

NBER WORKING PAPER SERIES

THE ROLE OF ORGANIZATION IN FIRM-SPONSORED TRAINING

David J. Deming
Raffaella Sadun
Christopher T. Stanton

Working Paper 35920
<http://www.nber.org/papers/w35920>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
October 2026

I have nothing to disclose beyond executive education activities that support firm training and my role as Chair of the Board of Harvard Business School Publishing. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research. This research was supported by a grant from the Russell Sage Foundation and by the Division of Research and Faculty Development at Harvard Business School. No other financial support was received for this research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2026 by David J. Deming, Raffaella Sadun, and Christopher T. Stanton. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

The Role of Organization in Firm-Sponsored Training
David J. Deming, Raffaella Sadun, and Christopher T. Stanton
NBER Working Paper No. 35920
October 2026
JEL No. J2, J24, M5, M51

ABSTRACT

Some firms are more likely than others to invest in workers' skills. We document this phenomenon and discuss its implications in light of models of firm-sponsored skills training. The data show that better managed firms, even in the same industries and labor markets, are more likely to train, while training clusters around either managerial or frontline worker development. We argue that organizational differences provide incentives for firms to train beyond traditional explanations that focus on labor market frictions. We illustrate how organizational heterogeneity can generate training investments through a knowledge hierarchy framework where top managers with different skills build bespoke organizations tailored to their unique needs. This in turn makes general skill firm-specific, allowing firms to capture returns on training investment.

David J. Deming
Harvard University
Harvard Kennedy School
and NBER
david_deming@harvard.edu

Christopher T. Stanton
Harvard University
Harvard Business School
and NBER
cstanton@hbs.edu

Raffaella Sadun
Harvard University
Harvard Business School
and NBER
rsadun@hbs.edu

Some firms invest heavily in worker development, while other observationally similar competitors invest almost nothing. Why do otherwise comparable firms make such different choices about whom to train, how much to invest, and what skills to build?

These differences matter. If wage growth reflects rising productivity, most life-cycle human capital accumulation happens at work rather than in school (for example, Deming 2025). A growing body of evidence on firm effects points to the importance of early career employment in high-wage-growth firms as a pronounced source of lifetime earnings variation. Even within the same industry and labor market, high-wage-growth firms offer distinct early career wage growth paths that persist for decades (Di Addario et al. 2023; Arellano-Bover and Saltiel 2026). Theories of within firm learning and task assignment (Gibbons and Waldman 2006) and structural analysis of the returns to training (Adda and Dustmann 2023) point to firm-sponsored skill investment as a leading candidate mechanism for these differences, though direct evidence remains thin.

Standard labor economic theories of on-the-job training start from Becker's (1964) famous argument that firms in perfectly competitive labor markets will not pay for general human capital because the benefits accrue only to workers who are being paid the marginal product of their labor. Subsequent work has described labor market frictions that can allow firms to appropriate part of the surplus from training investments: as one example, if a firm has monopsony power in the labor market or if employees face search frictions, wages are compressed relative to productivity (as in Acemoglu and Pischke 1999), and firm-sponsored training even for general human capital can be viable when labor market imperfections allow firms to capture some of the returns from firm outlays on training.

A different account of these facts starts from firms choosing to train workers as a response to organizational complementarities that increase the overall returns to skill investment, not only the share of those returns a firm can capture. Under this view, training is embedded in a broader bundle of productivity-enhancing practices (performance management, promotion ladders, and technology adoption) that vary systematically across firms. Some firms provide intensive worker development at entry and little thereafter; others invest continuously in frontline upskilling; still others concentrate resources on managerial and leadership training higher in the hierarchy.

The result is an employment landscape in which learning opportunities – and therefore human capital accumulation – are deeply shaped by organizational context. This perspective accounts for the within-industry variation in training across firms and shifts the question from *why firms train* to *why some firms build organizations in which training is valuable*.

We begin by characterizing some empirical patterns in on-the-job training, but in this task, we confront a problem with availability of data on training. The US government has not carried out a systematic survey of workplace training since the 1995 Survey of Employer-Provided Training. The only available US data on spending levels for on-the-job training is from industry sources not designed for research use. The limited available evidence suggests that US firms spend on the order of \$100 billion a year on training, with expenditures reaching roughly 4–5 percent of labor costs in the 1990s (Black, Skipper, and Smith 2023); about half of US adults report employer-sponsored training in the past year, and nearly 80 percent report access to firm-provided learning resources (Sandvik et al. 2025).

However, there are several sources of survey data on which workers within an organization received on-the-job training. We draw upon several of these surveys to document four empirical patterns. First, participation in training rises with workers' education, suggesting complementarity between schooling and on-the-job learning. Second, firm characteristics – such as industry of activity, establishment size and learning culture – explain a large fraction of the variation in participation in formal and informal training activity across workers, even after controlling for individual and job characteristics. Third, better-managed firms invest more in training and do so in more sophisticated forms. Fourth, training strategies cluster: some firms emphasize managerial development and internal promotion ladders, while others focus on standardized frontline skills. These patterns place firm-level capabilities at the center of adult skill formation.

What makes some organizations more valuable training contexts than others? We argue that scarce managerial talent is the key driver of the variation in organizational design choices and training investments. In the “knowledge hierarchies” framework of Garicano (2000) and Garicano and Rossi-Hansberg (2006), higher-level managers spend their limited time solving the most complex problems and lower-level managers address more common tasks. In this way, the productivity of the top of the management hierarchy depends on the filtering and problem-solving capacity of the

middle and lower layers. In this framework, training not only increases worker productivity directly, but also frees time higher up in the hierarchy, increasing managerial effectiveness and/or spans of control (Espinosa and Stanton 2026). As a result, we expect firms with stronger managerial capability to invest more in training, and the level of training should cluster with other organizational practices.

This perspective has implications for wage inequality and corresponding public policy. If firms differ in their capacity to generate and deploy human capital, then early-career sorting into high- or low-capability firms can produce persistent wage divergence, even among observationally similar workers. Moreover, policies that target individual barriers to training—such as tuition subsidies to help overcome credit constraints—may have limited effects if firm capability is the binding constraint for deployment of learning programs for employees. Improving human capital accumulation at scale may require attention to management quality, organizational systems, and complementarities between training and production.

How does the incentive to hire and train junior workers change with recent advances in generative artificial intelligence? The job market for entry-level workers has been weak in the years since ChatGPT was released, with relatively greater impacts in AI-exposed occupations (for example, Brynjolfsson, Chandar, and Chen 2025). Some authors have concluded that AI represents seniority-biased technological change that disproportionately reduces demand for junior workers (Hosseini Maasoum and Lichtinger 2025). Garicano and Rayo (2026) develop a model that explores how AI changes training incentives in a knowledge hierarchy. In their model, AI disrupts the implicit bargain in which junior workers “pay” for training by providing rote, entry-level work. If AI is a better substitute for junior work, firms have less of an incentive to train. In our framework, the value of firm training also depends on organization-specific context, which could make training valuable even as AI capabilities advance.

Throughout this essay, we will use “training” and “workplace learning” somewhat interchangeably, but the distinction matters and deserves acknowledgment. Formal training programs with structured curricula and credentialing are only one mechanism through which firms build human capital. Peer learning, mentorship, job rotation, and the day-to-day accumulation of know-how through experience or trial and error are likely just as important, but available data rarely permit a clean separation. We expect formal training programs and the informal learning that surrounds them to be strongly complementary, particularly where firms structure formal programs

deliberately around who participates, whether it is mandatory, and how training connects to promotion.

Who Gets Training?

