premium exceeds the male productivity premium by around 0.1 log points, a wedge that survives firm fixed effects, and neither the gender wedge nor the managerial wedge explains the other.

4.2 Identification

OLS estimation of (18) and (19) faces five identification problems. I discuss each and the steps taken to address it.

Endogeneity of inputs and human capital shares. Input choices are correlated with the productivity shock A_{ft} , which the firm observes and the econometrician does not (Olley and Pakes, 1996; Levinsohn and Petrin, 2003; Akerberg et al., 2015). Here the concern extends to the human capital shares: productive firms attract educated workers, biasing $\hat{\mu}_k^g$ upward in the cross-section. Rather than a control function, I rely on firm fixed effects, which absorb time-invariant productivity and the time-invariant component of sorting. They do not absorb the correlation between changes in composition and within-firm shocks. Appendix E.8 provides a check: last year's composition cannot have been caused by this year's shock, so estimates driven by contemporaneous shocks should attenuate sharply when current shares are replaced by lags. They largely do not.

⁷Entering experience in the efficiency-units aggregate in the same additive form as the education shares would require $\mu_{exper} \cdot \overline{exper}_{ft} \ll 1$, which fails at typical values of mean firm experience.

Ability bias and endogenous promotion. Two upward biases survive firm fixed effects. If high-ability workers are both more likely to hold degrees and more productive conditional on holding one, $\hat{\mu}_k^g$ absorbs ability. And if firms promote their most productive degree-holders in response to positive shocks, $\bar{x}_{k,ft}^m$ is endogenous to within-firm productivity; Appendix E.3 shows that education strongly predicts promotion within firms. The within-firm $\hat{\mu}_k^g$ are therefore conditional correlations. Instruments such as promotion vacancies created by managerial departures, or distance to universities of different types (following Card, 1995), are the natural next step.

Measurement error and selection in LinkedIn shares. The shares are estimated from a selected subsample of each workforce, which introduces two downward biases. Sampling error generates classical attenuation, which firm fixed effects exacerbate because demeaning removes signal from slowly varying shares and leaves the noise intact (Appendix B). I partially correct for this by weighting each observation by $\min(\text{Coverage}_{ft}, 1)$, which tilts the estimator toward firm-years with higher signal-to-noise. And because educated workers are more likely to have a profile, observed education shares are biased upward in low-coverage firms; since coverage and productivity are negatively related, this compresses the estimated returns. I address it with Revelio’s occupation reweighting and a quadratic coverage control. Three results confirm the direction of these biases: the human capital coefficients rise monotonically with the coverage threshold (Appendix E.4); IV estimates using lagged shares as instruments exceed OLS for nearly all coefficients (Appendix E.8); and long-difference estimates, which retain more signal than the annual within-firm estimator, lie between the fixed-effects and cross-sectional estimates (Appendix E.9).

Markup bias. Under imperfect competition, value-added productivity conflates efficiency with markups (Hall, 1988; De Loecker et al., 2020). Without output prices I cannot separate the two. Industry-year fixed effects absorb common markup trends within narrow industries, and the results are robust to region-year fixed effects, which absorb local concentration (Appendix E.5).

Taylor approximation error. The log-linearisation requires $z_{ft} \ll 1$. For the within-firm estimates, $\hat{z}_{ft} \approx 0.19$ at sample means, implying an error of order $\hat{z}_{ft}^2/2 \approx 0.02$ log points. For the cross-sectional estimates the larger premia push $\hat{z}_{ft}$ to around 0.41, implying an error of order 0.08 log points. Appendix E.7 drops the approximation with nonlinear least squares; the parameters move within this bound.

5. Results

The results come in three steps. Section 5.1 treats the workforce as a single pool and asks which types of education raise productivity, and whether wages track the answer. Section 5.2 splits the workforce by role. Section 5.3 asks whether the wage–productivity wedges differ by role.

Two points of interpretation apply throughout. The cross-sectional and within-firm

estimates answer different questions: the cross-section conflates the return to a characteristic with the sorting of workers who hold it into productive firms, and the within-firm estimates are attenuated by measurement error (Section 4.2). And every wedge is relative to the omitted category, a female non-manager with no degree and no firm experience. A positive wedge means a group’s relative wage premium exceeds its relative productivity premium against that baseline, not that the group is paid above its absolute marginal product.

All specifications include 4-digit industry-by-year fixed effects, are weighted by coverage with a quadratic coverage control, control for experience, its square, and the male share, and cluster standard errors by firm. The factor elasticities are conventional and stable: $\hat{\alpha} \approx 0.31$ – 0.34 on capital and $\hat{\beta} \approx 0.64$ – 0.67 on labour, close to constant returns, unchanged by firm fixed effects and across the robustness exercises of Appendix E. The stability reassures that the factor elasticities are not absorbing the workforce-composition effects that are the focus of the paper.

5.1 Which Types of Education Raise Productivity, and Do Wages Follow?

Table 5 reports the pooled estimates (Assumption 2): for each education characteristic, the productivity premium $\hat{\gamma} = \hat{\beta}\hat{\mu}$ from (5), the wage premium $\hat{\delta} = \hat{d}$ from (13), and the wedge $\hat{\phi} = \hat{\delta} - \hat{\gamma}$, in the cross-section and within firms. The implied efficiency-unit premia and the full regressions are in Appendix Tables 12–16.

Three facts stand out. First, in the cross-section the premia are large and institution quality rivals the degree itself. A degree raises efficiency units by 0.57 log points over no degree, and a top-20 institution adds a further 0.60; the postgraduate and STEM premia are an order of magnitude smaller. The elite premium primarily reflects sorting: graduates of selective universities work at productive firms, and Appendix D.2 shows that the premium is three to five times larger in knowledge-intensive services than in goods production.

Second, within firms the premia shrink but retain their ranking. Degree-holders are around 20% more productive than non-degree-holders in the same firm; a top-20 institution adds a further 10% and a postgraduate qualification 6%. The STEM premium is small and insignificant in the pooled specification; Section 5.2 shows that this masks a sharp role difference.⁸

Third, wages track productivity closely. In the cross-section the degree and postgraduate wage premia exceed their productivity premia by around 0.13 log points each, but these wedges primarily reflect sorting: high-wage firms employ educated workforces for reasons beyond what those workers produce. Within firms the only significant wedge is on the degree itself, at 0.06 log points. The postgraduate, STEM, and elite wedges are zero. On average, firms pay for

⁸These magnitudes can be benchmarked against the Mincerian literature, with the caveat that the comparison is between a relative productivity premium and an individual wage return. Causal estimates of the return to schooling cluster around 9–11% per year (Psacharopoulos and Patrinos, 2018; Bhuller et al., 2017), and UK studies place the undergraduate wage premium at roughly 20–40% (Blundell et al., 2005; Britton et al., 2020). The within-firm degree productivity premium of around 20% sits at the lower end of this range, as expected if the within-firm estimator is attenuated.

degree type roughly what it produces. This is the benchmark against which the role-specific wedges of Section 5.3 should be read: the wedges that appear there are hidden, not absent, in the pooled data.

5.2 Managers and Non-Managers

Figure 4 presents the headline estimates of the full specification (Assumption 3) with firm fixed effects: the manager premium $\hat{\lambda}$ and the efficiency-unit premia $\hat{\mu}_k^g$, estimated separately for managers and non-managers from a single regression. Figure 5 is the cross-sectional counterpart on the same axis. Appendix Table 12 tabulates the parameters and Table 16 the raw coefficients.

The manager premium. The manager share enters positively and significantly throughout. Within firms, a manager with no degree generates around 26% more efficiency units than an otherwise identical non-manager. Cross-sectionally, $\hat{\lambda}$ falls from 0.42 to 0.28 once education is controlled: much of the raw premium reflects educated managers at productive firms. Within firms it is essentially unmoved by conditioning on education, at 0.23 without and 0.26 with. Whatever managers contribute – coordination, decision-making, organisational knowledge – operates through channels that degree level, field, and institution do not capture.

The undergraduate degree pays in both roles. The degree premium is similar for managers and non-managers. Within firms the point estimates are $\hat{\mu}_{deg}^m = 0.11$ and $\hat{\mu}_{deg}^{nm} = 0.19$, and the difference is not statistically distinguishable from zero; in the cross-section they are 0.60 and 0.51. The role contrast that follows is about the type and level of a degree, not about holding one.

Specialist credentials pay for non-managers; breadth pays for managers. For non-managers, the elite institution and STEM premia dominate the cross-section and survive firm fixed effects: a STEM degree adds 12% in efficiency units within firms and a top-20 institution 11%. For managers, both premia are smaller in the cross-section and negligible within firms. What managers' education does predict is degree level. The postgraduate premium is large for managers in the cross-section ($\hat{\mu}_{post}^m = 0.26$, against 0.06 for non-managers), and substituting the MBA share for the postgraduate share raises the cross-sectional coefficient to 0.42 log points (Appendix E.1): the managerial postgraduate premium is substantially an MBA effect. Within firms the manager human capital estimates are individually insignificant. Manager shares interacted with education types are slow-moving and noisily measured, so firm fixed effects compound attenuation; because the two figures share an axis, the attenuation is visible as the shift of every estimate toward zero between them.

Formal tests confirm that returns differ by role. The null of equal premia, $H_0^\mu: \mu_k^m = \mu_k^{nm}$ for all k , is rejected cross-sectionally (0.1% level) and within firms (5% level). The null of equal marginal products, $H_0^{MP}: (1 + \lambda)\mu_k^m = \mu_k^{nm}$, is rejected in the cross-section (1% level): postgraduate qualifications raise productivity more when held by managers, and STEM and elite credentials less. Within firms only the STEM gap is individually significant.

These averages conceal heterogeneity, documented in Appendix D. Education premia rise steeply with firm size and are three to five times larger in knowledge-intensive services than in goods production. The manager premium is largest in the smallest firms, consistent with standardised processes substituting for individual managerial discretion in large organisations.

Two tracks and the Peter Principle. Why should the return to education depend on role? The pattern in Figures 4 and 5 is that the two roles draw on different skills. Non-managerial production rewards depth: the specialised technical knowledge that STEM training and selective institutions build. Managerial production rewards breadth: judgement across domains and the coordination of others' expertise (Garicano, 2000). STEM field and institution measure depth, so they predict productivity for workers and not for managers. Postgraduate education, and the MBA in particular, is the one measured credential that plausibly builds breadth, and it is the one whose premium is concentrated among managers. The undergraduate degree, which builds the general skills both roles use, pays in both.

The promotion process completes the story. Firms promote on performance in the current role, so they select managers on depth. This is the Peter Principle (Peter and Hull, 1969; Lazear, 2004), which Benson et al. (2019) document for salespeople: firms promote their best sellers although sales ability negatively predicts managerial quality. Appendix E.3 tests the promotion leg here. In individual regressions on the 2025 cross-section, every credential predicts promotion within firms, and STEM training – the largest within-firm non-manager productivity premium – is the strongest predictor of holding a managerial role, raising the probability by 3.8 percentage points against a base of 13%. Yet STEM carries no managerial productivity premium. The version here is milder than Benson et al.'s – specialist credentials stop paying on promotion rather than predicting bad management – but the selection mechanism is the same.

5.3 Wage–Productivity Wedges by Role

Figure 6 plots, for each education–role cell, the wage coefficient $\hat{\delta}$ against its production-function counterpart $\hat{\gamma}$. The shaded gap is the wedge $\hat{\phi} = \hat{\delta} - \hat{\gamma}$, labelled with its value and its SUR-based significance. Appendix Tables 13 and 14 tabulate the wage premia and wedges, and Table 15 the raw wage regressions.

The manager wedge is a sorting artefact. Cross-sectionally (left panel), managers' relative wage premium exceeds their relative productivity premium by around 0.2 log points. Three mechanisms from Section 1.1 predict this sign – monopsony, since managers have more outside options; bargaining, since managerial matches generate more surplus; and back-loading, since managers are later in their careers – and the cross-section cannot separate them. Within firms (right panel) the wedge disappears: -0.022 , insignificant. This points to a composition effect: high-paying firms employ proportionally more managers, so in the cross-section managers are well paid partly because of where they work, without any manager being paid above marginal product within their own firm.

The education wedges split by role. The lower panels let the wedges differ by role. For non-managers, the only significant within-firm wedge is on STEM: $\hat{\phi}_{stem}^{nm} = -0.070$ (0.1% level). STEM-trained non-managers earn a relative wage premium that falls short of their relative productivity premium within the same firm. This is the sign monopsony predicts if STEM skills are deployed at few local employers, leaving these workers with steep firm-specific labour supply and a larger markdown. It is also the sign [Dearden et al. \(2006\)](#) find for training in British industries. For managers the sign reverses: the degree wedge is 0.133, significant at the 5% level, and survives firm fixed effects, so unlike the overall manager wedge it is not a sorting artefact. Monopsony, bargaining, and dynamic pay for degree-holding managers all remain candidates. The null of equal wedges across roles, $H_0^{\phi,eq}: \phi_k^m = \phi_k^{nm}$, is rejected within firms at the 5% level.

The pooled wedges of Table 5 average these two opposite-signed gaps. The pooled STEM wedge is zero because a markdown on non-managers is offset by no markdown on managers; the pooled degree wedge of 0.06 reflects the manager degree wedge, since the non-manager degree wedge is zero.

A natural concern is that the managerial wedge is a gender wedge in disguise: managers are disproportionately male, and [Hellerstein et al. \(1999\)](#) document that women are paid below their relative marginal product. Appendix E.10 shows it is not. Managers are only six percentage points more male than non-managers, so the overlap is too small for one wedge to be a relabelling of the other, and both survive conditioning on the other.

6. Management, Human Capital, and Productivity

I link the LinkedIn–Orbis panel to the World Management Survey (WMS) to ask how much of the productivity–management relationship is accounted for by the education of managers and workers, and how much the management score captures about practices independently of who works at the firm.

6.1 The LinkedIn–Orbis–WMS Panel

The WMS conducts structured telephone interviews with plant managers, scoring firms on 18 practices across operations, monitoring, targets, and people management ([Bloom and Van Reenen, 2007](#)). I use the UK manufacturing waves, spanning 2006–2022 and covering 1,048 firms. The management measure is the average of the 18 standardised question scores, re-standardised within the matched sample, so a one-unit difference is one standard deviation of management quality.

The match proceeds in two steps. I collapse the WMS panel to the firm level, taking the mean score across waves; management practices are highly persistent ([Bloom et al., 2019](#)), and the limited overlap between WMS survey years (concentrated in 2006–2014) and the LinkedIn–Orbis panel (2015–2024) makes this the only feasible strategy. I then collapse the LinkedIn–Orbis panel to firm-level means over 2015–2022 and merge on the Bureau van Dijk

identifier. The match yields 367 firms, with 681 unmatched; the match rate reflects the sample restrictions of Section 3. Table 6 shows the matched sample is broadly representative of the WMS manufacturing sample, with most standardised mean differences near 0.1. Given the sample size, this section is a descriptive cross-sectional exercise.

As a validation of the LinkedIn shares, I compare them with the degree shares the WMS collects from the plant manager. The correlation is 0.33 for managers and 0.22 for non-managers, both significant, and 0.48 for the manager headcount share. These moderate correlations are consistent with both measures being noisy proxies: the WMS share is one manager’s estimate in a telephone interview, and the LinkedIn share is a mean across a selected subsample. The higher correlation for managers reflects their over-representation on LinkedIn.

6.2 Decomposing the Productivity–Management Relationship

The decomposition follows [Bender et al. \(2018\)](#), replacing their AKM-based ability measures with the education shares from LinkedIn. All specifications are cross-sectional with 2-digit industry fixed effects.

I first ask whether human capital predicts the management score, regressing the firm-level z -score on three progressively richer sets of controls: the manager share alone; pooled education shares; and separate manager and non-manager shares:

$$M_f = \rho^\lambda s_f^m + \sum_k \rho_k^{nm} \bar{x}_{k,f}^{nm} + \sum_k \rho_k^m \bar{x}_{k,f}^m + \phi_j + \varepsilon_f \quad (20)$$

I then regress log value added per worker on employment, the management score, and the same progressively richer controls:

$$\ln\left(\frac{VA}{L}\right)_f = \gamma_L \ln L_f + \psi M_f + \gamma_{man} s_f^m + \sum_k \gamma_k^{nm} \bar{x}_{k,f}^{nm} + \sum_k \gamma_k^m \bar{x}_{k,f}^m + \phi_j + \varepsilon_f \quad (21)$$

The attenuation of $\hat{\psi}$ across columns measures the share of the productivity–management correlation that human capital explains.

6.3 Results

Table 7 reports the management-score regressions. The manager share alone does not predict management quality. Pooled education shares do: degree share and STEM field both enter positively and significantly in column (2), so firms with more educated and technically trained workforces score higher on structured practices. Disaggregated by role in column (3), the associations are concentrated among non-managers. The postgraduate coefficient is larger for managers than non-managers, as in Section 5, though neither is individually significant. The association between human capital and management quality runs primarily through the general skill level of the workforce, not the credentials of managers.

Table 8 reports the decomposition. The baseline coefficient on the management z -score

is $\hat{\psi}_0 = 0.082$ (column 1), consistent with UK estimates in [Bloom and Van Reenen \(2007\)](#) and [Feng et al. \(2025\)](#). Let $\hat{\tau} = (\hat{\psi}_0 - \hat{\psi})/\hat{\psi}_0$ denote the explained share. Adding the manager share leaves the gradient unchanged. Pooled education shares attenuate it to 0.066 (column 3) and the full specification to 0.061 (column 4), so $\hat{\tau} \approx 26\%$.

This is close to the roughly 30% that [Bender et al. \(2018\)](#) report for Germany, and if anything conservative: education shares are a coarser proxy for ability than their AKM fixed effects, so the true human capital channel is likely larger. The pattern of attenuation differs. Here the explained share comes from overall workforce education; in [Bender et al. \(2018\)](#) managerial human capital is the dominant channel. The difference most likely reflects their proxy. They treat the highest-paid workers as managers, but highly paid specialists in non-managerial roles — common in STEM-intensive manufacturing — would be attributed to the managerial channel in their framework and to the non-managerial channel here. The coefficients in column (4) match [Section 5](#): non-manager degree and elite shares enter positively and significantly, and the manager postgraduate share is larger than its non-manager counterpart. With 367 firms the individual coefficients should be read with caution.

7. Conclusion

This paper asks three related questions. Which types of education raise firm productivity, and do wages follow? Do the returns differ between managers and non-managers, and does any misalignment between pay and productivity differ by role? And how much of the well-documented productivity–management relationship runs through the people rather than the practices? To answer them, I build a linked LinkedIn–Orbis panel of over 20,000 firms employing roughly a sixth of the UK workforce, with education observed by field, institution quality, and degree level, separately for managers and non-managers.

