

The Firm as a Labor-Allocation System

Frameworks, Facts, and Measurement

Organizations, Personnel, and Internal Labor Markets — Week 1

Sai Zhang

What becomes visible inside the firm?

Central question

What do we learn by observing workers, jobs, managers, and teams, rather than treating the firm as a single production unit?

- Two workers with the same education and occupation can face different tasks, customers, supervisors, and career opportunities.
- Equal output can mask unequal effort costs, task difficulty, or support.
- Ask which organizational decision generates the difference and what would change under another rule.

Core points: rules shape what workers can do

- The same workforce can produce different outcomes under different assignments, incentives, information, and team structures.
- A productivity change can reflect effort, selection, reassignment, learning, technology, or measurement.
- Wages, ratings, physical output, and worker welfare are distinct objects.
- Rich records improve measurement; causal claims still need credible variation.
- A study can reveal a general mechanism without establishing a universal magnitude.

Follow the organizational decision

Organizational decision → worker response
→ observable outcome → credible comparison
→ firm and worker incidence

- **Bridge:** connect labor supply, human capital, search, and wages to decisions within employment relationships.
- **Field Core:** distinguish mechanisms and map each to records and variation.
- **Research Lab:** audit outcomes, reconstruct a statistic, and diagnose competing explanations.

Opening maps: Lazear–Shaw (2007); Hoffman–Stanton (2025, published handbook chapter); Benson–Shaw (2025).

Choose the unit before choosing the model

Unit	What it can reveal	What it may hide
Worker	Pay, spells, learning, careers	Assignment and team production
Job or task	Difficulty, output, requirements	Who receives each task
Manager	Authority, teaching, evaluation	Endogenous worker assignment
Team	Coordination and peer exposure	Individual contributions
Establishment	Local production and practices	Employer-wide rules
Firm	Pay-setting and organizational policy	Internal heterogeneity

A legal employer, an establishment, and a team are different identifiers. A worker-month and a worker-task-day answer different questions.

Similar outcomes can arise through different mechanisms

Mechanism	Organizational margin to observe
Selection	Who applies, is hired, stays, or leaves?
Incentives	How does a reward change effort or task substitution?
Assignment	Who does which job, with what opportunity cost?
Information	What is learned, recorded, or withheld?
Management	Who teaches, delegates, evaluates, or blocks a move?
Coordination	Whose work depends on whom, and when?
Careers	Who receives training, promotion, and future opportunities?
Surplus division	Who receives earnings, amenities, or profits?

Map the internal labor market before explaining it

- Baker–Gibbs–Holmström use personnel records to reveal hierarchy, transitions, promotions, and pay variation.
- A job-level map and transition matrix are substantive economic evidence before a causal coefficient is estimated.
- Description alone does not show whether the hierarchy is efficient or discriminatory.

Measurement exercise

An employee becomes a “senior analyst.” What records distinguish a promotion, a change in tasks, and an administrative relabeling?

Baker, Gibbs, and Holmström (1994), *The Internal Economics of the Firm: Evidence from Personnel Data*.

Organizational rules define the feasible responses

$$\Omega_f = (\mathcal{S}_f, a_f, b_f, \mathcal{M}_f, G_f, \mathcal{H}_f)$$

- $\mathcal{S}_f$: workforce selected
- a_f : worker–job assignments
- b_f : pay rules
- $\mathcal{M}_f$: management and authority
- G_f : collaboration structure
- $\mathcal{H}_f$: training

A pedagogical framework

Changing one rule can change who works, what workers do, what they learn, and whether they remain. Observed organizations need not choose globally optimal rules.

Production depends on the worker and the organization

$$y_{it} = f_{k(i,t)}(\theta_i, h_{it}, e_{it}, M_{m(i,t),t}, G_{it}; \Omega_f) + \varepsilon_{it}$$

- i, t, k, m : worker, time, job/task, and manager.
- θ_i : persistent skill; h_{it} : accumulated knowledge; e_{it} : effort.
- M : manager inputs; G : team environment; ε : other shocks.
- A manager may change assignment or information, rather than simply add a constant to output.

Empirical implication

A worker residual can contain unobserved task difficulty, managerial support, and coordination. It is not automatically worker ability.

The firm chooses rules anticipating worker responses

$$\max_{\Omega_f} \mathbb{E} \left[R_f(\{y_{it}\}, \Omega_f) - \sum_i w_{it} - C_f(\Omega_f) - K_f^{\text{turnover}} \right]$$

- R_f : revenue; w_{it} : compensation; C_f : costs of organizational rules; K_f^{turnover} : turnover costs.
- Staffing, production, contracts, and legal constraints restrict choices.
- Limited information, coordination costs, and manager self-interest can prevent the firm from reaching this benchmark.
- Higher profit does not establish higher worker welfare.