Becker's (1964) baseline model of human capital accumulation, when applied to competitive labor markets, suggests that any firm-sponsored general skills training is paid for in some way by the employee, most likely through lower wages. Yet there appears to only be a modest or weak relationship between training investments and lower wages in analyses dating back to data from 1992 in the United States (Barron, Berger, and Black 1999). More recent analyses of workers in Brazil and Italy show that employees of firms that generate high lifetime earnings, plausibly through training, do not start at lower wage levels (Arellano-Bover and Saltiel 2026). Instead, productivity gains that come from training partially accrue to the employer but also appear to benefit workers, raising the question of which workers receive these opportunities.

This section uses three complementary data sources to explore this question. The Programme for the International Assessment of Adult Competencies (PIAAC) identifies how training participation varies across workers and firms in 29 countries. The World Management Survey links training provision to management practices in a large sample of manufacturing firms. A 2023 survey of chief human resources officers at US firms lets us look at how training is organized inside firms.

PIAAC

The Programme for the International Assessment of Adult Competencies (PIAAC) is a cross-national representative survey of workers administered by the OECD that measures adult skills and collects details on training participation, education, labor market status, job characteristics, and workplace practices. Specifically, we use data from Cycle 2, fielded in 2022-2023. Our analysis follows the sample restrictions in Black, Skipper, and Smith (2023): ages 25-59, employees in the private or non-profit sector, working full-time (at least 30 hours per week). The US sample contains 1,083 respondents, of whom 1,068 have complete information on respondent and firm characteristics; the pooled cross-country sample varies between 15 countries (when using a full set of individual characteristics) and 29 countries (when using a smaller restricted set of variables that are available across the larger set of countries).

Who Gets Training? Age, Education, and Establishment Size (US)

Adjusted training rates -- formal (left) vs. informal (right). US, PIAAC Cycle 2.

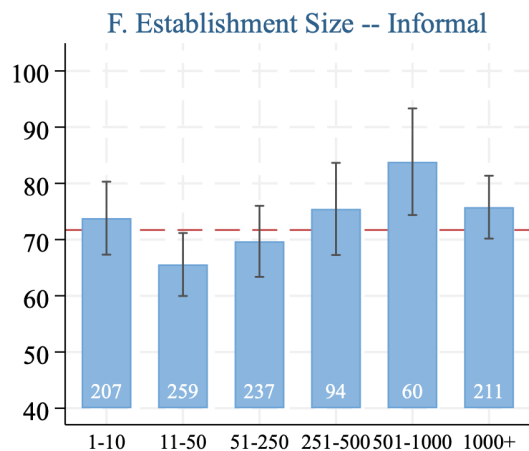
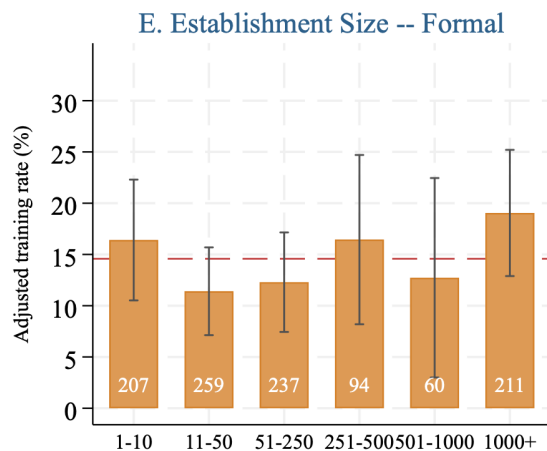
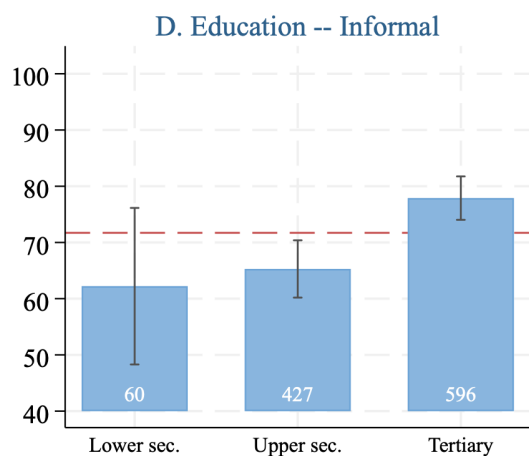
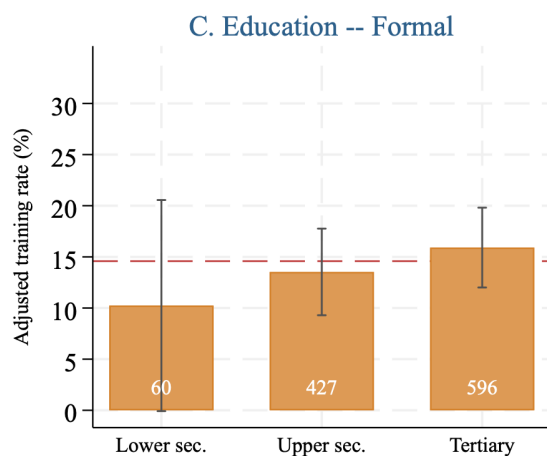
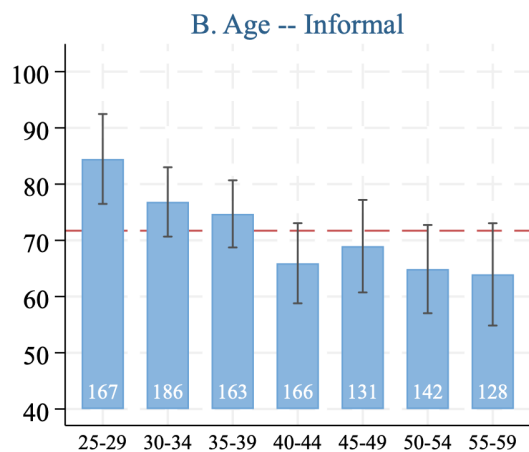
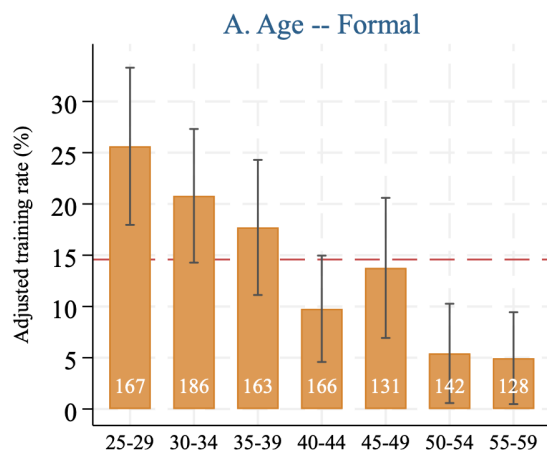


Figure 1: Correlates of training

Note: The figure plots adjusted training rates for formal and informal training by age, education, and establishment (workplace) size. The full specification controls for gender, age, education, dropout status, marital status, nativity, number of children, children under 6, occupation dummies, manager status, wage decile, numeracy, literacy, writing, ICT, teamwork, establishment size, part-of-larger-firm indicator, any workplace changes, industry dummies, learning at work, team voice, task discretion, and process improvement. The sample is PIAAC Cycle 2, United States (N = 1,083); bars show 95 percent confidence intervals from robust standard errors. The dashed line is the overall weighted training rate, and the numbers inside the bars are numbers of observations. Establishment size is the number of people at the respondent's work location.

Figure 1 shows participation rates separately for formal training (study toward a recognized qualification) and informal training (any organized non-formal training activity in the past 12 months, including open or distance courses, organized on-the-job training, workshops, and private lessons).¹ Across all age groups, 15 percent of workers participated in formal training during the last 12 months, while 72 percent participated in informal training. Formal training is largely a subset of informal training: of the 151 US workers in our sample who participated in formal training, 133 (88 percent) also participated in informal training, and only 18 did not. By contrast, informal training is often standalone: of the 789 workers who participated in informal training, 656 (83 percent) did not also participate in formal training.