On the first question, the type of education matters, and institution quality matters most. Within firms, a degree raises productivity by around 20%, a top-20 institution adds a further 10%, and a postgraduate qualification 6%. The cross-sectional premia are two to five times larger, reflecting the sorting of educated workers into productive firms. On average, wages track productivity: the only significant within-firm wedge is on the degree itself, and the type-of-degree wedges are zero.

On the second question, the returns split by role in a systematic way. Specialist credentials – a STEM field, an elite institution – raise productivity for non-managers, by around 12% and 11% within firms, and not for managers, whose premium loads instead on postgraduate qualifications, above all MBAs. A manager with no degree also contributes around 26% more than an otherwise identical non-manager, a premium that barely moves when education is controlled. A potential reading is that management rewards breadth where professional and technical work rewards depth, and that firms promote on depth: a STEM degree is the strongest predictor of reaching management and carries no premium once there. This is a version of

the Peter Principle. The wedges that cancel on average also split by role, and run in opposite directions: degree-holding managers earn a relative wage premium exceeding their relative productivity premium, while STEM-trained non-managers earn one that falls short. The framework identifies the sign and size of each wedge; it does not adjudicate between monopsony, bargaining, efficiency wages, or back-loading as the source. The mechanisms reviewed in Section 1.1 make distinct predictions across worker types, and the wedges estimated here are the empirical targets against which those predictions can be tested.

On the third question, for the manufacturing firms matched to the World Management Survey, workforce educational composition accounts for around a quarter of the productivity–management correlation, close to the roughly 30% that [Bender et al. \(2018\)](#) attribute jointly to workforce selection and pay premia in German manufacturing. The figures are not strictly comparable as education shares are a coarser proxy for ability than their AKM fixed effects. The two studies differ on the channel. Here the explained share operates through the general education of the workforce rather than managerial credentials, most likely because [Bender et al.](#) proxy managers as the highest-paid workers, a set that includes highly paid specialist non-managers. This echoes the production-function result: whatever good managers contribute, their education captures only part of it.

The estimates come with clear limits. They are conditional correlations, not causal returns. Firm fixed effects absorb time-invariant sorting, and the lagged-share results in Appendix E.8 argue against simple reverse causality, but ability bias and endogenous promotion may remain. Fixed effects also trade one bias for another, worsening attenuation in slow-moving, noisily measured shares; the long-difference estimates in Appendix E.9 place the human capital premia between the within-firm and cross-sectional figures. Credible instruments – e.g., distance to universities of different types (following [Card, 1995](#)) – and adapting the [Akerberg et al. \(2015\)](#) control function to the efficiency-units framework are the natural next steps.

Read with these limits, the results still carry a clear policy message: the composition of higher education expansion matters as much as its scale. Degree attainment alone carries a sizeable productivity premium, but the non-manager premium is also large for STEM fields and elite institutions, precisely the credentials whose share has remained flat even as participation has surged ([HESA, 2024](#)). The wedges sharpen the point. If STEM-trained workers are paid below their relative marginal product, the private incentive to acquire those skills falls short of their value to firms, before counting the spillovers that push the social return higher still.

The three strands of this paper converge on a single underexplored question: what does education do for productivity? Raising productivity is a stated aim of UK education policy, yet the evidence base that informs it consists almost entirely of wage returns. If wages equalled marginal products, the distinction would be harmless. This paper shows it is not. The UK pairs a world-class university sector with a long-running productivity problem, and firm-level evidence connecting the two remains scarce. The estimates here are an imperfect first step. They map the terrain: where the productivity returns to education are concentrated, where pay

fails to track them, and how much of management's effect on productivity is explained through education. The harder tasks – identifying these returns causally, and building the equilibrium models – can now begin on surveyed ground.

References

- Abowd, John M., Francis Kramarz, and David N. Margolis**, “High Wage Workers and High Wage Firms,” *Econometrica*, 1999, 67 (2), 251–333.
- Acemoglu, Daron and Jörn-Steffen Pischke**, “Beyond Becker: Training in Imperfect Labour Markets,” *Economic Journal*, 1999, 109 (453), F112–F142.
- Ackerberg, Daniel A., Kevin Caves, and Garth Frazer**, “Identification Properties of Recent Production Function Estimators,” *Econometrica*, 2015, 83 (6), 2411–2451.
- Altonji, Joseph G., Erica Blom, and Costas Meghir**, “Heterogeneity in Human Capital Investments: High School Curriculum, College Major, and Careers,” *Annual Review of Economics*, 2012, 4, 185–223.
- , **Peter Arcidiacono, and Arnaud Maurel**, “The Analysis of Field of Study and Career Outcomes,” in Eric A. Hanushek, Stephen Machin, and Ludger Woessmann, eds., *Handbook of the Economics of Education*, Vol. 5, Elsevier, 2016, pp. 305–396.
- Andrews, Dan, Chiara Criscuolo, and Peter N. Gal**, “The Best versus the Rest: The Global Productivity Slowdown, Divergence across Firms and the Role of Public Policy,” OECD Productivity Working Papers 5, OECD, Paris 2016.
- Bandiera, Oriana, Andrea Prat, Stephen Hansen, and Raffaella Sadun**, “CEO Behavior and Firm Performance,” *Journal of Political Economy*, 2020, 128 (4), 1325–1369.
- Becker, Gary S.**, *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education*, National Bureau of Economic Research, 1964.
- Bender, Stefan, Nicholas Bloom, David Card, John Van Reenen, and Stefanie Wolter**, “Management Practices, Workforce Selection, and Productivity,” *Journal of Labor Economics*, 2018, 36 (S1), S371–S409.
- Benson, Alan, Danielle Li, and Kelly Shue**, “Promotions and the Peter Principle,” *The Quarterly Journal of Economics*, 2019, 134 (4), 2085–2134.
- Benzell, Seth G., Guillermo Lagarda, and Daniel Rock**, “Do Labor Demand Shifts Occur Within Firms or Across Them? Non-Routine-Biased Technological Change, 2000–2016,” 2019. Working Paper, SSRN 3427396.
- Bertrand, Marianne and Antoinette Schoar**, “Managing with Style: The Effect of Managers on Firm Policies,” *Quarterly Journal of Economics*, 2003, 118 (4), 1169–1208.
- Bhuller, Manudeep, Magne Mogstad, and Kjell G. Salvanes**, “Life-Cycle Earnings, Education Premiums, and Internal Rates of Return,” *Journal of Labor Economics*, 2017, 35 (4), 993–1030.
- Bloom, Nicholas and John Van Reenen**, “Measuring and Explaining Management Practices Across Firms and Countries,” *Quarterly Journal of Economics*, 2007, 122 (4), 1351–1408.
- , **Erik Brynjolfsson, Lucia Foster, Ron Jarmin, Megha Patnaik, Itay Saporta-Eksten, and John Van Reenen**, “What Drives Differences in Management Practices?,” *American Economic Review*, 2019, 109 (5), 1648–1683.

- , **Mark Schankerman, and John Van Reenen**, “Identifying Technology Spillovers and Product Market Rivalry,” *Econometrica*, 2013, 81 (4), 1347–1393.
- , **Raffaella Sadun, and John Van Reenen**, “Management as a Technology?,” Working Paper 22327, National Bureau of Economic Research 2016.
- Blundell, Richard, Lorraine Dearden, and Barbara Sianesi**, “Evaluating the Effect of Education on Earnings: Models, Methods and Results from the National Child Development Survey,” *Journal of the Royal Statistical Society: Series A*, 2005, 168 (3), 473–512.
- Britton, Jack, Lorraine Dearden, Laura van der Erve, and Ben Waltmann**, “The Impact of Undergraduate Degrees on Lifetime Earnings,” IFS Report R167, Institute for Fiscal Studies 2020.
- , — , **Neil Shephard, and Anna Vignoles**, “How English Domiciled Graduate Earnings Vary with Gender, Institution Attended, Subject and Socio-Economic Background,” IFS Working Paper W16/06, Institute for Fiscal Studies 2016.
- Burdett, Kenneth and Dale T. Mortensen**, “Wage Differentials, Employer Size, and Unemployment,” *International Economic Review*, 1998, 39 (2), 257–273.
- Cahuc, Pierre, Fabien Postel-Vinay, and Jean-Marc Robin**, “Wage Bargaining with On-the-Job Search: Theory and Evidence,” *Econometrica*, 2006, 74 (2), 323–364.
- Card, David**, “Using Geographic Variation in College Proximity to Estimate the Return to Schooling,” in Louis N. Christofides, E. Kenneth Grant, and Robert Swidinsky, eds., *Aspects of Labour Market Behaviour: Essays in Honour of John Vanderkamp*, Toronto: University of Toronto Press, 1995, pp. 201–222.
- , “The Causal Effect of Education on Earnings,” in Orley Ashenfelter and David Card, eds., *Handbook of Labor Economics*, Vol. 3A, Elsevier, 1999, pp. 1801–1863.
- , **Ana Rute Cardoso, and Patrick Kline**, “Bargaining, Sorting, and the Gender Wage Gap: Quantifying the Impact of Firms on the Relative Pay of Women,” *Quarterly Journal of Economics*, 2016, 131 (2), 633–686.
- Dearden, Lorraine, Howard Reed, and John Van Reenen**, “The Impact of Training on Productivity and Wages: Evidence from British Panel Data,” *Oxford Bulletin of Economics and Statistics*, 2006, 68 (4), 397–421.
- Dorn, David, Florian Schoner, Moritz Seebacher, Lisa Simon, and Ludger Woessmann**, “Multidimensional Skills on LinkedIn Profiles: Measuring Human Capital and the Gender Skill Gap,” 2025. CESifo Working Paper No. 11846.
- Exadaktylos, Dimitrios, Massimo Riccaboni, and Armando Rungi**, “Talents from Abroad: Foreign Managers and Productivity in the United Kingdom,” *International Economics*, 2024, 177, 100474.
- Feng, Yuchen, Andrew Henley, and Anna Kochanova**, “The Role of Management Practices in Productivity: Does Family Ownership Matter?,” Working Paper 051, The Productivity Institute 2025.

- Galindo-Rueda, Fernando and Jonathan Haskel**, “Skills, Workforce Characteristics and Firm-Level Productivity: Evidence from the Matched ABI/Employer Skills Survey,” 2005. IZA Discussion Paper No. 1542.
- Garicano, Luis**, “Hierarchies and the Organization of Knowledge in Production,” *Journal of Political Economy*, 2000, 108 (5), 874–904.
- Hægeland, Torbjørn and Tor Jakob Klette**, “Do Higher Wages Reflect Higher Productivity? Education, Gender and Experience Premiums in a Matched Plant-Worker Data Set,” in John C. Haltiwanger, Julia I. Lane, James R. Spletzer, Jules J. M. Theeuwes, and Kenneth R. Troske, eds., *The Creation and Analysis of Employer-Employee Matched Data*, Vol. 241 of *Contributions to Economic Analysis*, Amsterdam: Elsevier, 1999, pp. 231–259.
- Hall, Robert E.**, “The Relation between Price and Marginal Cost in U.S. Industry,” *Journal of Political Economy*, 1988, 96 (5), 921–947.
- Hellerstein, Judith K., David Neumark, and Kenneth R. Troske**, “Wages, Productivity, and Worker Characteristics,” *Journal of Labor Economics*, 1999, 17 (3), 409–446.
- Hershbein, Brad and Melissa S. Kearney**, “Major Decisions: What Graduates Earn Over Their Lifetimes,” Technical Report, The Hamilton Project, Brookings Institution 2014.
- HESA**, “Higher Education Student Statistics: UK, 2022–23,” Technical Report, HESA 2024.
- Jeffers, Jessica**, “The Impact of Restricting Labor Mobility on Corporate Investment and Entrepreneurship,” *Review of Financial Studies*, 2024, 37 (1), 1–44.
- Kalemli-Ozcan, Sebnem, Bent Sørensen, Carolina Villegas-Sanchez, Vadym Volosovych, and Sevcan Yesiltas**, “How to Construct Nationally Representative Firm Level Data from the Orbis Global Database: New Facts and Aggregate Implications,” Working Paper 21558, National Bureau of Economic Research 2015.
- Kaplan, Steven N., Mark M. Klebanov, and Morten Sorensen**, “Which CEO Characteristics and Abilities Matter?,” *Journal of Finance*, 2012, 67 (3), 973–1007.
- Kirkeboen, Lars J., Edwin Leuven, and Magne Mogstad**, “Field of Study, Earnings, and Self-Selection,” *Quarterly Journal of Economics*, 2016, 131 (3), 1057–1111.
- Lazear, Edward P.**, “Why Is There Mandatory Retirement?,” *Journal of Political Economy*, 1979, 87 (6), 1261–1284.
- , “The Peter Principle: A Theory of Decline,” *Journal of Political Economy*, 2004, 112 (S1), S141–S163.
- Lebedinski, Lara and Vincent Vandenberghe**, “Assessing Education’s Contribution to Productivity Using Firm-Level Evidence,” *International Journal of Manpower*, 2014, 35 (8), 1116–1139.
- Levinsohn, James and Amil Petrin**, “Estimating Production Functions Using Inputs to Control for Unobservables,” *The Review of Economic Studies*, 2003, 70 (2), 317–341.
- Loecker, Jan De, Jan Eeckhout, and Gabriel Unger**, “The Rise of Market Power and the Macroeconomic Implications,” *The Quarterly Journal of Economics*, 2020, 135 (2), 561–644.

- Lucas, Robert E.**, “On the Size Distribution of Business Firms,” *Bell Journal of Economics*, 1978, 9 (2), 508–523.
- Manning, Alan**, *Monopsony in Motion: Imperfect Competition in Labor Markets*, Princeton, NJ: Princeton University Press, 2003.
- Mincer, Jacob**, *Schooling, Experience, and Earnings*, National Bureau of Economic Research, 1974.
- Moretti, Enrico**, “Workers’ Education, Spillovers, and Productivity: Evidence from Plant-Level Production Functions,” *American Economic Review*, 2004, 94 (3), 656–690.
- Olley, G. Steven and Ariel Pakes**, “The Dynamics of Productivity in the Telecommunications Equipment Industry,” *Econometrica*, 1996, 64 (6), 1263–1297.
- Peter, Laurence J. and Raymond Hull**, *The Peter Principle: Why Things Always Go Wrong*, New York: William Morrow, 1969.
- Psacharopoulos, George and Harry Anthony Patrinos**, “Returns to Investment in Education: A Decennial Review of the Global Literature,” *Education Economics*, 2018, 26 (5), 445–458.
- Shapiro, Carl and Joseph E. Stiglitz**, “Equilibrium Unemployment as a Worker Discipline Device,” *American Economic Review*, 1984, 74 (3), 433–444.
- Tambe, Prasanna, Lorin M. Hitt, Daniel Rock, and Erik Brynjolfsson**, “Digital Capital and Superstar Firms,” Working Paper 28285, National Bureau of Economic Research 2020.

Figures and Tables

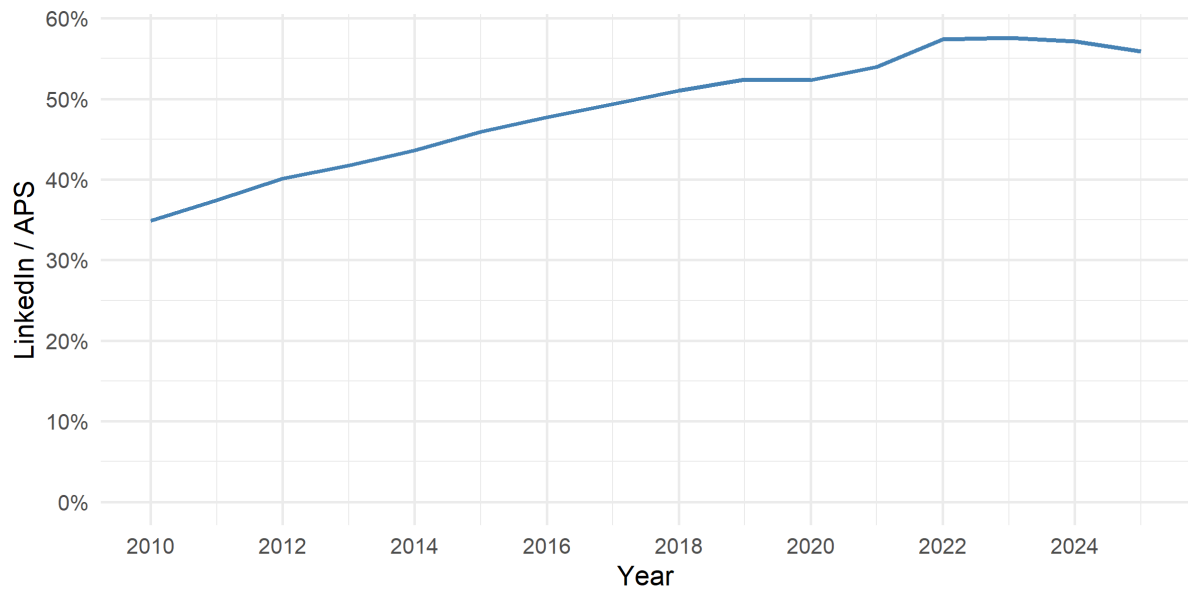


Figure 1: LinkedIn Employment Coverage Relative to APS, UK, 2010–2025

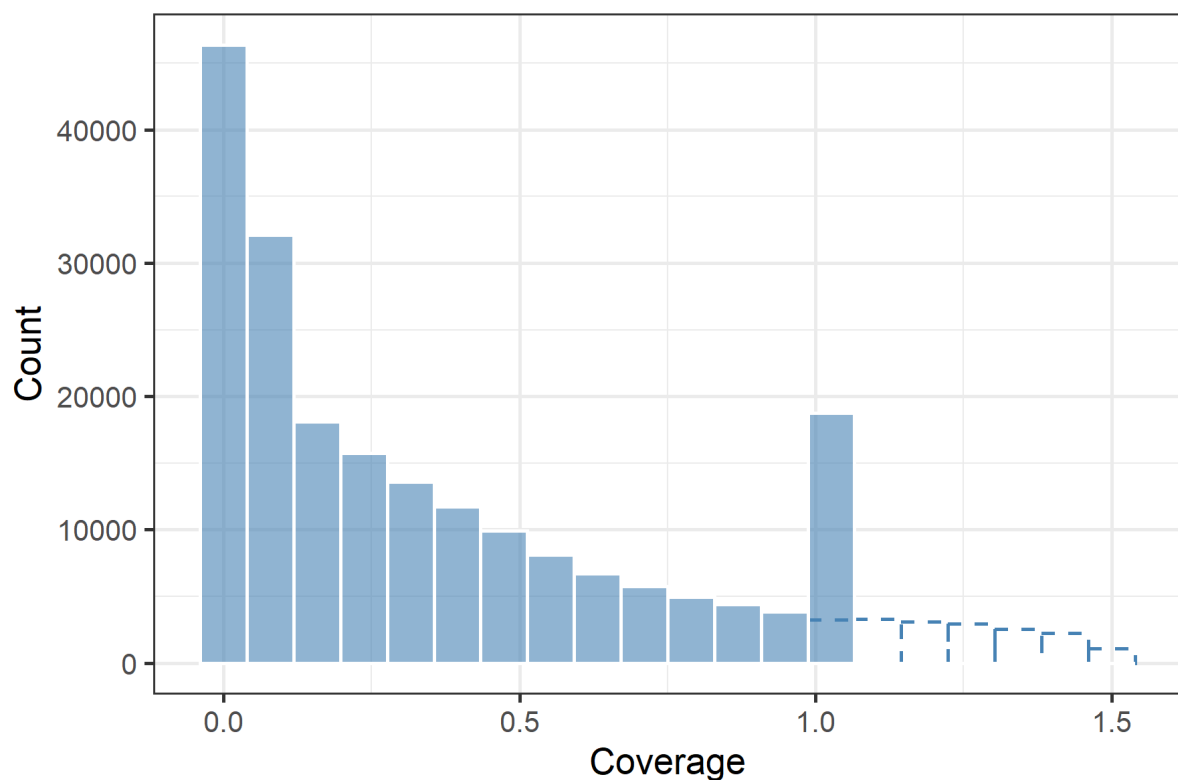


Figure 2: Employment Coverage Density, LinkedIn–Orbis Panel, UK, 2015–2024

Notes: Coverage = LinkedIn workers / Orbis employment. The solid bars top-code coverage at one, which is the transformation applied throughout the analysis.