We follow Black et al. (2023) and examine the correlates of training participation with cumulative linear probability models, adding four blocks of variables sequentially: (i) individual characteristics (gender, age, education, dropout, marital status, native-born, number of children, child under 6); (ii) occupation (measured by the ISCO 4-category from the International Standard Classification of Occupations) plus wage decile and manager status; (iii) skill use at work (standardized OECD-derived indices for numeracy, reading, writing, and information and communications technology) plus an indicator for team-based work; and (iv) firm characteristics, including structural features (size, organizational reach, recent workplace change, industry fixed effects) and proxies for organizational practices experienced at work (standardized indices of

¹ We follow common usage in calling this *informal* training, though PIAAC labels it *non-formal*; the truly unstructured category — informal learning in PIAAC's sense — is not the focus of this section.

learning at work, task discretion, team voice, and involvement in process improvement). The headline decomposition is estimated on the US sample.²

Both formal and informal training decline steeply with age. Workers aged 55-59 are about 21 percentage points less likely to participate in both formal and informal training than workers aged 25-29, although the proportional decline is much larger for formal training. Tertiary educational attainment is associated with 13 percentage point higher informal training participation, but its correlation with formal training is modest and statistically insignificant once the full set of controls is included. Both gradients are broadly consistent with human capital theory. Younger workers have longer expected horizons over which they (and their employer) can recoup investments in skills. More educated workers may both demand and generate greater returns to additional skill acquisition. Yet the education gradient also suggests complementarity, particularly for informal training: workers who arrive with stronger foundational skills appear to receive more, not less, of the on-the-job skill investment that informal training represents. This pattern has implications for inequality, since firm-provided informal training would then amplify rather than equalize the human-capital differences workers bring with them when they enter the firm.

Firm characteristics add explanatory power on top of individual and job controls. Industry controls are jointly significant. Controlling for industry fixed effects, occupation, wage decile, manager status, and standardized OECD-derived skill-use indices (that seek to measure numeracy, reading, writing, and information technology skills), the individual establishment-size effects are jointly insignificant for formal training ($F = 1.17$, $p = 0.32$); informal training indicators, however, are jointly significant ($F = 2.72$, $p = 0.02$), driven by a single sharp differential at the 501-1000 employee band. Yet focusing on establishment size (the measure available in the data) rather than firm size masks potential economies of scale in training delivery. What predicts training far more consistently is whether the local establishment is part of a larger organization: +5.2 percentage points ($p = 0.04$) for formal training, +8.3 percentage points ($p = 0.005$) for informal training. This pattern is consistent with scale economies in training delivery: larger organizations can amortize the fixed costs of curriculum design, deploy

² Analogous models on the cross-country pooled sample appear in the online Appendix. The cumulative R^2 across the four blocks lets us ask how much of the variation in training participation traces to characteristics of the worker versus characteristics of the firm.

training infrastructure across sites, and operate the more complex production processes that raise the value of worker skills, in the spirit of the knowledge hierarchy argument we develop in the next main section. The evidence in Figure 1 also complements quasi-experimental evidence that workers who begin their careers at large firms enjoy persistent, long-run advantages by age 35 (Arellano-Bover 2024).

World Management Survey

The World Management Survey is fielded through double-blind, semi-structured, open-ended telephone interviews with middle managers to measure the adoption of structured management practices across large samples of firms. Interviewers score firms on the adoption of 18 management practices related to operations, monitoring, target-setting, and people (talent) management on a 5 point scale, with higher scores denoting a higher and more systematic degree of adoption of the specific practice (Bloom and Van Reenen 2010).³ In 2022, the manufacturing survey was augmented with two questions asking 1,344 manufacturing firms across 11 countries about the type of training opportunities provided in the plant, distinguishing between no training, basic, and advanced (cross-functional) training to new and existing employees. Figure 2 examines how the provision of advanced training for new and existing employees varies with the management score, including controls for firm size and industry and country fixed effects. The patterns help to explain why, beyond establishment scale and organizational reach, some firms invest more in their workers than others.

³ On monitoring, for example, a low-scoring firm tracks a handful of measures that do not reveal whether its objectives are being met and reviews them only among senior staff; a high-scoring firm tracks a defined set of indicators continuously, makes them visible to the people doing the work, and reviews them on a regular schedule. On target-setting, a low score means that tracked goals are vague or purely financial, while high scores correspond to a balanced set of short- and long-run targets that individuals can connect to their own work. The individual scores are aggregated by simple averaging to obtain a single management score. Interviewers do not know the firm's performance and managers are not informed of the scoring scale.

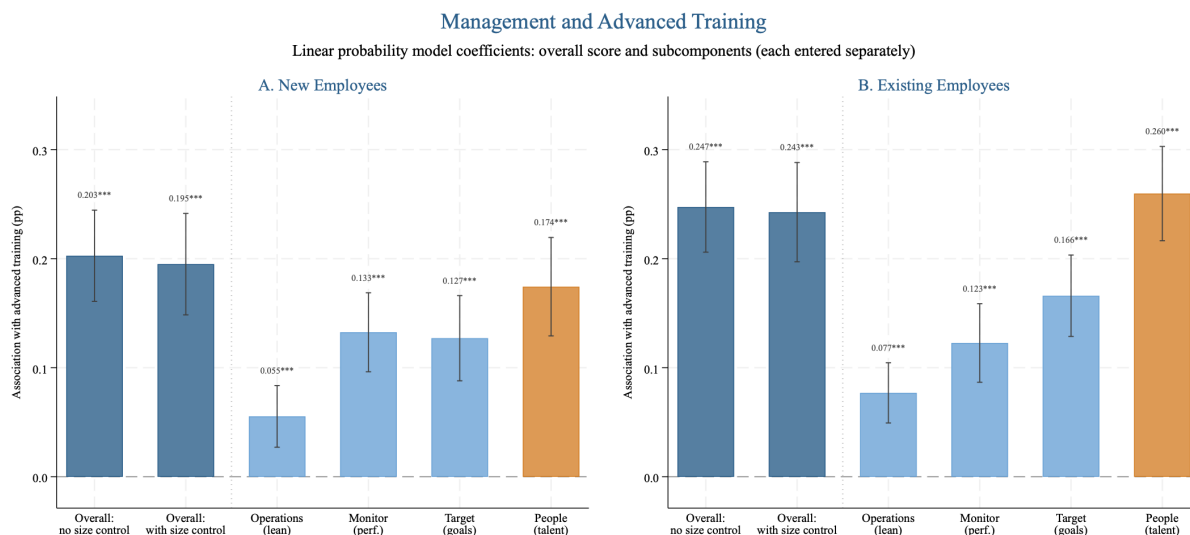


Figure 2: Management Practices and Advanced Training Provision

Note: The figure plots linear probability model coefficients on advanced training for overall management score and four subcomponents (each entered separately). Details available in the online Appendix. The figure is split by training for new and existing employees. The sample is the World Management Survey 2022 manufacturing wave ($N = 1,344$ firms per employee type); bars show 95 percent confidence intervals from robust standard errors. All models control for country and two-digit industry fixed effects (with missing industry as its own category) and, where indicated, the logs of firm and plant employment with a missing-value indicator.

The key result is that firms with greater adoption of structured management practices (the first two bars in each panel) are significantly more likely than others to provide advanced training, even after controlling for firm size and industry and country fixed effects.⁴ A one-point increase on the World Management Survey five-point overall management scale (roughly 1.7 standard deviations in this sample) is associated with a 19.5 percentage points (new employees) to 24.3 percentage points (existing employees) increase in the probability of providing advanced training. Decomposing the management index into its four subcomponents—operations, monitoring, targets, and people management—shows that people management, which captures the adoption of practices related to talent development, promotion, and performance evaluation, accounts for most of this association. In a specification that includes all four subcomponents jointly, a one-point increase in the people management score is associated with an 11-percentage point increase in advanced training for new

⁴ Size and productivity are positively correlated in the cross-section, but dynamic firms may have better management practices in place to prepare for growth.

employees and a 22-percentage point increase for existing employees, while the remaining subcomponents are weaker or statistically insignificant (monitoring retains a significant association for new employees). This pattern suggests that training is embedded in a broader system of people management processes. Firms that promote based on performance, evaluate talent rigorously, and address underperformance are also more likely to provide advanced training.

Figure 3 summarizes the explanatory power that firm-side variables add to standard worker and structural controls, using both the PIAAC and World Management Survey datasets. Panel A shows the PIAAC decomposition: adding the firm block (establishment size, organizational reach, industry fixed effects, and worker-reported organizational-practice proxies) on top of individual, occupation, and skill-use controls raises the adjusted R^2 from 6.8 to 10.7 percent for formal training (a 58 percent proportional gain) and from 24.3 to 28.1 percent for informal training (a 16 percent gain). The same qualitative ordering and approximate magnitudes appear in the great majority of OECD countries we examine.⁵ Panel B shows the World Management Survey analog: adding the overall management score on top of country, industry, and log-employment controls raises the adjusted R^2 for advanced training from 12.5 to 16.8 percent for new employees and from 9.7 to 16.7 percent for existing employees. Notably, management quality explains more of the training variation for *existing* employees than for *new* employees, consistent with the view that ongoing worker development requires deeper organizational capability.

⁵ In a specification covering 29 countries (which omits the OECD-derived skill-use indices and other individual variables that are not included uniformly in public-use files), the average gain in adjusted R -squared from adding firm and organizational-practice variables is roughly 11 percentage points for informal training, nearly an order of magnitude larger than what individual characteristics alone explain. Korea is the lone outlier, with very low explanatory power across all blocks. Details are available in the online Appendix.

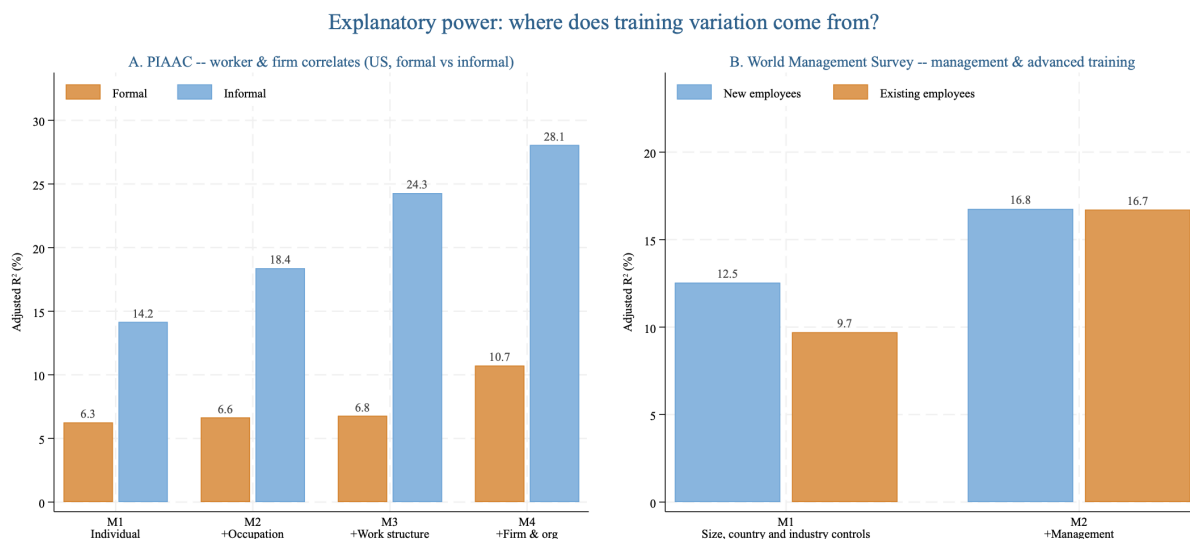


Figure 3: Explaining Training Participation and Provision

Note: The figure plots adjusted R-squared from cumulative four-block decompositions for the PIAAC sample in Panel A (Linear Probability Models, US sample) and the World Management Survey incremental contribution to adjusted R-squared from adding management practices in Panel B. All models within a panel are estimated on a constant sample (PIAAC United States, $N = 1,068$; World Management Survey, $N = 1,344$ firms per employee type); Panel B adds the overall management score to country, industry, and log-employment controls. Calculations are available in the online Appendix.

Although the two panels use different operationalizations of firm characteristics, they tell a coherent story: firm-level variation—and management quality in particular—adds meaningful explanatory power, even though most of the variation in training participation remains unexplained by observables in either dataset.

Where a person works and how their firm is organized are among the strongest predictors of whether they receive training. Although this evidence is only descriptive, it is consistent with quasi-experimental work that growth of wages and human capital depend to some extent on the firm in which a person ends up working. Recent examples include Arellano-Bover (2024), who uses labor-market entry conditions as a source of exogenous variation to identify firm effects on young workers' careers, and Arellano-Bover and Saltiel (2026), who use worker-mover designs (that is, shifts in wage growth as workers move, or do not move, between firms) to identify firm-specific human-capital growth. A growing body of evidence establishes that firm heterogeneity

is quantitatively important in explaining adult earnings inequality.⁶ Our results suggest that firms' investments in their workers' skills, and the way they organize those investments, are at least partially responsible.

Employer Reskilling and Upskilling Survey

To look more closely at what training bundles contain, and to examine how firms organize training itself, we turn to a 2023 survey of Chief Human Resources Officers and senior human resources leaders at 954 US firms with 100 or more employees (Doumi, Sadun, and Tamayo 2026). The survey was administered in July 2023 by Dynata, a professional survey provider, to one senior human resources leader per firm, with quotas on firm size, industry, and geography. The survey asks each respondent about a single flagship reskilling or upskilling program in detail (its participants, delivery model, requirements, funding, and incentives provided to trainees and their managers) together with the firm's broader human resources infrastructure. The 954 firms we analyze are those that report operating at least one upskilling or reskilling program.

Compared with samples used in the PIAAC and World Management Survey, the strength of these data is their granularity: they let us look beyond *whether* firms train and ask *how* they organize that investment. The trade-off is that the sample skews toward larger firms with some training in place. Hence, the data are best positioned to understand differences in how programs are targeted at different workers or designed and administered, rather than whether a firm trains at all.

We group firms into two clusters based on the characteristics of their flagship program, as shown in Figure 4. The groups come from a cluster analysis in which each firm is described by 86 program-design features covering delivery, participant selection, funding, incentives, assessment, and placement.⁷ The first cluster (about one-quarter of firms) runs programs that are local in scope and oriented toward frontline employees. They tend to take place at a single site, are typically voluntary, are most often funded entirely by the organization, and disproportionately oriented toward rank-and-file workers. The second cluster (about three-quarters of firms) runs programs that are

⁶ Estimates of this kind should be read with care. The firm effects recovered from two-way fixed-effects models in the tradition of Abowd, Kramarz, and Margolis (1999) also absorb sorting and match quality, so they are not a clean measure of a firm premium (Borovickova and Shimer 2025).