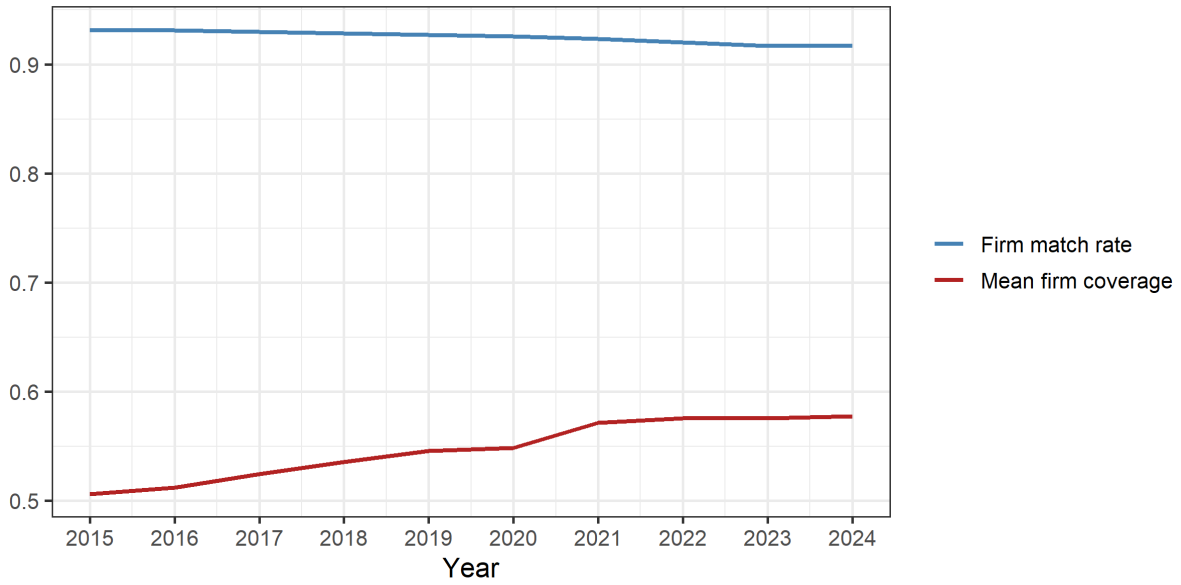


Figure 3: Firm Match Rate and Employment Coverage Rate, LinkedIn–Orbis Regression Panel, UK, 2015–2024

Notes: Computed on the final matched regression sample (coverage ≥ 0.1), top-coding coverage at one. Mean coverage over the full panel is 0.55; median is 0.48.

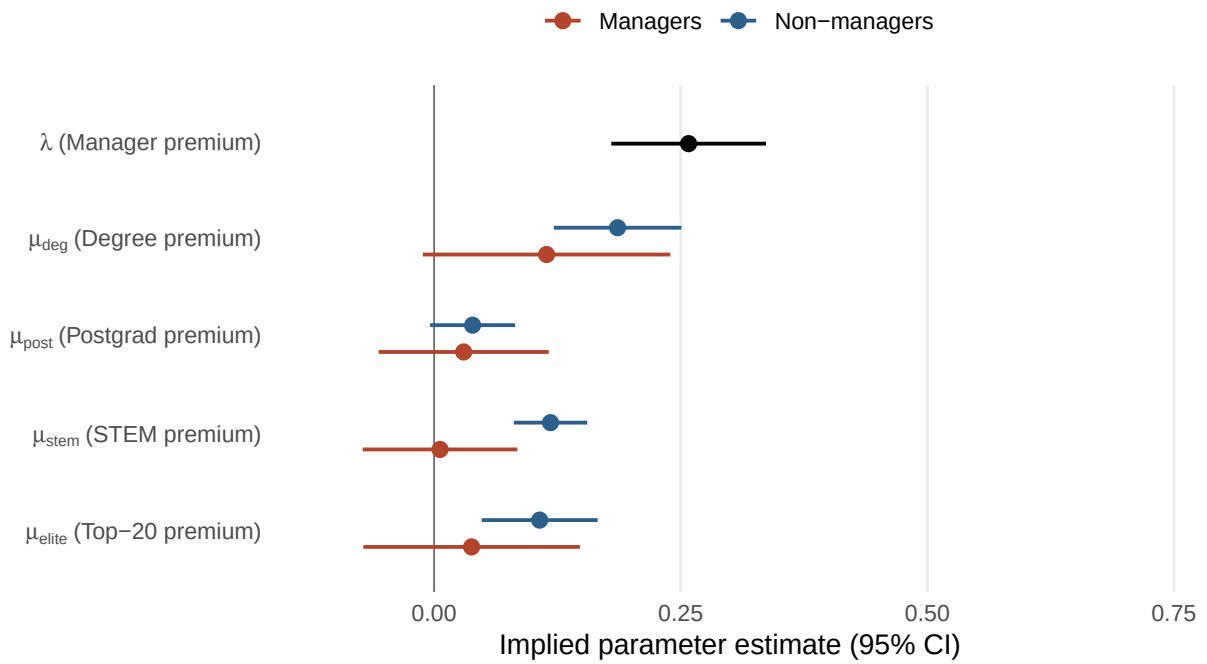


Figure 4: Implied Production Function Parameters, Within Firm

Notes: Implied structural parameters with 95% confidence intervals from the full specification (8). Standard errors clustered at the firm level. Full table in Appendix Table 12.

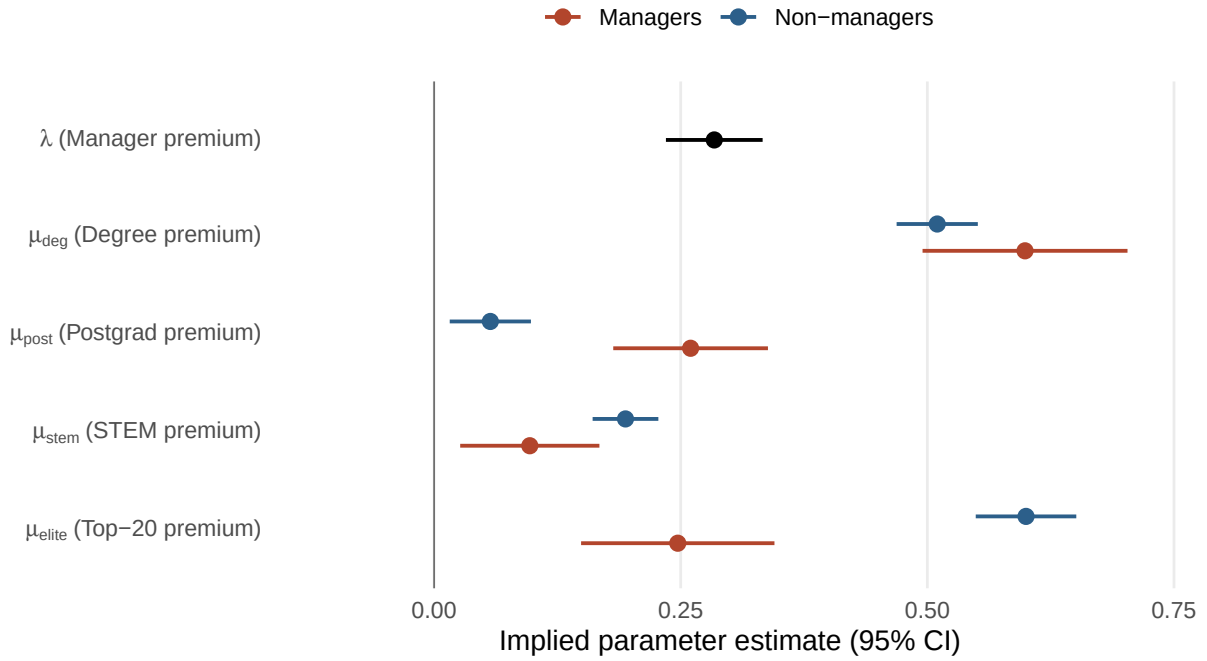


Figure 5: Implied Production Function Parameters, Cross-Section

Notes: Implied structural parameters with 95% confidence intervals from the full specification (8). Standard errors clustered at the firm level. Full table in Appendix Table 12.

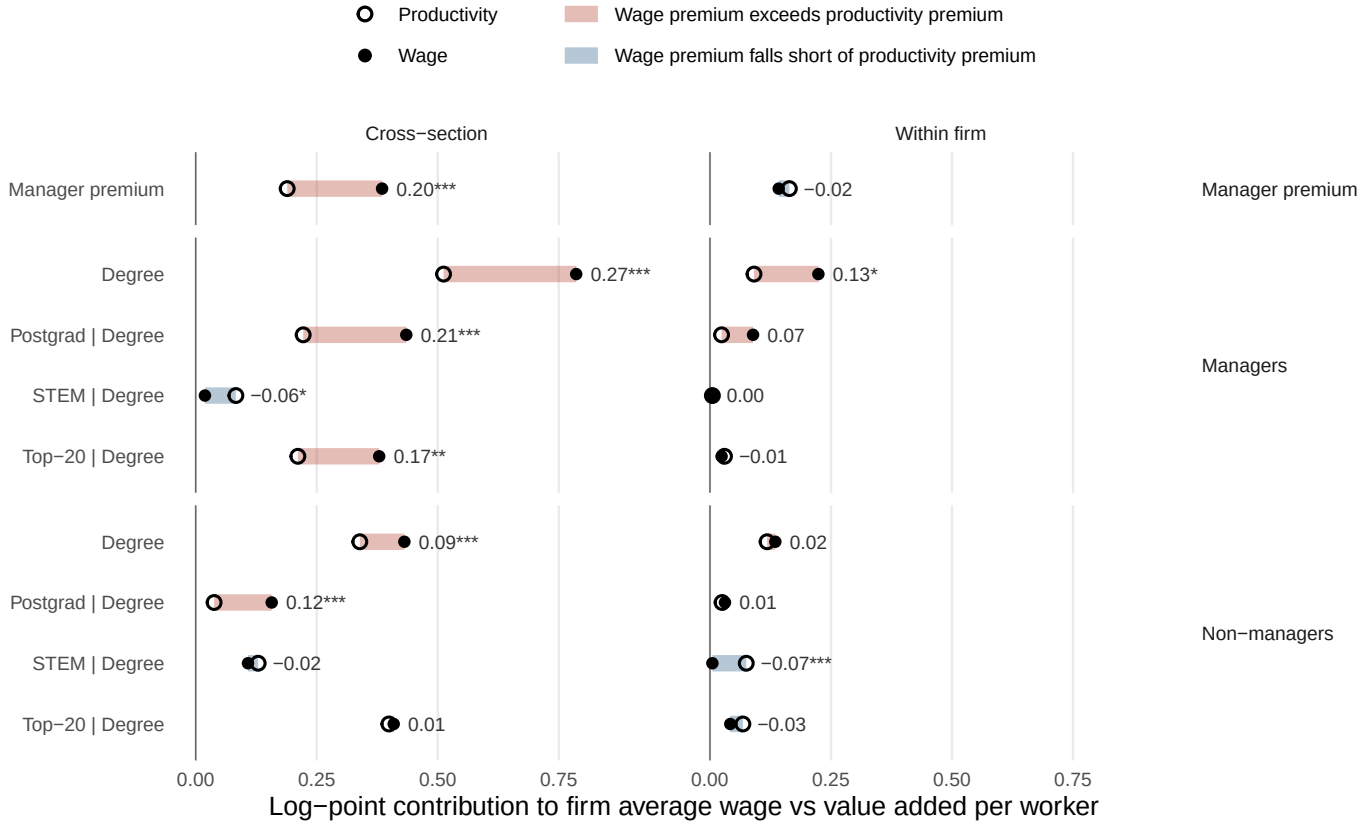


Figure 6: Wage Premia versus Productivity Premia, by Education–Role Cell

Notes: Each row plots the wage-equation coefficient $\hat{\delta}$ (filled) against the production-function coefficient $\hat{\gamma}$ (hollow) for the same cell, from columns (5)–(6) of Appendix Tables 15 and 16. The shaded gap is the wage–productivity wedge $\hat{\phi} = \hat{\delta} - \hat{\gamma}$, labelled with its value and significance from the SUR standard errors of Appendix Table 14. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table 1: Worker Group Shares in 2025

Worker Group	APS Share	LinkedIn Share
Managers	11%	25%
Professional	23%	26%
STEM & Healthcare	17%	21%
Sales & Office	9%	14%
All Other Service, Manual & Industrial	39%	14%
Total	100%	100%

Notes: Professional (SOC2 = 13 Business & Financial Operations, 23 Legal, 25 Educational Instruction & Library, 27 Arts, Design, Entertainment, Sports & Media); STEM & Healthcare (SOC2 = 15 Computer & Information Technology, 17 Architecture & Engineering, 19 Life, Physical & Social Science, 29 Healthcare Practitioners & Technical); Sales & Office (SOC2 = 41 Sales & Related, 43 Office & Administrative Support); Managers (SOC2 = 11 Management).

Table 2: Sample Construction

Step	<i>N</i> Firms	<i>N</i> Firm-Years	Emp	VA/Emp (£)	Staff/Emp (£)	London
(1) Employment ≥ 50	99,559	543,903				
(2) Child firms only	50,485	414,985				
(3) Non-missing key vars	31,694	230,288	378	102,443	65,873	20.6%
(4) Matched to LinkedIn	28,921	212,919	385	89,058	53,441	19.9%
(5) Extreme values	25,794	187,416	386	65,653	44,771	18.2%
(6) Coverage ≥ 0.1	20,168	130,577	290	66,536	45,457	17.4%

Notes: Each row reports the sample after applying the stated restriction cumulatively. Employment and financial variables are reported from step (3) onwards, when the Orbis financial data are fully available. London share is the fraction of firms with a registered address in Greater London.

Table 3: Average Coverage Rate by Firm Size, Industry, and Region

Group	Mean Coverage	Median Coverage	N Firm-Years
By Firm Size Quartile			
Q1 (Smallest)	0.60	0.57	32,622
Q2	0.55	0.47	32,633
Q3	0.53	0.44	32,639
Q4 (Largest)	0.53	0.45	32,683
By Industry (NACE Section)			
C — Manufacturing	0.44	0.37	29,169
D–F — Utilities & Construction	0.52	0.46	11,420
G–I — Trade, Transport & Hospitality	0.47	0.39	28,623
J–K — Finance & ICT	0.84	0.76	11,923
L–M — Professional Services	0.81	0.77	10,763
N — Administrative & Support	0.78	0.68	9,298
By NUTS1 Region			
North East (England)	0.45	0.37	3,511
North West (England)	0.51	0.43	13,194
Yorkshire and the Humber	0.49	0.40	9,634
East Midlands (England)	0.49	0.40	9,627
West Midlands (England)	0.48	0.39	12,137
East of England	0.53	0.45	13,427
London	0.72	0.59	24,161
South East (England)	0.58	0.50	21,071
South West (England)	0.48	0.39	8,769
Wales	0.46	0.36	3,272
Scotland	0.49	0.40	7,234
Northern Ireland	0.41	0.31	3,052

Notes: Computed on the final matched regression sample (coverage ≥ 0.1), top-coding coverage at one. Mean coverage over the full panel is 0.55; median is 0.48.

Table 4: Workforce and Manager Characteristics in the Estimation Panel

Panel A: Unweighted Workforce Human Capital Characteristics

Variable	All Workers		Managers	
	Mean	SD	Mean	SD
% Any Degree	51%	18%	56%	23%
% Postgrad Degree	25%	17%	29%	22%
% STEM Degree	48%	21%	55%	24%
% Elite Institution Degree	11%	9%	13%	11%
Mean firm experience (years)	5.69	2.17	7.41	4.37
Mean job experience (years)	4.81	1.97	6.19	4.17
Mean total experience (years)	15.30	3.58	18.56	6.28

Panel B: Firm-Level Financial Characteristics

Variable	Mean	SD	P25	Median	P75
ln(VA per worker)	10.88	0.66	10.48	10.84	11.25
ln(Staff cost per worker)	10.58	0.53	10.29	10.56	10.85
ln(Capital per worker)	11.65	1.06	11.06	11.64	12.23
ln(Employment)	4.92	0.93	4.28	4.74	5.36

Notes: $N = 130,577$. Unit of observation is the firm-year. Variables are means across observed LinkedIn profiles matched to the firm in that year. Education shares conditional on degree-holding where indicated. Managers = {Middle Managers, Executive Managers}; Frontline Managers included in ‘All Workers’.

Table 5: Productivity Premia, Wage Premia, and Wedges by Education Type, Pooled Workforce

	Cross-section			Within firm		
	Productivity $\hat{\gamma}$ (1)	Wage $\hat{\delta}$ (2)	Wedge $\hat{\phi}$ (3)	Productivity $\hat{\gamma}$ (4)	Wage $\hat{\delta}$ (5)	Wedge $\hat{\phi}$ (6)
Degree share	0.379*** (0.011)	0.513*** (0.009)	0.134*** (0.014)	0.129*** (0.019)	0.189*** (0.013)	0.060** (0.023)
Postgrad Deg.	0.081*** (0.012)	0.217*** (0.010)	0.136*** (0.016)	0.037* (0.015)	0.049*** (0.010)	0.012 (0.018)
STEM Deg.	0.075*** (0.010)	0.127*** (0.008)	0.052*** (0.013)	0.013 (0.013)	0.012 (0.009)	-0.001 (0.016)
Elite Deg.	0.396*** (0.015)	0.431*** (0.012)	0.035 (0.019)	0.065*** (0.019)	0.044*** (0.013)	-0.021 (0.023)
$\hat{\beta}$ (labour elasticity)	0.661			0.637		
Observations	111,960	111,960	111,960	111,960	111,960	111,960
Firm FE	No	No	No	Yes	Yes	Yes

Notes: Pooled specification (Assumption 2). Columns (1) and (4) report the productivity premium $\hat{\gamma}_k = \hat{\beta}\hat{\mu}_k$ in log points from (5); the efficiency-unit premium $\hat{\mu}_k = \hat{\gamma}_k/\hat{\beta}$ is reported in Appendix Table 12. Columns (2) and (5) report the wage premium $\hat{\delta}_k = \hat{d}_k$ from (13). Columns (3) and (6) report the wedge $\hat{\phi}_k = \hat{\delta}_k - \hat{\gamma}_k$, with standard errors from feasible SUR. Coefficients are from columns (3)–(4) of Appendix Tables 16, 15, and 14. All specifications include industry×year FE, a quadratic in experience, the male share, and a quadratic coverage control, and are weighted by coverage. Standard errors clustered at the firm level. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

Table 6: Covariate Balance: WMS Unmatched vs. LinkedIn–Orbis Matched Sample

Variable	Mean		Diff.	SMD	<i>p</i> -value
	Unmatched	Matched			
<i>Management z-scores</i>					
Management (composite)	−0.035	0.074	0.110	0.107	0.084
Operations	−0.057	0.120	0.177	0.174	0.005
Monitoring	−0.047	0.098	0.145	0.141	0.022
Targets	−0.013	0.027	0.040	0.040	0.525
People	−0.016	0.034	0.050	0.048	0.434
<i>Firm characteristics</i>					
Employment	960.3	506.4	−453.8	−0.140	0.008
Firm age (years)	49.5	47.0	−2.5	−0.055	0.360
Sales (£000s)	198,749	110,647	−88,102	−0.129	0.025
Multinational (%)	57.1	52.8	−4.2	−0.088	0.164
London (%)	3.4	1.6	−1.8	−0.097	0.096
<i>Workforce composition</i>					
Manager degree share (%)	42.5	43.5	1.0	0.034	0.613
Non-manager degree share (%)	9.3	9.2	−0.1	−0.004	0.949
Manager share (%)	12.4	10.7	−1.7	−0.189	0.002
<i>N</i> (firms)	681	367			

Notes: $SMD = (\text{Mean}_{\text{Matched}} - \text{Mean}_{\text{Unmatched}}) / SD_{\text{Unmatched}}$. $|SMD| > 0.1$ conventionally indicates meaningful imbalance. *p*-values from two-sample *t*-tests.

Table 7: Human Capital as a Predictor of Management Quality

	(1)	(2)	(3)	
		Pooled	Managers	Non-managers
<i>Dependent variable: WMS management z-score</i>				
Manager share (%)	−0.009 (0.008)			−0.011 (0.009)
Degree share		1.032** (0.343)	0.201 (0.361)	0.890* (0.443)
Postgrad Degree		0.529 (0.349)	0.349 (0.373)	0.188 (0.297)
STEM Degree		0.945** (0.287)	0.117 (0.288)	0.678* (0.312)
Elite Degree		0.701 (0.427)	0.928* (0.452)	0.246 (0.392)
<i>N</i>	356	347		318
<i>R</i> ²	0.168	0.206		0.205

Notes: OLS with 2-digit industry fixed effects absorbed. Column (3) reports manager and non-manager coefficients from a single regression. Manager share is the WMS-reported percentage of workers in managerial roles. Standard errors (robust) in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table 8: Decomposing the Productivity–Management Correlation

	(1)	(2)	(3)	(4)	
				Managers	Non-managers
<i>Dependent variable: ln(VA per worker)</i>					
ln(Employment)	0.006 (0.033)	0.012 (0.033)	0.003 (0.033)		0.004 (0.035)
Management z -score	0.082** (0.027)	0.083** (0.027)	0.066* (0.028)		0.061* (0.030)
Manager share (%)		0.002 (0.004)			0.003 (0.004)
Degree share			0.480** (0.163)	0.377* (0.166)	0.537** (0.205)
Postgrad Degree			0.131 (0.164)	0.259 (0.136)	0.084 (0.171)
STEM Degree			0.213 (0.136)	0.001 (0.132)	0.156 (0.144)
Elite Degree			0.196 (0.200)	0.117 (0.209)	0.356* (0.180)
N	367	356	347		318
R^2	0.355	0.359	0.388		0.432
$\hat{\psi}$	0.082	0.083	0.066		0.061
$\hat{\tau}$ (%)	—	—	19.1		25.9

Notes: OLS with 2-digit industry fixed effects absorbed. Column (4) reports manager and non-manager coefficients from a single regression. Manager share is the WMS-reported percentage of workers in managerial roles. Standard errors (robust) in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Appendix

A. Data Appendix

Covariate balance: LinkedIn–Orbis match. Table 9 compares the matched sample to the Orbis universe of child firms with ≥ 50 employees and non-missing financials (step 3 of Table 2). All standardised mean differences are well below 0.1.

Table 9: Covariate Balance: Orbis Universe vs. LinkedIn-Matched Sample

Variable	Mean (All)	Mean (Matched)	Difference	SMD
Employment	378	385	7	0.003
ln(Employment)	5.024	5.032	0.008	0.008
VA per worker (£)	102,443	89,058	−13,384	−0.002
ln(VA per worker)	10.854	10.857	0.003	0.004
Staff cost per worker (£)	65,873	53,441	−12,433	−0.002
ln(Staff cost per worker)	10.556	10.558	0.003	0.004
Share in London	0.206	0.199	−0.007	−0.017
<i>N</i> (firm-year)	230,288	212,919		

Notes: Difference is Mean (Matched) – Mean (All). Standardised mean difference (SMD) = $(\bar{X}_{\text{Matched}} - \bar{X}_{\text{All}})/SD_{\text{All}}$. $|SMD| > 0.1$ conventionally indicates meaningful imbalance.

Managers and non-managers in the 2025 LinkedIn cross-section. Table 10 reports descriptive statistics by occupation group. Managers are more male (60%) and more likely to be White (77%) than the average worker, and have the highest median connection count. They hold degrees at a similar rate to the average worker (58% against 54%) but at a lower rate than professionals (64%) or STEM and health workers (67%); conditional on a degree, they hold MBAs at three times the non-manager rate. Managers earn substantially more, partly by construction, since Revelio imputes earnings from title, seniority, experience, location, and occupation; imputed earnings are not used in the production function or wage equations. Additional compensation — bonuses, overtime, and incentive pay — is particularly large for managers, which is relevant to the wage–productivity comparison if managerial pay is back-loaded or performance-contingent. Managers also switch jobs more often and are promoted at substantially higher rates. Job switches are new LinkedIn positions; promotions are increases in the Revelio seniority score; industry is 2-digit NAICS.

Table 11 reports regression-adjusted differences, conditioning on cohort (year of labour market entry), 2-digit industry, and travel-to-work area fixed effects. With over 18.5 million observations every difference is significant at the 0.1% level, so the *D*-statistic — the raw gap relative to its pooled standard deviation — is the more informative metric. The demographic *D*-statistics are small (0.08 for gender, 0.11 for ethnicity). The earnings *D*-statistics are large (1.10 for salary): a manager drawn at random sits at around the 86th percentile of the non-

manager salary distribution. Among the education variables, the MBA gap stands out: despite a gap of only 6.4 percentage points, its *D*-statistic of 0.30 is the highest, because MBA holders are rare across the whole workforce. Firm and job tenure flip sign after adjustment, since the raw tenure advantage is an age-composition effect.

Table 10: Sample Descriptives by Occupation Group, UK 2025 LinkedIn Cross-Section

Panel A: Demographics and Connections

Occupation Group	% Male	% White	Med. Conn.
Managers	60%	77%	179
Professional	50%	75%	126
STEM & Health	66%	72%	134
Sales & Office	49%	75%	60
Other*	58%	77%	57

Panel B: Education

Occupation Group	% Any Deg.	% Postgrad Deg.	% MBA Deg.	% STEM Deg.	% Top-20 Deg.
Managers	58%	38%	9%	56%	19%
Professional	64%	34%	3%	37%	18%
STEM & Health	67%	44%	3%	59%	19%
Sales & Office	43%	24%	4%	47%	11%
Other*	42%	29%	2%	43%	12%

Panel C: Tenure and Earnings

Occupation Group	Mean Tenure (Years)			Salary Earnings		Total Compensation	
	Job	Firm	Total	Mean	Median	Mean	Median
Managers	7.39	8.74	18.38	£95,900	£80,400	£164,400	£108,600
Professional	6.48	7.40	16.05	£62,200	£51,900	£82,500	£60,500
STEM & Health	5.82	6.84	14.50	£63,400	£59,000	£79,200	£67,400
Sales & Office	6.79	7.54	14.31	£50,200	£42,100	£65,900	£50,500
Other*	5.54	8.42	15.65	£54,100	£44,200	£72,100	£48,000

Panel D: Job Mobility

Occupation Group	Mean Job Switches	Of Which			
		Promotions	In Local Firm	In Parent Firm	In Industry
Managers	3.32	40%	21%	23%	32%
Professional	3.24	32%	15%	19%	28%
STEM & Health	3.01	28%	19%	25%	38%
Sales & Office	2.39	21%	15%	18%	24%
Other*	2.11	22%	13%	14%	22%

Notes: Managers = executive + middle managers; *Other = all other service, manual, and industrial occupations + frontline managers. Occupation groups otherwise defined by SOC2 codes as in Table 1. Earnings are Revelio imputations.

Table 11: Comparison of Managers and Non-Managers, UK 2025 LinkedIn Cross-Section

	Non-Managers	Managers	Raw Gap		Adjusted Gap	
Outcome			Gap	D-stat	Gap	Std. Coef.
Demographics						
% Male	55%	60%	5.3%	0.08	0.4%	0.01
% White	74%	77%	2.9%	0.11	2.2%	0.07
% Asian	10%	8.7%	−1.7%	−0.07	−1.2%	−0.05
Education						
% Any Degree	54%	58%	3.6%	0.04	6.2%	0.13
% Postgrad Deg.	33%	38%	4.7%	0.06	4.0%	0.10
% MBA Deg.	2.6%	9.0%	6.4%	0.30	4.9%	0.28
% STEM Deg.	46%	56%	9.9%	0.15	8.8%	0.19
% Top-20 Deg.	14%	19%	5.3%	0.08	2.4%	0.07
Network						
Connections	153	222	69	0.42	78	0.48
Career Mobility						
Job Switches	2.7	3.3	0.6	0.16	0.5	0.16
% Promotions	24%	40%	16%	0.39	12%	0.41
% Within Firm	15%	21%	6.4%	0.24	7.8%	0.26
% Within Parent Firm	19%	23%	4.1%	0.24	8.2%	0.27
% Within Industry	28%	32%	4.0%	0.20	8.0%	0.22
Tenure						
Firm Tenure (years)	7.6	8.7	1.1	0.16	−0.6	−0.11
Job Tenure (years)	6.1	7.4	1.3	0.11	−1.0	−0.19
Total Experience (years)	15.2	18.4	3.2	0.27	0.3	0.04
Earnings						
Salary Earnings	£56,800	£95,900	£39,100	1.10	£30,900	0.98
Total Compensation	£75,650	£164,400	£88,750	1.00	£68,100	0.98
N	16,518,285	2,019,878				

Notes: Each row reports means by manager status, the raw gap (Managers – Non-Managers), and an adjusted gap. Adjusted gap estimated via OLS with cohort fixed effects (year of labour market entry), NAICS 2-digit industry fixed effects, and travel-to-work-area fixed effects. *D*-stat = raw gap / pooled SD. Standardised coefficient = adjusted gap / residual SD of outcome. Managers = executive + middle managers; non-managers = all other workers, including frontline managers.

B. Measurement Error: Formal Derivation

This appendix derives the attenuation bias from sampling error in the LinkedIn shares and shows how firm fixed effects compound it.

Abstracting from selection, model the observed share $\hat{x}_{k,ft}^g$ for a firm-year with n_{ft}^L LinkedIn observations as a sample mean of Bernoulli draws: an unbiased estimator of the true share $x_{k,ft}^g$ with sampling variance

$$\text{Var}(\hat{x}_{k,ft}^g) = \frac{x_{k,ft}^g(1 - x_{k,ft}^g)}{n_{ft}^L} \quad (22)$$

The observed share decomposes as $\hat{x}_{k,ft}^g = x_{k,ft}^g + \eta_{k,ft}^g$, with $\eta_{k,ft}^g$ mean-zero noise. This is classical errors-in-variables: substituting into (18), OLS attenuates toward zero,

$$\text{plim } \hat{\gamma}_k^g = \gamma_k^g \cdot \underbrace{\frac{\text{Var}(x_{k,ft}^g)}{\text{Var}(x_{k,ft}^g) + \text{Var}(\eta_{k,ft}^g)}}_{\in (0,1)} \quad (23)$$

The attenuation factor approaches one only as $n_{ft}^L \rightarrow \infty$. Firm fixed effects make it worse: demeaning removes within-firm signal variance, which is small for slowly varying shares, and leaves the noise variance unchanged,

$$\text{plim } \hat{\gamma}_{k,FE}^g = \gamma_k^g \cdot \frac{\text{Var}(x_{k,ft}^g - \bar{x}_{k,f}^g)}{\text{Var}(x_{k,ft}^g - \bar{x}_{k,f}^g) + \text{Var}(\eta_{k,ft}^g)} \quad (24)$$

One partial correction is to weight each observation by $w_{ft} = \min(\text{Coverage}_{ft}, 1) = \min(n_{ft}^L/n_{ft}^{Orb}, 1)$, proportional to LinkedIn sample size, which tilts the estimator toward firm-years with the highest signal-to-noise. The weighted attenuation factor

$$\text{plim } \hat{\gamma}_{k,WLS}^g = \gamma_k^g \cdot \frac{\sum_{ft} w_{ft} \text{Var}(x_{k,ft}^g)}{\sum_{ft} w_{ft} [\text{Var}(x_{k,ft}^g) + \text{Var}(\eta_{k,ft}^g)]} \quad (25)$$

exceeds the unweighted factor, reducing but not eliminating the bias. Consistent with this, the human capital coefficients rise monotonically with the coverage threshold (Appendix E.4), and IV estimates that instrument each share with its own lag exceed the weighted baseline (Appendix E.8).

C. Full Regression Output

C.1 Implied Structural Parameters

Table 12: Implied Production Function Parameters

Parameter	Assumption 1 only		Assumption 2 only		Assumptions 1–3			
					Managers (μ^m)		Non-Managers (μ^{nm})	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
α	0.338***	0.310***	0.332***	0.309***	0.329***	0.310***	0.329***	0.310***
(Capital elasticity)	(0.002)	(0.003)	(0.002)	(0.003)	(0.002)	(0.003)	(0.002)	(0.003)
β	0.661***	0.636***	0.661***	0.637***	0.665***	0.635***	0.665***	0.635***
(Labour elasticity)	(0.002)	(0.004)	(0.002)	(0.004)	(0.002)	(0.004)	(0.002)	(0.004)
λ	0.421***	0.234***			0.284***	0.258***	0.284***	0.258***
(Manager premium)	(0.025)	(0.039)			(0.025)	(0.040)	(0.025)	(0.040)
μ_{deg}			0.573***	0.202***	0.599***	0.114	0.510***	0.186***
(Degree premium)			(0.017)	(0.030)	(0.053)	(0.064)	(0.021)	(0.033)
μ_{post}			0.123***	0.058*	0.260***	0.030	0.057**	0.039
(Postgrad premium)			(0.018)	(0.024)	(0.040)	(0.044)	(0.021)	(0.022)
μ_{stem}			0.113***	0.020	0.097**	0.006	0.194***	0.118***
(STEM premium)			(0.015)	(0.020)	(0.036)	(0.040)	(0.017)	(0.019)
μ_{elite}			0.599***	0.102***	0.247***	0.038	0.600***	0.107***
(Top-20 premium)			(0.022)	(0.030)	(0.050)	(0.056)	(0.026)	(0.030)
Observations	111,960	111,960	111,960	111,960	111,960	111,960	111,960	111,960
Firm FE	No	Yes	No	Yes	No	Yes	No	Yes

Notes: Parameters derived from the raw regression in Table 16. All specifications include industry×year FE. Standard errors clustered at the firm level. Firms weighted by coverage; coverage polynomial included. *** p<0.001, ** p<0.01, * p<0.05.

Table 13: Wage Equation: Implied Wage Premia

Parameter	Assumption 1 only		Assumption 2 only		Assumptions 1–3			
	(1)	(2)	(3)	(4)	Managers (d^m)		Non-Managers (d^{nm})	
					(5)	(6)	(7)	(8)
d^l	0.485*** (0.014)	0.141*** (0.016)			0.385*** (0.013)	0.142*** (0.017)	0.385*** (0.013)	0.142*** (0.017)
d_{deg}			0.513*** (0.009)	0.189*** (0.013)	0.568*** (0.027)	0.196*** (0.030)	0.431*** (0.012)	0.135*** (0.014)
d_{post}			0.217*** (0.010)	0.049*** (0.010)	0.314*** (0.020)	0.078*** (0.020)	0.157*** (0.011)	0.031*** (0.009)
d_{stem}			0.127*** (0.008)	0.012 (0.009)	0.014 (0.018)	0.004 (0.018)	0.108*** (0.009)	0.005 (0.008)
d_{elite}			0.431*** (0.012)	0.044*** (0.013)	0.274*** (0.025)	0.021 (0.026)	0.409*** (0.014)	0.042*** (0.012)
Observations	111,960	111,960	111,960	111,960	111,960	111,960	111,960	111,960
Firm FE	No	Yes	No	Yes	No	Yes	No	Yes

Notes: Parameters derived from the raw regression in Table 15. All specifications include industry×year FE. Standard errors clustered at the firm level. Firms weighted by coverage; coverage polynomial included. *** p<0.001, ** p<0.01, * p<0.05.