⁷ Because those features mix binary and continuous items, dissimilarity is measured with the Gower metric and firms are grouped by agglomerative hierarchical clustering; standard criteria for choosing the number of groups favor a two-group split. The online Appendix reports robustness checks, which suggest the two configurations are best read as poles of a continuum rather than sharply bounded types.

much more structured in delivery. Nine in ten of these programs run across multiple locations, relative to fewer than one in ten of the first cluster. Nearly all include on-the-job activities (98 against 75 percent), most culminate in a final assessment or certificate (93 against 77 percent), and over 40 percent are mandatory rather than voluntary (44 against 16 percent). They reach farther up the hierarchy as well: 22 percent explicitly include C-suite executives and 54 percent include business unit leaders, against 12 and 40 percent in the first cluster.



Figure 4: Firms Cluster in How They Deliver Training

Note: Survey of Chief Human Resources Officers and senior human resources leaders, fielded mid-2023, at U.S. firms with 100 or more employees. N = 954 firms with structured upskilling or reskilling programs (the cluster sample). Firms whose only training is general or job-induction are excluded. Cluster 1 contains 260 firms and Cluster 2 contains 694; three Cluster 1 firms did not answer the program-design items, so the bars use 257 and 694 responses. Bars show the share of firms in each cluster whose program has the listed feature, with 95 percent confidence intervals.

A core finding from this survey is that the organization of training is closely tied to a firm's formal human-capital infrastructure. Skill taxonomies, skill inventories, and internal learning academies are the strongest predictors of whether firms have more structured investments in training. Having skill taxonomies and skill inventories are associated with 26 and 17 percentage point increases respectively in the probability that

a firm runs a structured training program.⁸ Multinationals are 13 percentage points more likely than domestic firms to run such programs. Firms with 10,000 or more employees are 13 percentage points more likely than firms with 100 to 999 employees to have structured training programs. Formal human resources infrastructure, multinational scope, and scale together account for most of the explanatory variation; among them, the human resources practices are the most important.

Overall, the firms running structured, hierarchy-spanning programs are precisely the firms with the organizational capacity that can make such investments worthwhile. Conversely, the firms running localized, frontline-focused programs are those with less formal human resources infrastructure and narrower organizational scope. The organization of training, in other words, correlates with the broader organizational design of the firm.

Summary

The figures in this section show four patterns that motivate our discussion in the next section of the organizational economics of training. First, firm-provided training is increasing in workers' education, suggesting a complementarity between on the job learning and formal education and skilling. Second, firm characteristics add substantial explanatory power beyond worker observables, and this pattern generalizes across countries, labor market institutions, and industries. Third, better managed firms, especially those with stronger people management practices, are closely linked to advanced training provision. Fourth, firms differ not only in *whether* they train but also in *how* they organize training. Training appears as part of a broader portfolio of organizational design choices, in which formal human resources infrastructure, hierarchical reach, and program institutionalization move together.

Our results suggest that firms differ in management capability and organizational design, and that these differences shape the returns to skill within the hierarchy. Heterogeneity in training investment is therefore not merely a response to worker demand or labor market frictions, but an endogenous component of how firms produce and compete. In that view, adult skill formation is tightly related to both individuals accumulating human capital, and firms deciding when, how, and for whom additional skill is valuable.

⁸ In the Appendix, we estimate linear probability models that hold constant firm size, multinational status, public listing, union density, industry of activity, and the firm's stated reasons for investing in training (upskilling or reskilling).

In the next section, we sketch a framework to help understand what drives firms facing the same conditions to make different choices about how to organize. The literature on the determinants of management practices has traditionally focused on factors like labor market conditions or competition. But other factors emphasized in some papers, like family ownership, suggest that the identity of a top manager also tends to matter (Bloom, Sadun, and Van Reenen 2010). We argue that differences in top-manager skill can explain differences in organizational structure and people-management practices and, through these, the heterogeneity in training investment we have just documented.

Training as an Organizational Design Choice

Standard theories of firm training grapple with the question of why firms might perceive such training as in their interest, rather than just raising the wages of workers who may leave for an alternative employer. Becker argued that firms would only pay for firm-specific skill acquisition, not general training. Later theories emphasized labor market frictions. Monopsony and wage-compression models such as Acemoglu and Pischke (1999) predict more training when firms have greater wage-setting power. When search frictions or labor market power compress the wage schedule, a worker's wage rises by less than their productivity, enabling firms to invest in skills and capture value. However, these models do not explain why firms from the same industries and local labor markets would make different training investments. The skill-weights framework developed by Lazear (2009) allows firms to value different bundles of general skills, making those bundles of skills effectively firm-specific, but takes skill weights for a firm as given rather than explaining their origin. Models of asymmetric information about worker productivity explain how firms capture rents on training investment that are unobserved by outside employers, but also do not explain heterogeneous firm investments.

More fundamentally, we argue that this framing of the firm's decision to offer training elides a basic point. Firm training should be viewed as an endogenous design choice. Organizational differences necessitate different training investments, which in turn depend on managerial capabilities. Firms that train prepare workers for organizational complexity; firms that don't train do less customization, enabling the training firms to capture value.

Knowledge Hierarchies

We build on a long tradition of models in organizational economics, particularly the “knowledge hierarchy” model of Garicano (2000) and the extensions by Caliendo and Rossi-Hansberg (2012) and others. To fix ideas, consider the knowledge hierarchy of a management consulting firm. When a client engagement begins, the day-to-day analytical work lands first with junior consultants. They conduct interviews, build spreadsheets, and structure analysis. Most problems that arise are well within their competence. But when junior consultants encounter an issue that exceeds their skill, they ask a senior consultant. If the problem is harder still, the senior consultant raises the question with the partner who owns the relationship. In this knowledge hierarchy, the partner handles only the most difficult problems. The defining feature of a knowledge hierarchy is that problems are indexed by difficulty, and only the higher levels of a pyramidal organization can solve the more difficult problems.

A knowledge hierarchy is an efficient way to deploy the partner's time and judgment, which is the firm's scarcest labor resource. The upward filtering of harder problems that happens via the knowledge hierarchy enables managerial leverage. A senior consultant who can resolve 70 percent of the problems that arrive passes only 30 percent upward; one who can resolve 60 percent passes 40 percent, increasing the burden on the next layer. Worker training that improves autonomous problem resolution creates efficiencies in upper layers, because fewer problems arrive there. This insight is not unique to consulting and other professional services. Technology firms, government agencies, and manufacturing firms (which Caliendo, Monte, and Rossi-Hansberg (2015) study through the lens of hierarchies) are organized in layers for many of the same reasons. The conditions hold whenever exceptions are escalated, like when a nurse asks an attending physician for help on a complicated patient or when a loan officer refers an anomalous balance sheet to a supervisor. The value of the time savings from training a lower-level worker simply depends on the expensiveness of the upper layer, which in turn depends on the skill of the people there and, ultimately, on the skill of whoever sits at the top.

Empirical work is consistent with the importance of this kind of filtering mechanism. For example, when a Colombian government organization trained lower-level workers, the volume of help requests filtering up through the hierarchy fell, allowing managers to shift their time from supporting existing workers to other tasks (Espinosa and Stanton 2026). However, we have not yet explained why managers with different abilities make different organizational choices.

How Manager Ability Shapes the Firm

To explore why managers build different hierarchies, we examine how top managers with different abilities organize their firms. We make one key change to the canonical knowledge hierarchies framework, which we lay out formally in Deming, Gilman, Sadun, and Stanton (2026). Our illustration builds on Garicano and Rossi-Hansberg (2006), who show that the most skilled workers in the economy will start firms when costs of knowledge acquisition vary between individuals. To build intuition, we focus on a stylized industry equilibrium where two entrepreneurs with the highest skills in the economy start firms. We set these workers' skills at a high level, while assuming that a fixed cost of entry prevents a third entrepreneur with lower skills from forming a third, organized firm. However, workers can always produce as individuals, selling their own solutions. Other than the candidate entrepreneurs, everyone else in the economy is ex-ante identical, which will potentially give rise to either firms or workers paying for training.