Table 14: Implied Wage–Productivity Wedges

Parameter	Assumption 1 only		Assumption 2 only		Assumptions 1–3			
	(1)	(2)	(3)	(4)	Managers (ϕ^m)		Non-Managers (ϕ^{nm})	
					(5)	(6)	(7)	(8)
ϕ^l	0.206*** (0.022)	−0.008 (0.030)			0.196*** (0.021)	−0.022 (0.030)	0.196*** (0.021)	−0.022 (0.030)
ϕ_{deg}			0.134*** (0.014)	0.060** (0.023)	0.274*** (0.058)	0.133* (0.061)	0.092*** (0.018)	0.017 (0.025)
ϕ_{post}			0.136*** (0.016)	0.012 (0.018)	0.213*** (0.044)	0.065 (0.042)	0.119*** (0.018)	0.006 (0.017)
ϕ_{stem}			0.052*** (0.013)	−0.001 (0.016)	−0.064* (0.027)	0.000 (0.024)	−0.021 (0.014)	−0.070*** (0.014)
ϕ_{elite}			0.035 (0.019)	−0.021 (0.023)	0.168** (0.055)	−0.006 (0.054)	0.010 (0.022)	−0.026 (0.022)
Observations	111,960	111,960	111,960	111,960	111,960	111,960	111,960	111,960
Firm FE	No	Yes	No	Yes	No	Yes	No	Yes

Notes: Wage–productivity wedge $\hat{\phi}_j = \hat{\delta}_j - \hat{\gamma}_j$. Standard errors computed via feasible SUR: $\widehat{\text{Var}}(\hat{\phi}_j) = \widehat{\text{Var}}(\hat{\delta}_j) + \widehat{\text{Var}}(\hat{\gamma}_j) - 2\hat{\rho} \widehat{\text{SE}}(\hat{\delta}_j)\widehat{\text{SE}}(\hat{\gamma}_j)$, $\hat{\rho} = \text{corr}(\hat{e}^{wg}, \hat{e}^{pf})$. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

C.2 Baseline Wage Equation Estimates

Table 15: Wage Equation Estimates: Full Regression Output

Variable	(1)	(2)	(3)	(4)	(5)	(6)
Manager share s^m	0.485*** (0.014)	0.141*** (0.016)			0.385*** (0.013)	0.142*** (0.017)
Degree share			0.513*** (0.009)	0.189*** (0.013)		
Postgrad Deg.			0.217*** (0.010)	0.049*** (0.010)		
STEM Deg.			0.127*** (0.008)	0.012 (0.009)		
Elite Deg.			0.431*** (0.012)	0.044*** (0.013)		
$s^m \times M$ Degree					0.786*** (0.037)	0.224*** (0.034)
$s^m \times M$ Postgrad					0.435*** (0.028)	0.089*** (0.023)
$s^m \times M$ STEM					0.019 (0.025)	0.005 (0.021)
$s^m \times M$ Elite					0.379*** (0.035)	0.024 (0.030)
$(1 - s^m) \times NM$ Degree					0.431*** (0.012)	0.135*** (0.014)
$(1 - s^m) \times NM$ Postgrad					0.157*** (0.011)	0.031*** (0.009)
$(1 - s^m) \times NM$ STEM					0.108*** (0.009)	0.005 (0.008)
$(1 - s^m) \times NM$ Elite					0.409*** (0.014)	0.042*** (0.012)
Experience	0.110* (0.043)	0.052 (0.038)	0.887*** (0.043)	0.049 (0.038)		
Experience ²	-0.007 (0.008)	0.016* (0.008)	-0.121*** (0.008)	0.023** (0.008)		
Male share	0.480*** (0.010)	0.217*** (0.015)	0.533*** (0.010)	0.224*** (0.015)	0.586*** (0.010)	0.228*** (0.015)
$s^m \times M$ Experience					0.865*** (0.155)	0.094 (0.120)
$s^m \times M$ Exp. ²					-0.217*** (0.027)	-0.016 (0.022)
$(1 - s^m) \times NM$ Experience					0.608*** (0.046)	0.077* (0.039)
$(1 - s^m) \times NM$ Exp. ²					-0.059*** (0.009)	-0.027** (0.008)
Observations	111,960	111,960	111,960	111,960	111,960	111,960
Firm FE	No	Yes	No	Yes	No	Yes

C.3 Baseline Production Function Estimates

Table 16: Production Function Estimates: Full Regression Output

Variable	(1)	(2)	(3)	(4)	(5)	(6)
ln(Capital/Worker)	0.338*** (0.002)	0.310*** (0.003)	0.332*** (0.002)	0.309*** (0.003)	0.329*** (0.002)	0.310*** (0.003)
ln(Employment)	-0.001 (0.002)	-0.054*** (0.003)	-0.007*** (0.002)	-0.054*** (0.003)	-0.006*** (0.002)	-0.055*** (0.003)
Manager share s^m	0.279*** (0.017)	0.149*** (0.025)			0.189*** (0.016)	0.164*** (0.025)
Degree share			0.379*** (0.011)	0.129*** (0.019)		
Postgrad Deg.			0.081*** (0.012)	0.037* (0.015)		
STEM Deg.			0.075*** (0.010)	0.013 (0.013)		
Elite Deg.			0.396*** (0.015)	0.065*** (0.019)		
$s^m \times$ M Degree					0.512*** (0.045)	0.091 (0.051)
$s^m \times$ M Postgrad					0.222*** (0.034)	0.024 (0.035)
$s^m \times$ M STEM					0.083** (0.031)	0.005 (0.032)
$s^m \times$ M Elite					0.211*** (0.043)	0.030 (0.045)
$(1 - s^m) \times$ NM Degree					0.339*** (0.014)	0.118*** (0.021)
$(1 - s^m) \times$ NM Postgrad					0.038** (0.014)	0.025 (0.014)
$(1 - s^m) \times$ NM STEM					0.129*** (0.011)	0.075*** (0.012)
$(1 - s^m) \times$ NM Elite					0.399*** (0.017)	0.068*** (0.019)
Experience	0.282*** (0.052)	0.085 (0.058)	0.783*** (0.052)	0.088 (0.058)		
Experience ²	-0.065*** (0.010)	-0.008 (0.012)	-0.137*** (0.010)	0.010 (0.012)		
Male share	0.374*** (0.012)	0.114*** (0.022)	0.426*** (0.013)	0.128*** (0.022)	0.457*** (0.013)	0.127*** (0.022)
$s^m \times$ M Experience					1.121*** (0.190)	0.253 (0.181)
$s^m \times$ M Exp. ²					-0.272*** (0.033)	-0.062 (0.034)
$(1 - s^m) \times$ NM Experience					0.518*** (0.057)	0.109 (0.059)
$(1 - s^m) \times$ NM Exp. ²					-0.078*** (0.011)	-0.018 (0.012)
Observations	111,960	111,960	111,960	111,960	111,960	111,960
Firm FE	No	Yes	No	Yes	No	Yes

D. Heterogeneity

The specifications in this section combine Assumptions 1 and 2: managers and non-managers are perfect substitutes with a fixed premium λ , and all workers share a common set of human capital returns μ_k . Imposing $(1 + \lambda)\mu_k^m = \mu_k^{nm}$ on (8) collapses it to

$$\ln\left(\frac{VA}{L}\right)_{ft} \approx \ln A_{ft} + \alpha \ln\left(\frac{K}{L}\right)_{ft} + (\alpha + \beta - 1) \ln L_{ft} + \beta \lambda s_{ft}^m + \beta \sum_k \mu_k \bar{x}_{k,ft} \quad (26)$$

This sacrifices the role-specific returns of the full specification but reduces the parameter count enough to estimate within the smaller cells that arise when the panel is split by size, industry, or region. I report raw coefficients from (26); the implied structural parameters are each coefficient divided by $\hat{\beta}$. All specifications are cross-sectional, with industry-by-year fixed effects, coverage weights, and a quadratic coverage control.

D.1 By Firm Employment Size

Table 17 reports results by employment quartile of the estimation sample. Human capital effects rise with firm size: the degree coefficient climbs from 0.257 in Q1 to 0.492 in Q4, and the elite coefficient from 0.261 to 0.582. Experience returns are also larger in Q4. One explanation is complementarity between workers, where the value of a skilled worker's knowledge depends on how many others can draw on it; another is specialisation, where larger organisations let educated workers concentrate on the tasks that use their skills. Industry composition plays a supporting role, since service sectors are modestly larger and have larger returns (Section D.2). The manager share coefficient runs the other way: it is largest in Q1 (0.201), consistent with managers mattering more where decision-making is concentrated and standardised processes have not yet substituted for individual discretion.

Table 17: Heterogeneity by Firm Size

Variable	All	Q1 (< 71)	Q2 (71 - 114)	Q3 (115 - 212)	Q4 (>212)
ln(Capital/Worker)	0.333*** (0.001)	0.335*** (0.003)	0.330*** (0.003)	0.319*** (0.003)	0.320*** (0.003)
ln(Employment)	-0.007*** (0.001)	-0.076*** (0.010)	0.008 (0.020)	-0.024 (0.014)	-0.007 (0.004)
Manager share s^m	0.147*** (0.015)	0.201*** (0.028)	0.116*** (0.030)	0.020 (0.032)	0.103* (0.040)
Degree share	0.337*** (0.010)	0.257*** (0.020)	0.345*** (0.020)	0.360*** (0.021)	0.492*** (0.027)
Postgrad Deg.	0.082*** (0.011)	0.047* (0.020)	0.007 (0.020)	0.109*** (0.022)	0.181*** (0.032)
STEM Deg.	0.053*** (0.009)	0.056*** (0.016)	0.065*** (0.016)	0.060** (0.018)	0.074** (0.027)
Elite Deg.	0.344*** (0.013)	0.261*** (0.025)	0.301*** (0.025)	0.467*** (0.027)	0.582*** (0.037)
Experience	0.685*** (0.047)	0.529*** (0.087)	0.674*** (0.092)	0.583*** (0.098)	1.255*** (0.130)
Experience ²	-0.123*** (0.009)	-0.110*** (0.016)	-0.119*** (0.017)	-0.097*** (0.018)	-0.196*** (0.025)
Male share	0.440*** (0.011)	0.403*** (0.022)	0.395*** (0.022)	0.448*** (0.024)	0.461*** (0.030)
Observations	111,960	27,924	27,983	28,022	28,031

Notes: All specifications include industry×year FE. Standard errors clustered at the firm level. Firms weighted by coverage; coverage polynomial included. *** p<0.001, ** p<0.01, * p<0.05.

D.2 By Industry

Table 18 reports results by NACE section. The sharpest pattern is the divide between goods-producing and hospitality sectors (C–I) and knowledge-intensive services (J–N). In manufacturing and trade, degree returns are modest and the postgraduate, STEM, and experience coefficients are small and largely insignificant. In finance, ICT, professional services, and administrative support, degree returns are three to five times larger, and postgraduate qualifications carry a substantial premium in J–K (0.365) that is absent elsewhere. The manager premium is concentrated in manufacturing and in financial and professional services, and is essentially zero in trade and hospitality.

Table 18: Heterogeneity by Industry (NACE Section)

Variable	All	C	D-F	G-I	J-K	L-M	N
ln(Capital/Worker)	0.333*** (0.001)	0.368*** (0.004)	0.368*** (0.005)	0.368*** (0.003)	0.324*** (0.005)	0.279*** (0.005)	0.300*** (0.005)
ln(Employment)	-0.007*** (0.001)	-0.007* (0.003)	-0.012** (0.004)	-0.031*** (0.003)	0.010 (0.005)	0.003 (0.005)	-0.033*** (0.006)
Manager share s^m	0.147*** (0.015)	0.314*** (0.029)	0.075 (0.053)	0.013 (0.030)	0.362*** (0.059)	0.368*** (0.051)	0.132* (0.057)
Degree share	0.337*** (0.010)	0.174*** (0.020)	0.041 (0.028)	0.091*** (0.022)	0.544*** (0.040)	0.583*** (0.042)	0.419*** (0.036)
Postgrad Deg.	0.082*** (0.011)	-0.012 (0.018)	-0.016 (0.028)	0.006 (0.021)	0.365*** (0.050)	0.126** (0.044)	0.054 (0.046)
STEM Deg.	0.053*** (0.009)	-0.025 (0.015)	0.011 (0.020)	0.038** (0.016)	0.095* (0.043)	0.114** (0.037)	0.160*** (0.036)
Elite Deg.	0.344*** (0.013)	0.180*** (0.025)	0.102* (0.040)	0.058* (0.027)	0.515*** (0.053)	0.873*** (0.049)	0.654*** (0.062)
Experience	0.685*** (0.047)	0.318** (0.099)	0.021 (0.162)	0.105 (0.088)	0.726*** (0.192)	0.967*** (0.204)	1.406*** (0.215)
Experience ²	-0.123*** (0.009)	-0.062*** (0.018)	-0.041 (0.030)	-0.011 (0.017)	-0.106** (0.036)	-0.167*** (0.038)	-0.257*** (0.041)
Male share	0.440*** (0.011)	0.194*** (0.025)	0.336*** (0.038)	0.102*** (0.023)	0.592*** (0.050)	0.588*** (0.046)	0.554*** (0.043)
Observations	111,960	28,596	11,143	28,049	11,877	10,028	9,195

Notes: Column labels are NACE section codes: C = Manufacturing; D-F = Utilities & Construction (combined); G-I = Trade, Transport & Hospitality; J-K = Finance & ICT; L-M = Professional Services; N = Administrative & Support. All specifications include industry×year FE. Standard errors clustered at the firm level. Firms weighted by coverage; coverage polynomial included. *** p<0.001, ** p<0.01, * p<0.05.

D.3 By Region

Table 19 reports results by NUTS1 region. Several regions — the North East, Wales, and Northern Ireland — contribute fewer than 2,500 firm-years, so I focus on the larger cells. The striking result is the geography of the elite premium. In London the elite coefficient is 0.554, the largest in the table; in the South East it is 0.068 and insignificant, despite a larger degree premium (0.408 against 0.354) and a similar STEM coefficient. The elite premium is thus not a feature of high-skill regions in general but of London in particular. The most likely explanation is sectoral: finance, professional services, and ICT, where the elite premium is largest (Table 18), are concentrated in the capital. Distinguishing a sectoral from a spatial explanation would require within-sector regional variation that the sample does not support with precision.

Table 19: Heterogeneity by Region (NUTS1)

	NE	NW	Y + H	East.Mid	West.Mid	East.E
In(Capital/Worker)	0.267*** (0.013)	0.325*** (0.005)	0.317*** (0.006)	0.336*** (0.006)	0.332*** (0.006)	0.317*** (0.005)
In(Employment)	0.004 (0.012)	-0.029*** (0.005)	-0.023*** (0.006)	0.008 (0.005)	-0.019*** (0.006)	-0.005 (0.006)
Manager share s^m	0.012 (0.111)	0.219*** (0.046)	0.312*** (0.058)	0.093 (0.058)	0.135* (0.056)	0.069 (0.054)
Degree share	0.365*** (0.072)	0.077* (0.031)	0.275*** (0.037)	-0.007 (0.038)	0.250*** (0.037)	0.235*** (0.035)
Postgrad Deg.	-0.019 (0.072)	0.058 (0.033)	0.153*** (0.037)	-0.069 (0.038)	0.074* (0.037)	0.068 (0.035)
STEM Deg.	0.003 (0.057)	0.085*** (0.025)	0.113*** (0.028)	0.018 (0.028)	0.051 (0.029)	0.065* (0.029)
Elite Deg.	-0.092 (0.074)	0.186*** (0.046)	0.193*** (0.052)	0.118* (0.048)	0.065 (0.047)	0.084 (0.044)
Experience	0.819* (0.393)	1.251*** (0.165)	0.956*** (0.179)	0.679*** (0.190)	0.542** (0.181)	1.241*** (0.176)
Experience ²	-0.161* (0.073)	-0.216*** (0.030)	-0.169*** (0.032)	-0.117*** (0.035)	-0.082* (0.034)	-0.223*** (0.032)
Male share	0.505*** (0.085)	0.273*** (0.034)	0.268*** (0.043)	0.152*** (0.040)	0.242*** (0.042)	0.385*** (0.040)
Observations	2,550	11,894	8,471	8,355	9,827	10,050
	London	SE	SW	Wales	Scotland	N.Ireland
In(Capital/Worker)	0.302*** (0.003)	0.342*** (0.004)	0.325*** (0.006)	0.298*** (0.012)	0.319*** (0.007)	0.318*** (0.012)
In(Employment)	-0.013*** (0.004)	-0.001 (0.004)	-0.012 (0.007)	0.046*** (0.013)	0.028*** (0.008)	0.004 (0.011)
Manager share s^m	0.228*** (0.038)	0.072 (0.044)	0.252*** (0.068)	-0.018 (0.122)	0.327*** (0.074)	0.216 (0.117)
Degree share	0.354*** (0.031)	0.408*** (0.030)	0.187*** (0.045)	-0.036 (0.083)	0.039 (0.046)	0.514*** (0.068)
Postgrad Deg.	0.157*** (0.031)	0.019 (0.031)	0.052 (0.045)	-0.062 (0.078)	0.067 (0.047)	-0.091 (0.071)
STEM Deg.	0.124*** (0.027)	0.169*** (0.026)	0.039 (0.036)	0.161* (0.064)	0.065 (0.038)	0.298*** (0.060)
Elite Deg.	0.554*** (0.036)	0.068 (0.038)	0.031 (0.051)	0.187 (0.128)	0.184** (0.067)	0.352* (0.163)
Experience	1.064*** (0.113)	0.966*** (0.136)	0.888*** (0.187)	0.803* (0.385)	1.041*** (0.228)	0.339 (0.295)
Experience ²	-0.191*** (0.021)	-0.158*** (0.026)	-0.120*** (0.035)	-0.129 (0.070)	-0.165*** (0.043)	-0.045 (0.059)
Male share	0.538*** (0.033)	0.323*** (0.033)	0.419*** (0.047)	0.836*** (0.088)	0.330*** (0.051)	0.892*** (0.080)
Observations	23,129	16,983	7,466	2,381	6,100	2,222

Notes: All specifications include industry×year FE. Standard errors clustered at the firm level. Firms weighted by coverage; coverage polynomial included. *** p<0.001, ** p<0.01, * p<0.05.

E. Robustness

E.1 Alternative Specifications

Table 20 reports raw estimates from the joint model (26), with industry-by-year fixed effects, coverage weights, and a quadratic coverage control; implied structural parameters are each coefficient divided by $\hat{\beta}$.

Panel A varies the experience measure: firm-spell tenure (columns 1–2), job-spell tenure (3–4), and total labour-market experience (5–6). Firm and job tenure generate larger returns than general experience (0.685 and 0.663 against 0.336 cross-sectionally), consistent with match-specific capital paying more than general experience. All other coefficients are stable.

Panel B varies the human capital definitions. Columns (1)–(2) replace the combined STEM-and-Economics share with pure STEM; the coefficient falls to 0.026 cross-sectionally, so Economics graduates contribute independently. Columns (3)–(4) replace the top-20 dummy with continuous log rank, $\log(1 + \text{rank})$; the negative sign means that firms whose graduates attended higher-ranked universities have higher value added per worker. The postgraduate premium falls to zero in the cross-section when log rank is included: postgraduate holders tend to come from higher-ranked institutions (the firm-level correlation between the postgraduate share and mean log rank is -0.36), so without a rank control the postgraduate coefficient absorbs part of the prestige premium. Columns (5)–(6) replace the postgraduate share with the MBA share. The MBA coefficient is large in the cross-section (0.418) but small and insignificant within firms (0.053): the MBA premium reflects primarily sorting into productive firms rather than within-firm productivity differences.

Table 20: Robustness: Alternative Specifications

Panel A: Alternative Experience Measures						
	Firm Tenure		Job Tenure		Total Exp.	
	(1)	(2)	(3)	(4)	(5)	(6)
$\ln\left(\frac{K}{L}\right)$	0.333*** (0.001)	0.311*** (0.003)	0.332*** (0.001)	0.311*** (0.003)	0.332*** (0.001)	0.311*** (0.003)
$\ln(L)$	-0.007*** (0.001)	-0.055*** (0.003)	-0.005*** (0.001)	-0.052*** (0.003)	-0.004* (0.001)	-0.052*** (0.003)
Manager%	0.147*** (0.015)	0.159*** (0.022)	0.131*** (0.015)	0.182*** (0.022)	0.149*** (0.015)	0.179*** (0.022)
Degree %	0.337*** (0.010)	0.109*** (0.017)	0.395*** (0.011)	0.135*** (0.018)	0.381*** (0.011)	0.134*** (0.018)
Post Deg%	0.082*** (0.011)	0.036** (0.012)	0.050*** (0.011)	0.039** (0.012)	0.050*** (0.011)	0.039** (0.012)
STEM Deg%	0.053*** (0.009)	0.004 (0.010)	0.053*** (0.009)	0.005 (0.010)	0.055*** (0.009)	0.006 (0.010)
Elite Deg%	0.344*** (0.013)	0.049** (0.015)	0.357*** (0.013)	0.050** (0.015)	0.357*** (0.013)	0.050** (0.015)
Exper	0.685*** (0.047)	0.088 (0.050)	0.663*** (0.035)	0.163*** (0.045)	0.336*** (0.034)	0.163*** (0.046)
Exper ²	-0.123*** (0.009)	-0.012 (0.010)	-0.091*** (0.009)	-0.036** (0.012)	-0.070*** (0.009)	-0.039** (0.013)
Male%	0.440*** (0.011)	0.111*** (0.019)	0.418*** (0.011)	0.103*** (0.019)	0.423*** (0.012)	0.103*** (0.019)
Obs.	111,960	111,960	111,960	111,960	111,960	111,960
Firm FE	No	Yes	No	Yes	No	Yes
Panel B: Alternative Human Capital Definitions						
	Pure STEM		Log Rank		MBA	
	(1)	(2)	(3)	(4)	(5)	(6)
$\ln\left(\frac{K}{L}\right)$	0.334*** (0.001)	0.311*** (0.003)	0.330*** (0.001)	0.311*** (0.003)	0.333*** (0.001)	0.311*** (0.003)
$\ln(L)$	-0.006*** (0.001)	-0.055*** (0.003)	-0.006*** (0.001)	-0.055*** (0.003)	-0.007*** (0.001)	-0.055*** (0.003)
Manager%	0.152*** (0.015)	0.155*** (0.022)	0.113*** (0.015)	0.165*** (0.022)	0.117*** (0.015)	0.156*** (0.022)
Degree%	0.338*** (0.010)	0.108*** (0.017)	0.296*** (0.010)	0.108*** (0.017)	0.350*** (0.010)	0.115*** (0.017)
Post/(MBA) Deg%	0.097*** (0.011)	0.035** (0.012)	0.000 (0.011)	0.030* (0.012)	0.418*** (0.040)	0.053 (0.043)
STEM/(Econ) Deg%	0.026** (0.009)	0.020 (0.012)	0.067*** (0.009)	0.005 (0.010)	0.046*** (0.009)	0.009 (0.010)
Prestige	0.347*** (0.013)	0.048** (0.015)	-0.154*** (0.004)	-0.025*** (0.004)	0.367*** (0.013)	0.052*** (0.015)
Exper	0.694*** (0.047)	0.088 (0.050)	0.751*** (0.047)	0.080 (0.051)	0.676*** (0.046)	0.099* (0.050)
Exper ²	-0.125*** (0.009)	-0.012 (0.010)	-0.135*** (0.009)	-0.010 (0.010)	-0.121*** (0.009)	-0.014 (0.010)
Male%	0.474*** (0.012)	0.107*** (0.019)	0.431*** (0.011)	0.113*** (0.019)	0.437*** (0.011)	0.111*** (0.019)
Obs.	111,960	111,960	111,960	111,960	111,960	111,960
Firm FE	No	Yes	No	Yes	No	Yes

Notes: All specifications include industry×year fixed effects. *** p<0.001, ** p<0.01, * p<0.05.

E.2 The Managerial Hierarchy

The main text partitions the workforce into managers and non-managers. This appendix extends the labour aggregate to four groups — non-managers, frontline, middle, and executive managers, as defined in Section 3.1:

$$\tilde{L}_{ft} = L_{ft}^{nm} \left(1 + \sum_k \mu_k^{nm} \bar{x}_{k,ft}^{nm} \right) + \sum_{g \in \{fr, mid, exec\}} (1 + \lambda^g) L_{ft}^g \left(1 + \sum_k \mu_k^g \bar{x}_{k,ft}^g \right) \quad (27)$$

where λ^g is the premium of tier g over a non-manager and μ_k^g the within-tier premium of characteristic k . Writing $s_{ft}^g = L_{ft}^g / L_{ft}$ and log-linearising as in Section 2:

$$\begin{aligned} \ln \left(\frac{VA}{L} \right)_{ft} &\approx \ln A_{ft} + \alpha \ln \left(\frac{K}{L} \right)_{ft} + (\alpha + \beta - 1) \ln L_{ft} \\ &+ \beta \lambda^{fr} s_{ft}^{fr} + \beta \lambda^{mid} s_{ft}^{mid} + \beta \lambda^{exec} s_{ft}^{exec} \\ &+ \beta \sum_g (1 + \lambda^g \cdot \mathbf{1}[g \neq nm]) s_{ft}^g \sum_k \mu_k^g \bar{x}_{k,ft}^g \end{aligned} \quad (28)$$

estimated as a single cross-sectional OLS regression with a human capital block for each group, industry-by-year fixed effects, coverage weights, and a quadratic coverage control. I report raw coefficients; implied parameters are recovered by dividing by $\hat{\beta}$, and by $\hat{\beta}(1 + \hat{\lambda}^g)$ for within-tier premia.

Table 21 reports results. The three tier shares are positive but imprecise. The human capital results are more informative. For non-managers the pattern mirrors the main results: degree (0.364), STEM (0.168), and elite (0.473) are large and significant, and the postgraduate premium is not. For middle managers all four premia are large and significant. For executives the degree premium is large (1.374) but the postgraduate, STEM, and elite coefficients are insignificant: beyond holding a degree, the type and quality of education matter less at the top. Relative to the CEO literature (Bertrand and Schoar, 2003; Bandiera et al., 2020), the contribution is not a cleaner executive effect — their movers design is better identified — but an executive premium estimated net of the full educational composition of every other tier.

Table 21: Robustness: Heterogeneity by Managerial Tier

Variable	Non-managers	Frontline	Middle	Exec
ln(Capital/Worker)	0.322*** (0.002)	0.322*** (0.002)	0.322*** (0.002)	0.322*** (0.002)
ln(Employment)	0.005* (0.002)	0.005* (0.002)	0.005* (0.002)	0.005* (0.002)
Male share	0.393*** (0.019)	0.393*** (0.019)	0.393*** (0.019)	0.393*** (0.019)
Tier share s^g		0.045 (0.381)	0.141 (0.316)	0.164 (0.378)
$s^g \times \text{Degree}$	0.364*** (0.023)	-0.425*** (0.121)	1.077*** (0.104)	1.374*** (0.117)
$s^g \times \text{Postgrad} \mid \text{Deg.}$	0.021 (0.025)	-0.104 (0.094)	0.230** (0.082)	0.123 (0.088)
$s^g \times \text{STEM} \mid \text{Deg.}$	0.168*** (0.020)	0.154* (0.071)	0.246*** (0.072)	-0.142 (0.085)
$s^g \times \text{Elite} \mid \text{Deg.}$	0.473*** (0.031)	-0.049 (0.131)	0.922*** (0.116)	0.210 (0.112)
$s^g \times \text{Exper.}$	0.806*** (0.092)	0.885* (0.386)	1.015** (0.311)	0.468 (0.366)
$s^g \times \text{Exp.}^2$	-0.134*** (0.018)	-0.346*** (0.098)	-0.277*** (0.081)	0.050 (0.089)
Observations	63,933			
Firm FE	No			

Notes: All specifications include industry×year FE. Standard errors clustered at the firm level. Firms weighted by coverage; coverage polynomial included. HC variables pre-multiplied by group headcount share following (28). Sample restricted to firm-years with LinkedIn coverage across all four managerial tiers. *** p<0.001, ** p<0.01, * p<0.05.

E.3 Education and Promotion

Table 22 reports linear probability models on the 2025 LinkedIn cross-section. The first outcome equals one if the worker records at least one increase in Revelio seniority score across position changes; the second equals one for executive and middle managers. The education variables follow the conditional structure of the production function. Columns add fixed effects sequentially: cohort; 2-digit industry and travel-to-work area; firm. All columns control for a quadratic in total experience and a male dummy.

Every credential predicts both outcomes, within firms as well as across them. A degree raises the probability of ever being promoted by around 14 percentage points within firm, against a non-graduate promotion rate of around a quarter. The managerial margin looks

different: the degree itself adds little, and the field of study dominates. Within firms, a STEM degree raises the probability of holding a managerial role by 3.8 percentage points, nearly a third of the 13% managerial share, with postgraduate and top-20 credentials adding smaller but significant amounts. Read against Table 12, this completes the Peter-Principle mechanism of Section 5.2: STEM training carries the largest within-firm non-manager productivity premium, is the strongest predictor of reaching management, and carries no managerial productivity premium.

Table 22: Education, Promotion, and Managerial Attainment

	Ever promoted			In management		
	(1)	(2)	(3)	(4)	(5)	(6)
Degree	0.1711*** (0.0010)	0.1691*** (0.0010)	0.1436*** (0.0008)	0.0043*** (0.0006)	0.0037*** (0.0005)	0.0046*** (0.0004)
Postgrad Degree	0.0266*** (0.0013)	0.0301*** (0.0008)	0.0342*** (0.0008)	0.0114*** (0.0010)	0.0119*** (0.0007)	0.0108*** (0.0005)
STEM/Econ Degree	0.0451*** (0.0013)	0.0355*** (0.0016)	0.0240*** (0.0006)	0.0498*** (0.0009)	0.0437*** (0.0008)	0.0377*** (0.0005)
Top-20 Degree	0.0845*** (0.0019)	0.0779*** (0.0013)	0.0507*** (0.0011)	0.0183*** (0.0012)	0.0135*** (0.0009)	0.0085*** (0.0007)
Cohort FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry, TTWA FE	No	Yes	Yes	No	Yes	Yes
Firm FE	No	No	Yes	No	No	Yes
Observations	13,524,549	13,524,548	12,754,934	13,524,549	13,524,548	12,754,934
R^2	0.184	0.194	0.299	0.026	0.035	0.139

Notes: Linear probability models on the 2025 LinkedIn cross-section. *Ever promoted* = 1 if the worker records at least one increase in Revelio seniority score across observed position changes. *In management* = 1 for executive and middle managers. Education variables are conditional on holding a degree, matching the share structure of (18). All columns control for a quadratic in total experience and a male dummy. Number of job switches is excluded as a mediator. Standard errors clustered at the firm level. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

E.4 LinkedIn Coverage and Selection Bias

Section 4.2 identifies two downward biases in the LinkedIn shares: attenuation from sampling error, and selection from the over-representation of educated workers in thinly covered firms. Both predict that human capital coefficients should rise as the sample is restricted to better-covered firms, and both motivate the baseline choices of a coverage floor of 0.1, coverage weighting, Revelio sampling weights, and a quadratic coverage control. This appendix varies each in turn.

Table 23 estimates (26) on subsamples with coverage floors of > 0 , ≥ 0.1 (baseline), ≥ 0.3 , and ≥ 0.5 . The elite and degree premia rise monotonically across thresholds in the cross-section, as does the manager premium. The STEM premium is stable, so the attenuation and

selection mechanisms operate mainly on the institution-quality and degree-attainment margins.

Table 24 isolates the two weighting corrections: both coverage weights and Revelio weights (baseline); coverage weights only; Revelio weights only; neither. Human capital coefficients are larger with coverage weights, consistent with partial correction of attenuation. Revelio weights have little additional effect.

Table 25 varies the coverage control: none, linear, quadratic (baseline). The coverage coefficient is positive and significant: better-covered firms have higher value added per worker conditional on human capital. Including the control reduces the degree coefficient from 0.419 to 0.337 in the cross-section. The direction may seem at odds with the selection story, which predicts a negative covariance between observed education shares and coverage and so an upward correction. But the firm-level correlation between the degree share and coverage is strongly positive (0.36): better-covered firms employ more educated workforces, and this sorting dominates. Because coverage is positively associated with both education and productivity, omitting it biases the education coefficient upward, and the control absorbs the part of the raw correlation that runs through the coverage channel. Linear and quadratic controls give near-identical coefficients.

Table 23: Robustness: Coverage Threshold

	> 0		≥ 0.1 (baseline)		≥ 0.3		≥ 0.5	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ln(Capital/Worker)	0.335*** (0.001)	0.312*** (0.002)	0.333*** (0.001)	0.311*** (0.003)	0.332*** (0.002)	0.306*** (0.003)	0.331*** (0.002)	0.303*** (0.004)
ln(Employment)	-0.006*** (0.001)	-0.054*** (0.003)	-0.007*** (0.001)	-0.055*** (0.003)	-0.008*** (0.002)	-0.055*** (0.004)	-0.007** (0.002)	-0.054*** (0.005)
Manager share s^m	0.137*** (0.013)	0.132*** (0.018)	0.147*** (0.015)	0.159*** (0.022)	0.166*** (0.019)	0.206*** (0.029)	0.202*** (0.024)	0.258*** (0.038)
Degree share	0.326*** (0.009)	0.098*** (0.014)	0.337*** (0.010)	0.109*** (0.017)	0.358*** (0.013)	0.103*** (0.023)	0.384*** (0.016)	0.110*** (0.030)
Postgrad Deg.	0.082*** (0.009)	0.037*** (0.010)	0.082*** (0.011)	0.036** (0.012)	0.085*** (0.014)	0.038* (0.017)	0.072*** (0.019)	0.059* (0.024)
STEM Deg.	0.054*** (0.007)	0.005 (0.008)	0.053*** (0.009)	0.004 (0.010)	0.055*** (0.012)	0.009 (0.014)	0.071*** (0.016)	0.022 (0.021)
Elite Deg.	0.326*** (0.011)	0.046*** (0.013)	0.344*** (0.013)	0.049** (0.015)	0.405*** (0.017)	0.083*** (0.022)	0.489*** (0.023)	0.156*** (0.030)
Experience	0.624*** (0.039)	0.081* (0.041)	0.685*** (0.047)	0.088 (0.050)	0.771*** (0.061)	0.091 (0.068)	0.757*** (0.079)	0.046 (0.091)
Experience ²	-0.111*** (0.007)	-0.010 (0.008)	-0.123*** (0.009)	-0.012 (0.010)	-0.139*** (0.011)	-0.013 (0.014)	-0.137*** (0.015)	-0.009 (0.019)
Male share	0.412*** (0.010)	0.102*** (0.016)	0.440*** (0.011)	0.111*** (0.019)	0.468*** (0.015)	0.122*** (0.026)	0.496*** (0.019)	0.164*** (0.035)
Observations	162,732	162,732	111,960	111,960	88,697	88,697	61,700	61,700
Firm FE	No	Yes	No	Yes	No	Yes	No	Yes

Notes: All specifications include industry×year FE, are weighted by $\min(\text{Coverage}_{ft}, 1)$, use Revelio individual sampling weights, and include a quadratic coverage control. The baseline sample restriction is coverage ≥ 0.1 . Standard errors clustered at the firm level. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