The key difference from the literature involves wage determination. In our model, workers' pay is based on their outside option. Firms have a training incentive if melding a worker's skill to fit the organization's needs has a higher payoff than if the worker went to a rival or if the worker produced alone as a sole proprietor.⁹ Firms with simpler organizations are less likely to have an idiosyncratic match-specific fit that beats a rival firm and a worker's option to produce independently, explaining why competing firms make different training choices.

For concreteness, consider two consulting firms operating in the same city, competing for the same clients and to hire the same staff. The more capable partner — who we will call the founder of Firm 1 — can solve a wider range of problems than the founder of Firm 2. However, if the founder of Firm 1 is constantly interrupted by problems that a strong senior consultant could resolve, the firm is wasting its advantage. The firm with the more capable manager has more to gain from building a strong middle layer, so that fewer problems filter upward. Similarly, a more capable manager can scale personal influence with a larger frontline workforce, but only if the intermediate layers beneath can handle the increased volume. Firms with better top managers have a greater incentive to build out structure, add middle layers, and adopt

⁹ This follows a tradition in labor economics from Postel-Vinay and Robin (2002) that treats the wage determination process like an auction. As far as we know, papers in the knowledge hierarchies literature have used different wage determination processes. Our treatment differs from the model in Postel-Vinay and Robin, which features a competence parameter on the worker side, as well as recent contributions on worker and firm dynamics by Gualdani et al. (2026) and Di Addario et al. (2023), by focusing on firms' investment incentives in light of the wage setting mechanism.

the talent management practices that empirically correlate with deep training investments.

The key result is a complementarity between managerial talent and organizational depth: the more capable the leader sitting atop the firm, the more layers it builds. Figure 5 considers a benchmark numerical example to illustrate the relationship between the number of layers in an organization, top manager talent, and training. We consider a mass of 100 workers that can be allocated between the firms or solo operations, but we have stripped away all product-side differentiation. Firms or sole proprietors compete based simply on the number of solutions to problems they can solve, and we look for a Cournot equilibrium in quantities where firms choose how to organize (worker skills, numbers and layers) and sole proprietors make their own skill investments. Frontline workers and sole proprietors address problems that vary in difficulty, illustrated by the density curve in Figure 5, which we parameterize as draws from an exponential distribution with parameter $\lambda = 0.9$. The respective CEO skills, 3.8 and 3.0, imply that the leader of Firm 1 can solve $1 - \exp(-0.9 \times 3.8) = 97$ percent of problems, while the leader of Firm 2 can solve 93 percent. Sole proprietors and frontline workers have skills of 1.2, enabling them to solve 66 percent of problems they draw.

The relatively greater ability to solve exceptional problems leads to very different organizations. Firm 1's more capable CEO operates at a much larger scale, with about 49 frontline workers, compared to Firm 2, which hires only 12 frontline workers. The more capable entrepreneur (Firm 1) builds a deeper hierarchy with a middle layer of managers above frontline workers, while the less capable entrepreneur (Firm 2) operates a flatter structure with only frontline workers under the entrepreneur. By training a relatively small middle layer (here 4.2 workers with skill 2.8) to filter problems beyond frontline workers' capacity, Firm 1 can support more frontline workers while the CEO's time constraint just binds due to problems that require ability beyond the middle layer's skills. Ultimately, Firm 1 produces a quantity of 47.8 problems that is more than four times as large as Firm 2 because of the complementarity between organization and managerial skill. Each of these choices is stable for the firms. Adding a fourth layer to Firm 1 and adjusting the skill mix and numbers across layers would add more payroll and training expenses, reducing profit to 8.15 from 9.28 with three layers. If Firm 2 were to add a third layer and re-optimize, profit would fall from 3.96 to 3.77.

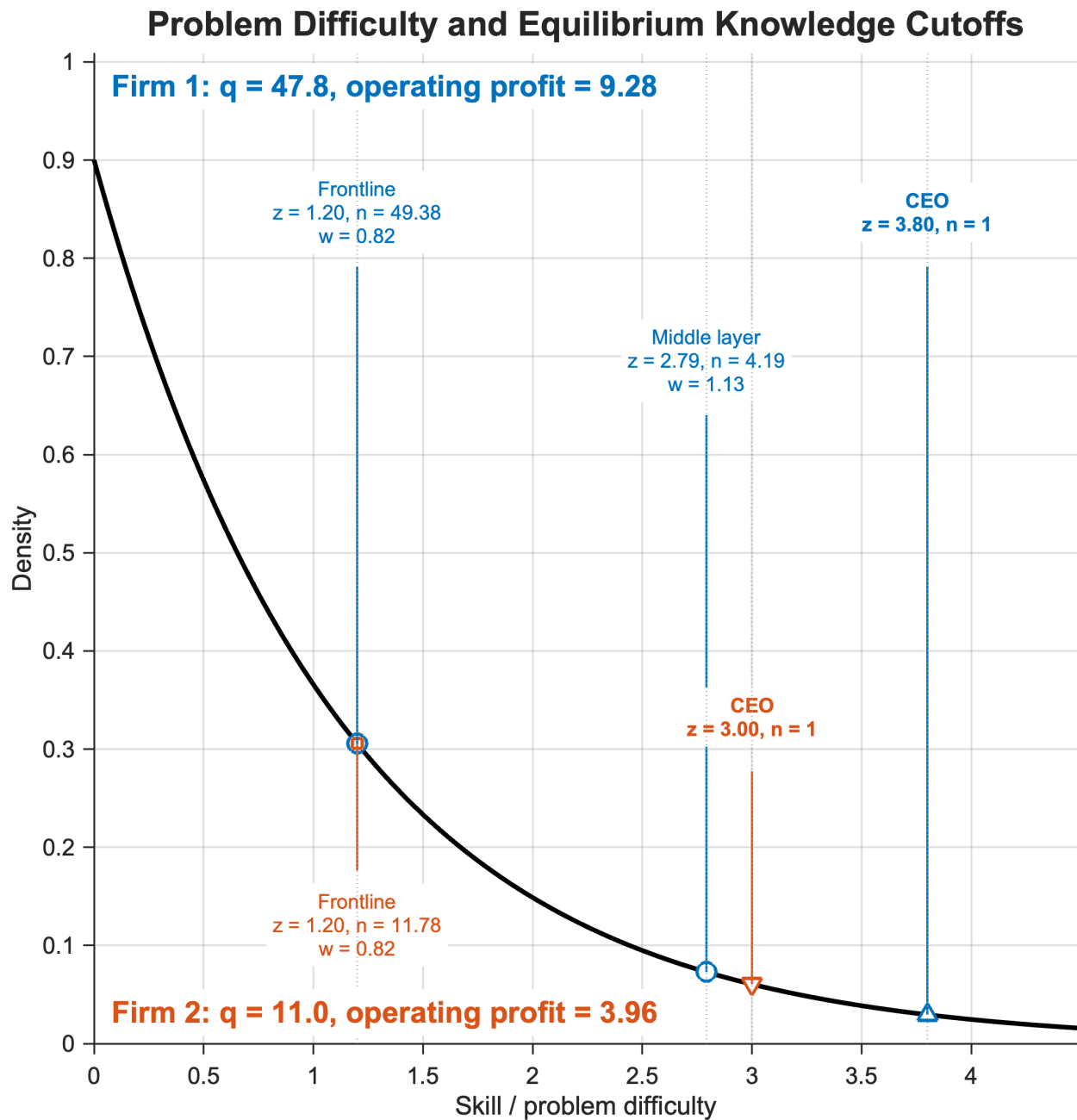


Figure 5: Optimal hierarchy structure for two firms with different entrepreneur skill levels.