Table 24: Robustness: Weights

	Baseline		Cov. wt. only		Revelio wt. only		Neither	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ln(Capital/Worker)	0.333*** (0.001)	0.311*** (0.003)	0.334*** (0.001)	0.311*** (0.003)	0.336*** (0.001)	0.312*** (0.003)	0.336*** (0.001)	0.312*** (0.003)
ln(Employment)	-0.007*** (0.001)	-0.055*** (0.003)	-0.007*** (0.001)	-0.055*** (0.003)	-0.004** (0.001)	-0.052*** (0.003)	-0.004** (0.001)	-0.052*** (0.003)
Manager share s^m	0.147*** (0.015)	0.159*** (0.022)	0.151*** (0.015)	0.158*** (0.022)	0.084*** (0.013)	0.101*** (0.018)	0.087*** (0.013)	0.100*** (0.018)
Degree share	0.337*** (0.010)	0.109*** (0.017)	0.339*** (0.010)	0.105*** (0.017)	0.245*** (0.009)	0.079*** (0.014)	0.247*** (0.009)	0.078*** (0.014)
Postgrad Deg.	0.082*** (0.011)	0.036** (0.012)	0.080*** (0.011)	0.035** (0.012)	0.075*** (0.008)	0.032*** (0.009)	0.071*** (0.008)	0.029** (0.009)
STEM Deg.	0.053*** (0.009)	0.004 (0.010)	0.048*** (0.009)	0.003 (0.010)	0.041*** (0.007)	0.012 (0.007)	0.037*** (0.007)	0.012 (0.008)
Elite Deg.	0.344*** (0.013)	0.049** (0.015)	0.340*** (0.014)	0.044** (0.016)	0.206*** (0.011)	0.001 (0.012)	0.200*** (0.011)	-0.006 (0.012)
Experience	0.685*** (0.047)	0.088 (0.050)	0.671*** (0.047)	0.090 (0.050)	0.561*** (0.038)	0.095* (0.041)	0.551*** (0.038)	0.096* (0.041)
Experience ²	-0.123*** (0.009)	-0.012 (0.010)	-0.120*** (0.009)	-0.012 (0.010)	-0.098*** (0.007)	-0.013 (0.008)	-0.097*** (0.007)	-0.013 (0.008)
Male share	0.440*** (0.011)	0.111*** (0.019)	0.441*** (0.011)	0.111*** (0.019)	0.365*** (0.010)	0.076*** (0.016)	0.366*** (0.010)	0.076*** (0.016)
Observations	111,960	111,960	111,960	111,960	111,960	111,960	111,960	111,960
Firm FE	No	Yes	No	Yes	No	Yes	No	Yes
Firm wt.	Cov.	Cov.	Cov.	Cov.	None	None	None	None
Share wt.	Revelio	Revelio	Raw	Raw	Revelio	Revelio	Raw	Raw

Notes: All specifications include industry×year FE and include a quadratic coverage control. Sample restriction is coverage ≥ 0.1 . Standard errors clustered at the firm level. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

Table 25: Robustness: Coverage Quadratic Control

	No control		Linear		Quadratic (baseline)	
	(1)	(2)	(3)	(4)	(5)	(6)
ln(Capital/Worker)	0.344*** (0.001)	0.313*** (0.003)	0.334*** (0.001)	0.311*** (0.003)	0.333*** (0.001)	0.311*** (0.003)
ln(Employment)	-0.013*** (0.001)	-0.063*** (0.003)	-0.007*** (0.001)	-0.055*** (0.003)	-0.007*** (0.001)	-0.055*** (0.003)
Manager share s^m	0.117*** (0.015)	0.159*** (0.022)	0.148*** (0.015)	0.159*** (0.022)	0.147*** (0.015)	0.159*** (0.022)
Degree share	0.419*** (0.010)	0.129*** (0.017)	0.344*** (0.010)	0.108*** (0.017)	0.337*** (0.010)	0.109*** (0.017)
Postgrad Deg.	0.099*** (0.011)	0.037** (0.012)	0.082*** (0.011)	0.036** (0.012)	0.082*** (0.011)	0.036** (0.012)
STEM Deg.	0.060*** (0.009)	0.006 (0.010)	0.053*** (0.009)	0.004 (0.010)	0.053*** (0.009)	0.004 (0.010)
Elite Deg.	0.349*** (0.013)	0.050** (0.015)	0.345*** (0.013)	0.049** (0.015)	0.344*** (0.013)	0.049** (0.015)
Experience	0.766*** (0.047)	0.065 (0.050)	0.681*** (0.047)	0.088 (0.050)	0.685*** (0.047)	0.088 (0.050)
Experience ²	-0.139*** (0.009)	-0.002 (0.010)	-0.122*** (0.009)	-0.012 (0.010)	-0.123*** (0.009)	-0.012 (0.010)
Male share	0.441*** (0.012)	0.111*** (0.019)	0.443*** (0.011)	0.111*** (0.019)	0.440*** (0.011)	0.111*** (0.019)
Coverage			0.203*** (0.006)	0.102*** (0.010)	0.045 (0.029)	0.084* (0.038)
Coverage ²					0.195*** (0.022)	0.013 (0.028)
Observations	111,960	111,960	111,960	111,960	111,960	111,960
Firm FE	No	Yes	No	Yes	No	Yes

Notes: All specifications include industry×year FE, are weighted by $\min(\text{Coverage}_{ft}, 1)$, use Revelio individual sampling weights. Sample restriction is coverage ≥ 0.1 . Standard errors clustered at the firm level. *** p<0.001, ** p<0.01, * p<0.05.

E.5 Sample Construction and Functional Form Choices

Table 26 reports results across four progressively restricted samples: a valid match and non-missing financials only (columns 1–2); dropping VA above the 99th percentile and adding the coverage floor (3–4); dropping large within-firm jumps, the baseline (5–6); and excluding multinationals (7–8). The human capital coefficients are stable throughout, so the baseline estimates are not driven by outliers or by consolidated multinational accounts.

Table 26: Robustness: Sample Construction

Variable	(i) Match only		(ii) +p99, cov $\geq$ 0.1		(iii) +No jumps		(iv) +No multi	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ln(K/L)	0.335*** (0.001)	0.312*** (0.002)	0.311*** (0.001)	0.297*** (0.003)	0.333*** (0.001)	0.311*** (0.003)	0.316*** (0.001)	0.295*** (0.003)
ln(L)	-0.006*** (0.001)	-0.054*** (0.003)	-0.006*** (0.001)	-0.054*** (0.003)	-0.007*** (0.001)	-0.055*** (0.003)	-0.004** (0.001)	-0.058*** (0.003)
Manager $^{\circ}$	0.137*** (0.013)	0.132*** (0.018)	0.184*** (0.022)	0.099*** (0.015)	0.147*** (0.015)	0.159*** (0.022)	0.122*** (0.015)	0.093*** (0.020)
Degree $^{\circ}$	0.326*** (0.009)	0.098*** (0.014)	0.319*** (0.010)	0.082*** (0.017)	0.337*** (0.010)	0.109*** (0.017)	0.323*** (0.010)	0.112*** (0.016)
Post Deg $^{\circ}$	0.082*** (0.009)	0.037*** (0.010)	0.063*** (0.010)	0.032** (0.012)	0.082*** (0.011)	0.036** (0.012)	0.075*** (0.010)	0.030** (0.011)
STEM Deg $^{\circ}$	0.054*** (0.007)	0.005 (0.008)	0.060*** (0.008)	0.003 (0.010)	0.053*** (0.009)	0.004 (0.010)	0.056*** (0.008)	0.010 (0.009)
Elite Deg $^{\circ}$	0.326*** (0.011)	0.046*** (0.013)	0.310*** (0.013)	0.042** (0.015)	0.344*** (0.013)	0.049** (0.015)	0.317*** (0.013)	0.045** (0.014)
Exper	0.624*** (0.039)	0.081* (0.041)	0.739*** (0.045)	0.074 (0.050)	0.685*** (0.047)	0.088 (0.050)	0.825*** (0.046)	-0.062 (0.048)
Exper ^{2}	-0.111*** (0.007)	-0.010 (0.008)	-0.130*** (0.008)	-0.008 (0.010)	-0.123*** (0.009)	-0.012 (0.010)	-0.144*** (0.009)	0.011 (0.010)
Male $^{\circ}$	0.412*** (0.010)	0.102*** (0.016)	0.430*** (0.011)	0.120*** (0.019)	0.440*** (0.011)	0.111*** (0.019)	0.419*** (0.011)	0.100*** (0.018)
Observations	162,732	162,732	126,364	126,364	111,960	111,960	109,671	109,671
Firm FE	No	Yes	No	Yes	No	Yes	No	Yes

Notes: All specifications include industry $\times$ year FE, are weighted by $\min(\text{Coverage}_{ft}, 1)$, and include a quadratic coverage control. Standard errors clustered at the firm level. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

Table 27 examines three functional form choices. Columns (1)–(2) replace log value added per worker with log sales per worker. Columns (3)–(4) restrict to firms with directly reported value added. Columns (5)–(6) add NUTS2 region-by-year fixed effects, and column (6) NUTS3 fixed effects, to address the concern that value-added productivity conflates efficiency with local markups (Hall, 1988; De Loecker et al., 2020). The sales specification includes intermediate inputs and so loads on capital intensity and scale: the capital coefficient and manager premium are higher in the cross-section and the human capital premia diluted. The direct value-added specification is close to the baseline, so the EBITDA-plus-staff-costs imputation is not driving the results. With region-year effects all four human capital premia stay close to baseline, so the education–productivity associations are not driven by geographic sorting; the manager premium falls, consistent with management-intensive firms locating in high-productivity areas.

Table 27: Robustness: Functional Form

	(i) Log sales per worker		(ii) Direct VA only		(iii) NUTS2×yr	(iv) + NUTS3
	(1)	(2)	(3)	(4)	(5)	(6)
ln(Capital/Worker)	0.485*** (0.002)	0.374*** (0.002)	0.347*** (0.002)	0.312*** (0.003)	0.325*** (0.001)	0.322*** (0.001)
ln(Employment)	-0.047*** (0.002)	-0.050*** (0.003)	-0.013*** (0.002)	-0.055*** (0.004)	-0.011*** (0.001)	-0.013*** (0.002)
Manager share s^m	0.417*** (0.019)	0.018 (0.020)	0.196*** (0.019)	0.003 (0.030)	0.107*** (0.015)	0.096*** (0.015)
Degree share	0.033* (0.013)	0.054*** (0.016)	0.343*** (0.013)	0.131*** (0.023)	0.275*** (0.010)	0.260*** (0.011)
Postgrad Deg.	0.098*** (0.013)	0.011 (0.011)	0.066*** (0.014)	0.007 (0.016)	0.058*** (0.011)	0.047*** (0.011)
STEM Deg.	0.098*** (0.011)	0.014 (0.009)	0.017 (0.011)	0.016 (0.013)	0.062*** (0.009)	0.069*** (0.009)
Elite Deg.	0.058*** (0.016)	0.009 (0.014)	0.290*** (0.017)	0.003 (0.021)	0.288*** (0.013)	0.274*** (0.014)
Experience	0.209*** (0.058)	0.111* (0.047)	0.489*** (0.065)	0.101 (0.074)	0.827*** (0.046)	0.807*** (0.047)
Experience ²	-0.063*** (0.011)	-0.016 (0.010)	-0.076*** (0.012)	-0.006 (0.015)	-0.143*** (0.009)	-0.139*** (0.009)
Male share	0.386*** (0.014)	0.085*** (0.018)	0.344*** (0.015)	0.050 (0.027)	0.430*** (0.011)	0.419*** (0.012)
Observations	127,323	125,726	77,079	75,180	96,455	91,311
Firm FE	No	Yes	No	Yes	No	No

Notes: All specifications include industry×year FE, are weighted by $\min(\text{Coverage}_{ft}, 1)$, and include a quadratic coverage control. Standard errors clustered at the firm level. *** p<0.001, ** p<0.01, * p<0.05.

E.6 Relaxing Perfect Substitutability: CES and Translog Specifications

Assumption 1 treats managers and non-managers as perfect substitutes in efficiency units. Theory points the other way: in span-of-control (Lucas, 1978) and knowledge-hierarchy models (Garicano, 2000), managerial input raises the productivity of the workers it coordinates. This appendix assesses how far the data depart from perfect substitutability. Estimates are in Table 28.

A natural generalisation nests the two within-group aggregates inside a CES:

$$\tilde{L}_{ft} = \left[\theta (\tilde{L}_{ft}^{nm})^{\frac{\sigma-1}{\sigma}} + (1-\theta) (\tilde{L}_{ft}^m)^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}}, \quad \tilde{L}_{ft}^g = L_{ft}^g \left(1 + \sum_k \mu_k^g \bar{x}_{k,ft}^g \right), \quad (29)$$

where σ is the elasticity of substitution between groups and θ a distribution parameter. The linear aggregate (2) is the limit $\sigma \rightarrow \infty$, with the manager premium recovered as $(1-\theta)/\theta = 1 + \lambda$; $\sigma < 1$ corresponds to gross complements. Point identification of σ from the value-added

equation alone is weak: σ enters only through the curvature of the labour aggregate in the group shares, which within-firm variation pins down imprecisely, and it is not separable from θ and λ without further structure. I therefore test the perfect-substitutes restriction with flexible approximations.

Translog approximation. A translog in $K_{ft}, L_{ft}^{nm}, L_{ft}^m$ provides a second-order approximation that nests Cobb–Douglas:

$$\ln V A_{ft} = \beta_0 + \sum_i \beta_i \ln X_{i,ft} + \frac{1}{2} \sum_i \sum_j \beta_{ij} \ln X_{i,ft} \ln X_{j,ft} + \varepsilon_{ft}, \quad X \in \{K, L^{nm}, L^m\}. \quad (30)$$

Variables are demeaned, so the first-order coefficients are output elasticities at the sample mean. The relevant object is the Allen elasticity of substitution, $\hat{\sigma}_{nm,m}^A \approx 1 + \hat{\beta}_{nm,m} / (\hat{s}_{nm} \hat{s}_m)$, where $\hat{s}_g$ are factor shares: values above unity indicate more substitutability than Cobb–Douglas, values below unity gross complements. Columns (5)–(6) give $\hat{\sigma}_{nm,m}^A = 1.19$ without and 1.08 with firm fixed effects: modestly above unity and far below the perfect-substitutes benchmark. The first-order labour coefficients are negative at the mean (-0.04 to -0.09 for $\ln L^m$ and -0.28 to -0.31 for $\ln L^{nm}$), implying implausible output elasticities, a symptom of collinearity among the logged inputs and their interactions. I read the translog as a descriptive bound rather than a structural estimate of σ .

Curvature in the manager share. Columns (3)–(4) augment the baseline with $(s_{ft}^m)^2$. Perfect substitutability implies the manager share enters linearly. Cross-sectionally the quadratic is positive and strongly significant (0.67); within firms it is small and insignificant (0.08) while the linear term remains positive (0.10, $p < 0.05$). The within-firm estimates support the linear-in-shares approximation; the cross-sectional convexity most plausibly reflects selection across firms.

Within-group additivity. Columns (7)–(8) add pairwise interactions among the conditional shares to probe Assumption 2. Additivity is rejected in both specifications ($F = 144$ and 5.4), far more strongly in the cross-section. The dominant departure is a positive Degree $\times$ Elite interaction (0.95 without, 0.15 with fixed effects): graduate quantity and graduate quality are complementary. Within firms the remaining interactions are small.

Relative-demand identification. Under (29) and competitive labour markets, the relative marginal product equals the relative wage,

$$\ln \left(\frac{w_{ft}^m}{w_{ft}^{nm}} \right) = \ln \frac{1 - \theta}{\theta} - \frac{1}{\sigma} \ln \left(\frac{L_{ft}^m}{L_{ft}^{nm}} \right) + \text{controls}_{ft}, \quad (31)$$

so the slope on log relative quantity is a sufficient statistic for σ . I do not pursue this: I observe neither managerial nor non-managerial wage bills, and the competitive wage-setting it requires is rejected for non-managers in Section 5.3.

Taken together, the specifications reject both restrictions statistically, but the economi-

cally relevant within-firm departures are modest: the manager/non-manager elasticity sits just above unity, and within-firm human capital interactions are small. The baseline linear specification is a reasonable approximation for the within-firm estimates that anchor the paper.

Table 28: Robustness: Relaxing Perfect Substitutability

	Baseline		Quad. s^m		Translog		Educ. interactions	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\ln(K/L)$	0.333*** (0.001)	0.311*** (0.003)	0.333*** (0.001)	0.311*** (0.003)			0.333*** (0.001)	0.311*** (0.003)
$\ln(L)$	-0.007*** (0.001)	-0.055*** (0.003)	-0.006*** (0.001)	-0.055*** (0.003)			-0.006*** (0.001)	-0.055*** (0.003)
s^m	0.147*** (0.015)	0.159*** (0.022)	0.187*** (0.042)	0.102* (0.053)			0.149*** (0.015)	0.158*** (0.022)
$(s^m)^2$			0.670*** (0.078)	0.083 (0.093)				
Exper	0.685*** (0.047)	0.088 (0.050)	0.698*** (0.047)	0.087 (0.050)	0.701*** (0.047)	0.082 (0.051)	0.709*** (0.047)	0.092 (0.051)
Exper ²	-0.123*** (0.009)	-0.012 (0.010)	-0.124*** (0.009)	-0.012 (0.010)	-0.125*** (0.009)	-0.012 (0.010)	-0.125*** (0.009)	-0.013 (0.010)
Male share	0.440*** (0.011)	0.111*** (0.019)	0.435*** (0.012)	0.111*** (0.019)	0.424*** (0.012)	0.106*** (0.019)	0.425*** (0.012)	0.114*** (0.019)
Degree share	0.337*** (0.010)	0.109*** (0.017)	0.333*** (0.010)	0.108*** (0.017)	0.343*** (0.010)	0.108*** (0.017)	0.313*** (0.010)	0.104*** (0.017)
Postgrad Deg.	0.082*** (0.011)	0.036** (0.012)	0.081*** (0.011)	0.036** (0.012)	0.085*** (0.011)	0.035** (0.012)	0.055*** (0.011)	0.038** (0.012)
STEM Deg.	0.053*** (0.009)	0.004 (0.010)	0.054*** (0.009)	0.004 (0.010)	0.052*** (0.009)	0.005 (0.010)	0.056*** (0.009)	0.002 (0.010)
Elite Deg.	0.344*** (0.013)	0.049** (0.015)	0.343*** (0.013)	0.049** (0.015)	0.337*** (0.013)	0.048** (0.015)	0.269*** (0.013)	0.055*** (0.016)
$\ln K$					0.339*** (0.001)	0.319*** (0.003)		
$\ln L^m$					-0.041*** (0.002)	-0.087*** (0.003)		
$\ln L^{nm}$					-0.307*** (0.003)	-0.279*** (0.004)		
$(\ln K)^2$					-0.010*** (0.001)	-0.019*** (0.001)		
$(\ln L^m)^2$					-0.004*** (0.000)	-0.007*** (0.000)		
$(\ln L^{nm})^2$					-0.018*** (0.001)	-0.015*** (0.001)		
$\ln K \times \ln L^m$					-0.001 (0.001)	-0.004*** (0.001)		
$\ln K \times \ln L^{nm}$					0.017*** (0.001)	0.015*** (0.002)		
$\ln L^m \times \ln L^{nm}$					0.020*** (0.001)	0.029*** (0.002)		
Degree $\times$ Elite							0.946*** (0.047)	0.152** (0.057)
Degree $\times$ Postgrad							0.225*** (0.037)	0.028 (0.045)
Elite $\times$ STEM							0.170*** (0.049)	-0.147** (0.049)
Postgrad $\times$ STEM							0.046 (0.036)	-0.014 (0.037)
Obs	111,960	111,960	111,960	111,960	111,960	111,960	111,960	111,960
Firm FE	No	Yes	No	Yes	No	Yes	No	Yes

Notes: All specs include industry $\times$ year FE, are weighted by coverage, and include a quadratic coverage control. Translog inputs are demeaned log. Standard errors clustered at the firm level. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

E.7 Nonlinear Least Squares Estimates

The baseline (8) replaces $\ln(1 + z_{ft})$ in (7) with z_{ft} , an approximation that is accurate within firms but strained in the cross-section, where the premia are larger (Section 4.2). I therefore re-estimate the cross-sectional specification by nonlinear least squares (NLS), retaining the exact term:

$$\min_{\alpha, \beta, \lambda, \{\mu_k^g\}} \sum_{ft} \omega_{ft} (\ln VA_{ft} - \xi_{jt} - \alpha \ln K_{ft} - \beta \ln L_{ft} - \beta \ln(1 + z_{ft}))^2, \quad (32)$$

where z_{ft} is defined in (7), ξ_{jt} are industry-year effects, and $\omega_{ft} = \min(\text{Coverage}_{ft}, 1)$. NLS also imposes the structural restrictions that the linearised form recovers only by division: a single β scales the whole labour aggregate, and the managerial premia operate on the common base $(1 + \lambda)$. I initialise at the linearised estimates.