Note: The figure plots equilibrium organizational structure for Firm 1 and Firm 2 against the density of problems in the economy. Skills are represented as segments of the real line and workers with skill z can solve any problem that has a difficulty index less than their skill. Workers are born with skill equal to 1.2 such that schooling prepares workers with exactly what is required to be a sole proprietor, while further training costs to acquire new skills are proportional to the interval length of the change in skill. There is a mass of 100 workers that can be hired by firms or work alone. Firms and individuals compete in quantities, taking into

account a linear, downward sloping industry demand curve. Labels in blue show the skill levels (z), number of workers (n), and wages (w) for workers at different layers for Firm 1. Labels in orange show the same thing for Firm 2. Firm 1's quantity is 47.8 and operating profits net of training costs are 9.28. Firm 2's quantity is 11 and operating profits are 3.96.

This framework illustrates the endogenous relationship between managerial capability and organization design. A strong senior consultant is worth more as a buffer beneath a highly capable partner than beneath a less capable one, simply because the partner's time is more valuable. This result holds with a simple conception of “skill” which is generally valued in the labor market: the worker's cumulative knowledge for handling consulting tasks of varying degrees of complexity.

Wages, Training Incentives, and the Optimal Training Investment

Before turning to the training decision itself, we need to ask whether firms have any incentive to train in the first place or whether trained workers will be poached by other firms or, alternatively, leave to set up their own shops.

The allocation in Figure 5 is stable because the complementarity between worker skill and organizational structure makes a trained worker more valuable inside Firm 1 than in any other firm or working on their own. Firm 2 might want to poach the trained employee from Firm 1, but it is only willing to pay for the skill it can actually use. However, the training in Firm 1 was calibrated to the needs of a hierarchy built beneath a skilled partner, and those needs are not replicated in Firm 2's shallower structure.¹⁰ The trained middle layer workers' wages of 1.13 reflect what they would make as a sole proprietor with these skills, but they do earn a premium relative to the wage of 0.82 that they would make without the firm's training investment. Yet the worker is more valuable inside Firm 1 than outside, because middle-level workers can handle more problems than if they had to work a problem end-to-end. Firm 1 therefore earns rents on its training investment—not because the labor market is imperfect in any standard sense, but because the organizational context itself creates value that is not portable. The labor market is effectively specific, even though the skills are not. In this example,

¹⁰ The reason a trained worker can be worth more inside the firm than as an independent provider is due to leverage: a trained worker can support several workers in the hierarchy below, which can be more valuable than solving a problem on her own. In knowledge hierarchies models, leverage is determined by a helping cost parameter, which represents the fraction of a time period that a higher-layer worker spends addressing a problem that gets passed up. Lower-layer workers and sole proprietors spend their time on a single problem. We set the helping cost parameter, h , to 0.25, following Garicano (2000).

skills are entirely one dimensional and general; the weights on skills are not combinations of different skills, but instead arise because some firms are willing to pay to make or buy a more capable generalist.

Given that the firm has an incentive to train, how does it decide where to invest? Consider Firm 1's decision about whether to train its senior consultants – the middle layer of its hierarchy. Every senior consultant who receives training resolves more of the problems escalated to the partner by junior consultants, so fewer problems reach the partner's desk and the partner can handle more of them, raising firm productivity. The firm trains when the gains from protecting the partner's time exceed the cost of raising the skills of every person in that layer.

The same logic applies one layer down, but the tradeoffs change. Training the entire group of frontline workers pays off in a smaller layer above – the savings accrue to the senior layer rather than directly to the partner. Whether such is worthwhile for the firm depends on how expensive the senior layer is and how many juniors each senior oversees. In Firm 1, where the senior layer is more skilled and the ratios steeper, the firm still concentrates its investment in the upper layers where the bottleneck is most binding. This insight mirrors the clustering of training strategies in the survey of chief human resources officers: some firms train mainly managers, others focus on frontline workers, and the pattern reflects organizational structure rather than arbitrary preference.

Two further implications follow. First, firms with more capable leadership invest more in training overall and capture larger rents from doing so – not because workers are stuck, but because the trained worker is genuinely worth more inside a hierarchy with more capable leadership. Firm 1's trained middle-layer workers earn higher wages than they could earn in Firm 2 or on their own; the organizational complementarity benefits both sides because these workers' skills expand after firm investment. Second, a survey that records only whether a firm has a training program will miss the most important variation: how much it invests, in which layers, and why.

What the Framework Explains and What it Misses

In this framework, differences in firm training investments are driven by variation in managerial capability and endogenous organizational design choices. The framework also explains the long-standing empirical puzzle that firms often pay for their workers to receive general skills training and that these workers generally stay at the firm that trained them. In addition, there is a sharp prediction that training

investments often matter most in the middle and upper layers of the management hierarchy.

In standard monopsony or search-friction models, firm rents from training arise because the market works imperfectly – workers cannot readily or at low cost find alternative employers who would pay them their true value. In contrast, in this hierarchy model, rents arise because the organizational context itself creates value that is not fully portable. The thinness of the post-training labor market is endogenous to the organizational structure. This distinction has important policy implications. If training incentives arise from market imperfections, then competition policy and labor market policies that encourage training may be at odds. On the other hand, if training arises from organizational heterogeneity, public policies to promote worker training that ignore firm heterogeneity and organizational complementarities may see very little returns from generic training investments.

More broadly, firms that invest more in training also create richer internal labor markets, with more rungs on the promotion ladder and faster potential wage growth for workers who enter the firm. This finding is consistent with natural experiments where early-career sorting into high-wage-growth firms generates persistent wage differences (for example, Arellano-Bover and Saltiel 2026). Better-managed firms probably also invest more in leadership development. The people-management practices captured in the World Management Survey include monetary and non-monetary rewards for high performance and opportunities for internal promotion, both of which reflect leadership development investments.

What happens to firm training if AI alters incentives to hire and train junior workers? Ide and Talamas (2025) develop a knowledge hierarchies model where AI agents can act either as copilots or autonomous coworkers. Copilot AI reduces inequality, while autonomous AI increases it. Garicano and Rayo (2026) show that AI can blunt training incentives by substituting for the rote work that formerly paid for the training of junior workers. In their model, the impact of AI depends on how much it raises the minimum skill at which an entry-level worker becomes useful (the “entry floor”) versus increasing the value of a fully-trained worker (the “expert ceiling”). In our framework, the impact of AI also depends on the organizational surplus created by embedding a trained worker beneath scarce managerial talent.

This distinction generates firm heterogeneity in the response to increasing AI capabilities. Firms with simple organizations will see the training period compress and eventually disappear, as in Garicano and Rayo (2026). If training takes time and

exposure, rather than the instantaneous investment our static framework considers, firms with talented managers and strong organizational complementarities may continue to train even after junior production ceases to finance the investment. But training in these firms may change, with fewer entrants, a higher “entry floor”, more intensive and formal instruction, and greater concentration on the middle and upper layers of the hierarchy.

One limitation of our framework is that it treats these internal labor markets as arising from managerial quality differences alone, but there may also be economies of scale and other sources of comparative advantage in training provision that shape which firms invest in workforce development. Some organizations may simply be better than others at designing, organizing, and delivering training programs at low cost. If on-the-job experience and formal learning are complements, then firms may need to participate actively in skill formation even when the underlying skills are portable across employers. In that environment, firms with superior organizational capabilities in training provision may offer development opportunities as a workplace amenity that helps attract and retain workers (Maestas et al. 2023).

Our framework is also not intended to provide a comprehensive explanation for all forms of firm-sponsored training. Instead, it is designed primarily to illustrate “what is in it for firms” when they invest in worker skill. Other mechanisms may support employer-sponsored learning even absent strong organizational complementarities. What is important is the interaction between internal structure and outside wages or opportunities, a logic that also underlies why firms train more when they have non-compete agreements at their disposal (Starr 2019) or when contracts with repayment provisions are available (Hoffman and Burks 2020). These arrangements effectively allow firms to finance general skill acquisition while partially protecting themselves against immediate worker turnover.

Finally, the basic hierarchy framework intentionally abstracts from worker heterogeneity, treating workers as homogeneous before training and workplace learning occur. Relaxing that assumption would likely generate sorting and matching patterns that interact importantly with firms’ training incentives. We conjecture that models with heterogeneous workers and heterogeneous firms would generate positive assortative matching between highly educated or highly skilled workers and firms with the greatest skill demands and strongest incentives to invest in training. In such environments, firms with the highest managerial capability or organizational complexity may both select and develop workers with the greatest learning potential,

amplifying observed differences in career trajectories across firms and helping to explain the positive relationship between education and on-the-job training.

Conclusion

Firm-sponsored general skills training is very common. Much of the literature has explained this phenomenon through frictional labor markets. For example, in studying how Indian garment firms can support the provision of social skills training, Adhvaryu et al. (2023) find extremely high returns on investment. They point to labor market imperfections to justify minimal effects on wages or retention. Yet the garment sector in India has many firms, raising a puzzle about how training can operate in sectors with many firms, each of which makes different training investments.

In this paper, we have suggested an alternative form of labor market friction that stems from organizational design. In the social skills context, the gains from training appear to come from improved collaboration and team effectiveness, alleviating the need for managerial attention. If all firms had the same organization, we would expect these sources of advantage to dissipate. But instead, heterogeneous organization generates different skill demands from workers – whether in the Indian garment industry or elsewhere--which make general skills effectively firm specific and justify firms' investment in training.

References

- Abowd, John M., Francis Kramarz, and David N. Margolis. "High wage workers and high wage firms." *Econometrica* 67, no. 2 (1999): 251-333.
- Acemoglu, Daron, and Jörn-Steffen Pischke. "The structure of wages and investment in general training." *Journal of Political Economy* 107, no. 3 (1999): 539-572.
- Adda, Jérôme, and Christian Dustmann. "Sources of wage growth." *Journal of Political Economy* 131, no. 2 (2023): 456-503.
- Adhvaryu, Achyuta, Namrata Kala, and Anant Nyshadham. "Returns to on-the-job soft skills training." *Journal of Political Economy* 131, no. 8 (2023): 2165-2208.
- Arellano-Bover, Jaime. "Career consequences of firm heterogeneity for young workers: First job and firm size." *Journal of Labor Economics* 42, no. 2 (2024): 549-589.
- Arellano-Bover, Jaime, and Fernando Saltiel. "Differences in on-the-job learning across firms." *Journal of Labor Economics* 44, no. 1 (2026): 149-188.
- Barron, John M., Mark C. Berger, and Dan A. Black. "Do workers pay for on-the-job training?." *Journal of Human Resources* (1999): 235-252.
- Becker, Gary (1964) *Human Capital*. Chicago: The University of Chicago Press.
- Black, Dan A., Lars Skipper, and Jeffrey A. Smith. "Firm training." *Handbook of the Economics of Education* 7 (2023): 287-468.
- Bloom, Nicholas, Raffaella Sadun, and John Van Reenen. "Recent advances in the empirics of organizational economics." *Annual Review of Economics* 2, no. 1 (2010): 105-137.
- Bloom, Nicholas, and John Van Reenen. "Why Do Management Practices Differ across Firms and Countries?" *Journal of Economic Perspectives* 24, no. 1 (2010): 203-224.
- Borovickova, Katarina, and Robert Shimer. "Assortative Matching and Wages: The Role of Selection." NBER Working Paper 33184 (2025).
- Brynjolfsson, Erik, Bharat Chandar, and Ruyu Chen. "Canaries in the coal mine?: Six facts about the recent employment effects of artificial intelligence." Stanford Digital Economy Lab Working Paper (2025).

Caliendo, Lorenzo, Ferdinando Monte, and Esteban Rossi-Hansberg. "The anatomy of French production hierarchies." *Journal of Political Economy* 123, no. 4 (2015): 809-852.

Caliendo, Lorenzo, and Esteban Rossi-Hansberg. "The impact of trade on organization and productivity." *The Quarterly Journal of Economics* 127, no. 3 (2012): 1393-1467.

Deming, David J. "Why Do Wages Grow Faster for Educated Workers?" , *Journal of Labor Economics*, forthcoming (2025).

Deming, David J., Julia Gilman, Raffaella Sadun, and Christopher Stanton. "Organization as a Rationale for the Provision of General Skills Training." Working paper (2026).

Di Addario, Sabrina, Patrick Kline, Raffaele Saggio, and Mikkel Sølvsten. "It ain't where you're from, it's where you're at: hiring origins, firm heterogeneity, and wages." *Journal of Econometrics* 233, no. 2 (2023): 340-374.

Doumi, Leila, Raffaella Sadun, and Jorge Tamayo. "How do firms train their workforce in the age of AI? Evidence from a survey of medium and large U.S. firms." Working paper (2026).

Espinosa, Miguel, and Christopher Stanton. "Training, communications patterns, and spillovers inside organizations." *Journal of Political Economy* 134, no. 7 (2026).

Garicano, Luis. "Hierarchies and the Organization of Knowledge in Production." *Journal of Political Economy* 108, no. 5 (2000): 874-904.

Garicano, Luis, and Luis Rayo. "Training in the age of AI: A theory of career viability." *CEPR Working Paper*, DP20634 (2026).

Garicano, Luis, and Esteban Rossi-Hansberg. "Organization and inequality in a knowledge economy." *The Quarterly Journal of Economics* 121, no. 4 (2006): 1383-1435.

Gualdani, Cristina, Elena Pastorino, Áureo de Paula, and Sergio Salgado. "Identifying Uncertainty, Learning about Productivity, and Human Capital Acquisition: A Reassessment of Labor Market Sorting and Firm Monopsony Power." NBER Working Paper 34973 (2026).

Gibbons, Robert, and Michael Waldman. "Enriching a theory of wage and promotion dynamics inside firms." *Journal of Labor Economics* 24, no. 1 (2006): 59-107.

Hoffman, Mitchell, and Stephen V. Burks. "Worker overconfidence: Field evidence and implications for employee turnover and firm profits." *Quantitative Economics* 11, no. 1 (2020): 315-348.

Hosseini Maasoum, Seyed Mahdi, and Guy Lichtinger. "Generative AI as seniority-biased technological change: Evidence from US résumé and job posting data." *Available at SSRN* 5425555 (2025).

Ide, Enrique, and Eduard Talamas. "Artificial intelligence in the knowledge economy." *Journal of Political Economy* 133, no. 12 (2025).

Lazear, Edward P. "Firm-specific human capital: A skill-weights approach." *Journal of Political Economy* 117, no. 5 (2009): 914-940.

Maestas, Nicole, Kathleen J. Mullen, David Powell, Till Von Wachter, and Jeffrey B. Wenger. "The value of working conditions in the United States and implications for the structure of wages." *American Economic Review* 113, no. 7 (2023): 2007-2047.

Postel-Vinay, Fabien, and Jean-Marc Robin. "Equilibrium wage dispersion with worker and employer heterogeneity." *Econometrica* 70, no. 6 (2002): 2295-2350.

Sandvik, Jason, Richard Saouma, Nathan Seegert, and Christopher Stanton. "Should Human Capital Development Programs be Mandatory or Voluntary? Evidence from a Field Experiment on Mentorship." *Management Science* (2025).

Starr, Evan. "Consider this: Training, wages, and the enforceability of covenants not to compete." *ILR Review* 72, no. 4 (2019): 783-817.