Table 29 reports the estimates. The production function parameters are essentially unchanged. The manager premium $\hat{\lambda}$ falls from 0.284 to 0.249 but remains large and significant. All parameters keep their sign and significance and the ordering of premia is preserved. The shifts are within the approximation error of order 0.08 log points implied by $\hat{z}_{ft} \approx 0.41$.

Table 29: Robustness: Nonlinear least squares estimates of the structural production function

Parameter	(1) Linearised	(2) NLS
α	0.329*** (0.002)	0.331*** (0.001)
β	0.665*** (0.002)	0.666*** (0.002)
λ	0.284*** (0.025)	0.249*** (0.034)
μ_{deg}^{nm}	0.510*** (0.021)	0.525*** (0.029)
μ_{post}^{nm}	0.057** (0.021)	0.109*** (0.024)
μ_{stem}^{nm}	0.194*** (0.017)	0.123*** (0.019)
μ_{elite}^{nm}	0.600*** (0.026)	0.542*** (0.034)
μ_{deg}^m	0.599*** (0.053)	0.513*** (0.066)
μ_{post}^m	0.260*** (0.040)	0.285*** (0.048)
μ_{stem}^m	0.097** (0.036)	0.138*** (0.040)
μ_{elite}^m	0.247*** (0.050)	0.218*** (0.062)
Observations	111,960	111,960
Taylor approximation	Yes	No
Ind. $\times$ Year FE	Yes	Yes

Notes: Cross-sectional estimates; industry $\times$ year fixed effects, no firm fixed effects. Firms weighted by $\omega_{ft} = \min(\text{Coverage}_{ft}, 1)$. Column (1) reports structural parameters backed out from the linearised estimating equation (8). Column (2) estimates these parameters directly by nonlinear least squares on equation (7), retaining $\ln(1 + z_{ft})$ rather than its first-order approximation. Experience quadratic, male share, and a quadratic in coverage enter as controls in both columns, unreported. Standard errors via the delta method in column (1) and from the NLS estimator in column (2), clustered at the firm level. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

E.8 Lagged and Instrumental Variable Estimates

This appendix uses the panel structure to probe two concerns: that contemporaneous productivity shocks drive the hiring or promotion of educated workers, and that measurement error attenuates the coefficients. Table 30 reports four specifications on the pooled model (26): the baseline (columns 1–2); lagged shares added alongside contemporaneous shares (3–4); lagged shares only (5–6); and contemporaneous shares instrumented with their own lags (7–8).

Reverse causality. If the associations were driven by contemporaneous shocks, using only lagged shares should attenuate them sharply. Column (5) shows it does not: the cross-sectional coefficients under the lag-only specification are close to baseline (degree 0.281 against 0.337, elite 0.315 against 0.344, postgraduate and STEM unchanged). Last year’s composition cannot have been caused by this year’s shock, so the persistence argues against simple reverse causality. The within-firm estimates are more fragile: the elite premium falls from 0.049 to an insignificant 0.002. Once both the firm mean and contemporaneous timing are removed, the identifying variation is thin.

Dynamics. Column (3), with contemporaneous and lagged shares together, suggests the productivity association builds over more than one period. The within-firm contemporaneous degree coefficient is 0.103 and the lagged coefficient a further 0.121, both significant; the same holds for postgraduate qualifications. This is consistent with educated hires raising measured productivity with a delay, through training, knowledge diffusion, or reorganisation.

Measurement error. If classical measurement error is attenuating the coefficients, instrumenting each share with its lag, whose error is independent of the current period’s, should raise them. Columns (7)–(8) are consistent with this. The IV estimates exceed baseline for the manager share, postgraduate, STEM, and elite coefficients in both the cross-section and within firms, and the cross-sectional elite premium rises to 0.418 from 0.344; the cross-sectional degree share is unchanged. The within-firm IV estimates are indicative only: the degree coefficient of 0.640 (column 8) is implausibly large and imprecise (standard error 0.097 against 0.017 at baseline).

Table 30: Robustness: Lagged and Instrumental Variable Estimates

	Baseline		Baseline + Lag HC		Lag HC only		IV: HC shares	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ln(Capital/Worker)	0.333*** (0.001)	0.311*** (0.003)	0.334*** (0.002)	0.311*** (0.003)	0.336*** (0.002)	0.311*** (0.003)	0.333*** (0.002)	0.317*** (0.003)
ln(Employment)	-0.007*** (0.001)	-0.055*** (0.003)	-0.007*** (0.002)	-0.053*** (0.003)	-0.007*** (0.002)	-0.051*** (0.003)	-0.007*** (0.002)	-0.046*** (0.003)
Manager share s^m	0.147*** (0.015)	0.159*** (0.022)	0.114*** (0.027)	0.127*** (0.025)			0.168*** (0.020)	0.412*** (0.047)
Degree share	0.337*** (0.010)	0.109*** (0.017)	0.307*** (0.020)	0.103*** (0.019)			0.324*** (0.013)	0.640*** (0.097)
Postgrad Deg.	0.082*** (0.011)	0.036** (0.012)	0.077*** (0.017)	0.045** (0.014)			0.084*** (0.015)	0.108** (0.041)
STEM Deg.	0.053*** (0.009)	0.004 (0.010)	0.045** (0.014)	0.029* (0.012)			0.061*** (0.012)	0.064 (0.035)
Elite Deg.	0.344*** (0.013)	0.049** (0.015)	0.296*** (0.022)	0.054** (0.018)			0.418*** (0.018)	0.073 (0.057)
Lag manager share			0.039 (0.025)	0.045* (0.020)	0.128*** (0.014)	0.079*** (0.019)		
Lag degree share			0.063*** (0.019)	0.121*** (0.016)	0.281*** (0.010)	0.095*** (0.015)		
Lag postgrad Deg.			0.006 (0.015)	0.046*** (0.012)	0.083*** (0.010)	0.033** (0.011)		
Lag STEM Deg.			0.013 (0.013)	0.021* (0.010)	0.045*** (0.008)	0.014 (0.009)		
Lag elite Deg.			0.087*** (0.020)	0.019 (0.016)	0.315*** (0.013)	0.002 (0.015)		
Experience	0.685*** (0.047)	0.088 (0.050)	0.780*** (0.050)	0.116* (0.055)	0.707*** (0.049)	0.116* (0.055)	0.763*** (0.050)	0.091 (0.057)
Experience ²	-0.123*** (0.009)	-0.012 (0.010)	-0.141*** (0.009)	-0.016 (0.011)	-0.131*** (0.009)	-0.009 (0.011)	-0.138*** (0.009)	-0.008 (0.012)
Male share	0.440*** (0.011)	0.111*** (0.019)	0.447*** (0.012)	0.109*** (0.021)	0.427*** (0.012)	0.107*** (0.021)	0.443*** (0.013)	0.075** (0.023)
Observations	111,960	111,960	109,207	107,717	109,493	108,020	109,207	107,717
Firm FE	No	Yes	No	Yes	No	Yes	No	Yes
Estimator	OLS	OLS	OLS	OLS	OLS	OLS	IV	IV

Notes: All specifications include industry×year FE, are weighted by $\min(\text{Coverage}_{ft}, 1)$, and include a quadratic coverage control. Standard errors clustered at the firm level. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

E.9 Long Differences

Long differences offer a partial remedy to the attenuation that firm fixed effects induce. Comparing a firm with itself h years apart recovers more of the genuine variation in composition, which accumulates over time, while still purging time-invariant heterogeneity. Table 31 reports estimates of (26) over 3-, 5-, and 7-year windows, alongside the cross-sectional and fixed-effects baselines, on a common sample of firms observed at the endpoints of every window.

The coefficients behave as the attenuation argument predicts: the long-difference estimates lie above the fixed-effects estimates and below the cross-section. The degree premium

is 0.185 at three years, against 0.109 under fixed effects and 0.337 in the cross-section. The capital coefficient is stable across all five columns (0.29–0.33), as expected for a well-measured input.

Two qualifications apply. The human capital coefficients decline modestly as the window lengthens. A pure measurement-error account predicts the opposite, since longer differences carry more signal; the decline more likely reflects compositional shifts toward firms with longer histories and more scope for confounding trends. The long differences therefore bound the attenuation rather than correct it. And the STEM premium, the least robust coefficient in the pooled specification, attenuates to zero at seven years. The manager premium is only significant at seven years (0.111), reflecting limited within-firm variation in managerial composition over short windows. Overall, the within-firm estimates understate the returns to education through attenuation, the cross-sectional estimates overstate them through sorting, and the truth lies between, closer to the long differences.

Table 31: Robustness: Long Differences

Variable	Levels (baseline)		Long differences		
	(1)	(2)	(3)	(4)	(5)
ln(Capital/Worker)	0.333*** (0.001)	0.311*** (0.003)	0.305*** (0.004)	0.297*** (0.003)	0.291*** (0.004)
ln(Employment)	-0.007*** (0.001)	-0.055*** (0.003)	-0.034*** (0.004)	-0.041*** (0.004)	-0.031*** (0.004)
Manager share s^m	0.147*** (0.015)	0.159*** (0.022)	0.005 (0.030)	0.041 (0.029)	0.111*** (0.031)
Degree share	0.337*** (0.010)	0.109*** (0.017)	0.185*** (0.024)	0.153*** (0.023)	0.124*** (0.024)
Postgrad Deg.	0.082*** (0.011)	0.036** (0.012)	0.066*** (0.019)	0.054** (0.018)	0.050** (0.019)
STEM Deg.	0.053*** (0.009)	0.004 (0.010)	0.062*** (0.016)	0.040** (0.015)	0.005 (0.016)
Elite Deg.	0.344*** (0.013)	0.049** (0.015)	0.108*** (0.025)	0.069** (0.025)	0.067** (0.025)
Experience	0.685*** (0.047)	0.088 (0.050)	-0.154* (0.076)	-0.055 (0.070)	-0.062 (0.072)
Experience ²	-0.123*** (0.009)	-0.012 (0.010)	-0.019 (0.015)	-0.009 (0.014)	-0.013 (0.014)
Male share	0.440*** (0.011)	0.111*** (0.019)	0.140*** (0.027)	0.217*** (0.026)	0.149*** (0.027)
Observations	111,960	111,960	34,115	33,563	32,760
Estimator	X-sec	firm FE	Δ_3	Δ_5	Δ_7

Notes: All long difference specifications estimated on a common sample of firms observed at both endpoints of all three difference windows, restricted to coverage ≥ 0.1 ($N_{\text{firms}} = 10,574$). Columns (3)–(5) are long-difference estimates: $\Delta_h[x]_{ft} = x_{ft} - x_{f,t-h}$ for $h \in \{3, 5, 7\}$. All differenced specifications include industry×year fixed effects on the end year t and a linear change-in-coverage control $\Delta_h[\text{cov}]$. Observations weighted by coverage at the end year. Standard errors clustered at the firm level. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

E.10 Gender and Hellerstein–Neumark Replication

A natural concern about the wedges in Section 5.3 is that the managerial wedge is the gender wedge of Hellerstein et al. (1999) in disguise. Three facts invite the concern. The male share enters both equations positively, with a wage coefficient that exceeds its productivity coefficient (Tables 16–15), so a gender wedge already sits in the estimates. Managers are disproportionately male: 60% against 55% in the population, and 62% against 56% in the regression sample. And the managerial wedge is positive in the cross-section.

This appendix makes gender the object of interest. Hellerstein et al. (1999) estimate that women in US manufacturing are around 16% less productive than men but paid around 45% less, a pay shortfall of around 29 percentage points in excess of the productivity shortfall. I recover the UK analogue and ask whether it can account for the managerial wedge. It cannot. The gender wedge is large and survives firm fixed effects, but the difference in male share between managers and non-managers is too small for one wedge to explain the other, in either direction.

Estimating equations. I estimate four nested SUR specifications, each pairing a production function with a wage equation. The pooled production function (Specification A) is

$$\ln\left(\frac{VA}{L}\right)_{ft} = \gamma_K \ln\left(\frac{K}{L}\right)_{ft} + \gamma_L \ln L_{ft} + \gamma_{male} \bar{x}_{ft}^{male} + \Gamma' \mathbf{X}_{ft} + \phi_{jt} + \phi_f + \varepsilon_{ft}, \quad (33)$$

where $\mathbf{X}_{ft}$ collects experience and coverage controls, with the wage equation replacing γ by δ and the dependent variable by $\ln(W/L)_{ft}$. The gender wedge is $\hat{\phi}_{male} = \hat{\delta}_{male} - \hat{\gamma}_{male}$. Specifications B and C add the pooled education shares and the manager share. Specification D interacts gender and education with role:

$$\begin{aligned} \ln\left(\frac{VA}{L}\right)_{ft} = & \gamma_K \ln\left(\frac{K}{L}\right)_{ft} + \gamma_L \ln L_{ft} + \gamma_{man} s_{ft}^m \\ & + (1 - s_{ft}^m) \left(\gamma_{male}^{nm} \bar{x}_{ft}^{nm,male} + \sum_k \gamma_k^{nm} \bar{x}_{k,ft}^{nm} \right) \\ & + s_{ft}^m \left(\gamma_{male}^m \bar{x}_{ft}^{m,male} + \sum_k \gamma_k^m \bar{x}_{k,ft}^m \right) + \Gamma' \mathbf{X}_{ft} + \phi_{jt} + \phi_f + \varepsilon_{ft}, \end{aligned} \quad (34)$$

which yields role-specific gender wedges $\hat{\phi}_{male}^g$ for $g \in \{m, nm\}$.

Role and gender are correlated but distinct cuts of the workforce, which bounds how much of one wedge the other can explain. Managers are only six percentage points more male than non-managers, so even if the managerial wedge were entirely a gender artefact it could be no larger than $0.06 \times 0.156 \approx 0.01$, about 5% of the cross-sectional managerial wedge of 0.206.

Results. Column A of Table 32 delivers the Hellerstein et al. (1999) result for UK firms over 2015–2024: the male wage premium exceeds the male productivity premium by 0.156 log points cross-sectionally, so women are paid below their relative marginal product. The wedge falls to 0.103 under firm fixed effects but remains large and significant, unlike the managerial

wedge, which collapses to zero. Because value added and the wage bill are per head, the level premia partly reflect men's longer hours; the wedge nets this out insofar as hours scale output and pay proportionally.

The gender wedge is robust to richer controls. Within firms it is flat across the education and role specifications (0.096 in B, 0.101 in C); cross-sectionally, education absorbs part of it (0.156 to 0.107) but it stays large, and conditioning on role does not absorb it further. Symmetrically, the managerial wedge is estimated net of the male share throughout the main text: dropping the gender control raises the cross-sectional managerial wedge by only about 0.01, in line with the bound above. Cross-sectionally the gender wedge is larger for managers than non-managers (0.224 against 0.130, Specification D), consistent with gendered bargaining over match surplus or with more elastic labour supply among male managers. The within-firm point estimates preserve the ordering (0.116 against 0.064) but neither is distinguishable from zero.

Table 32: Implied Wage–Productivity Wedges

Parameter	A: Male share		B: +Education		C: Role-interacted		D: Gender×role	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
<i>Gender wage–productivity wedge</i>								
ϕ_{male}	0.156*** (0.016)	0.103*** (0.027)	0.107*** (0.016)	0.096*** (0.027)	0.129*** (0.016)	0.101*** (0.027)		
ϕ_{male}^M							0.224* (0.090)	0.116 (0.083)
ϕ_{male}^{NM}							0.130*** (0.030)	0.064 (0.047)
<i>Manager wage–productivity wedge</i>								
ϕ^l					0.196*** (0.021)	−0.022 (0.030)	0.171*** (0.040)	0.007 (0.047)
<i>Education wage–productivity wedges</i>								
ϕ_{deg}			0.134*** (0.014)	0.060** (0.023)				
ϕ_{post}			0.136*** (0.016)	0.012 (0.018)				
ϕ_{stem}			0.052*** (0.013)	−0.001 (0.016)				
ϕ_{elite}			0.035 (0.019)	−0.021 (0.023)				
ϕ_{deg}^M					0.274*** (0.058)	0.133* (0.061)	0.292*** (0.076)	0.134 (0.076)
ϕ_{post}^M					0.213*** (0.044)	0.065 (0.042)	0.216*** (0.050)	0.067 (0.046)
ϕ_{stem}^M					−0.064* (0.027)	0.000 (0.024)	−0.112** (0.042)	0.069 (0.042)
ϕ_{elite}^M					0.168** (0.055)	−0.006 (0.054)	0.163* (0.066)	−0.056 (0.062)
ϕ_{deg}^{NM}					0.092*** (0.018)	0.017 (0.025)	0.092*** (0.026)	0.015 (0.032)
ϕ_{post}^{NM}					0.119*** (0.018)	0.006 (0.017)	0.119*** (0.021)	0.006 (0.019)
ϕ_{stem}^{NM}					−0.021 (0.014)	−0.070*** (0.014)	0.025 (0.018)	0.001 (0.017)
ϕ_{elite}^{NM}					0.010 (0.022)	−0.026 (0.022)	0.009 (0.040)	−0.026 (0.027)
Observations	111,960	111,960	111,960	111,960	111,960	111,960	111,960	111,960
Firm FE	No	Yes	No	Yes	No	Yes	No	Yes

Notes: Standard errors computed via feasible SUR. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.