Teaching objective, not an estimated model or a welfare theorem.

Workers value a bundle and a future

$$U_{it} = \mathbb{E}[u(w_{it})] - c_i(e_{it}) \\ + A_i(\text{schedule, safety, autonomy, treatment}) + \delta \mathbb{E}[V_{i,t+1}].$$

- u : utility of pay; c_i : effort cost; A_i : amenities; δ : discount factor.
- $V_{i,t+1}$ includes learning, promotion prospects, and future outside options.
- A pay rise, better supervisor, and task transfer can work through different channels.
- Mas–Pallais use choices over job offers to value work arrangements; a single wage comparison cannot do the same job.

Mas and Pallais (2017), *Valuing Alternative Work Arrangements*.

The best worker need not be best assigned to Job 1

Worker	Job 1	Job 2
A	12	11
B	10	3

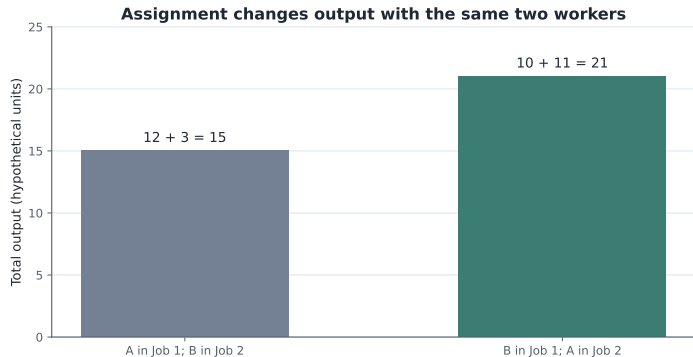
- A has an absolute advantage in both jobs.
- A's advantage over B is much larger in Job 2.
- Assigning A to Job 1 uses up the worker who is especially valuable in Job 2.

Predict before calculating

Which assignment maximizes total output when each worker fills one job?

Hypothetical output matrix: constructed teaching example, not empirical estimates.

Assignment changes total output from 15 to 21



Constructed teaching example — not an empirical estimate.

$$12 + 3 = 15$$

$$10 + 11 = 21$$

Same workers.

Same job capacities.

Different assignment.

Hypothetical teaching example. Figure generated from the displayed output matrix.

Current performance does not reveal every potential match

$$\max_{a_{ik} \in \{0,1\}} \sum_i \sum_k a_{ik} v_{ik}, \quad \sum_k a_{ik} \leq 1, \quad \sum_i a_{ik} \leq n_k.$$

- a_{ik} assigns worker i to job k ; v_{ik} is match value; n_k is capacity.
- We usually observe outcomes only in assigned jobs, not every v_{ik} .
- Benson–Li–Shue distinguish sales performance from subsequent managerial performance.
- Promotion can both allocate talent and motivate workers. A tradeoff does not by itself imply irrational decisions.

Benson, Li, and Shue (2019), *Promotions and the Peter Principle*.

A team's objective can obstruct firm-wide allocation

- A manager may want to retain a strong worker even when another role would use that worker better.
- Haegele studies talent hoarding through personnel records and surveys, linking manager incentives to internal applications.
- Observe available jobs, eligibility, applications, manager actions, and transfers; successful moves alone miss blocked opportunities.
- The application margin is one part of internal allocation. Week 6 develops the design and career consequences.

Distinguish the decision-makers

What benefits the current team, the organization, and the worker may differ.

Haegele (2026), *Talent Hoarding in Organizations*, published AER version.

A higher output mean leaves several explanations open

Candidate mechanism	Additional record or comparison
Incumbent effort changes	Repeated workers, tasks, hours, and quality
Worker selection changes	Entrants, leavers, applicant pool, baseline types
Tasks are reassigned	Task difficulty and assignment histories
Recording changes	Metric definitions, coverage, dashboard rollout
Learning or technology changes	Experience, training, tool access and use

An increase after a bonus reform is compatible with each mechanism. A large sample measures the association precisely without selecting the explanation.

Lazear and Shearer sharpen different comparisons

	Lazear (2000)	Shearer (2004)
Setting	Safelite compensation reform	Tree-planting field experiment
Variation	Organizational move toward piece rates	Compensation varied for observed workers
What to separate	Incumbent response and workforce selection	Within-worker response to compensation
Question to carry	Why this timing, tasks, and workforce?	Which tasks and conditions support transport?

The designs address different margins. Experimental variation narrows selection concerns for its comparison; a broader reform can change the organization and the workforce together.

Composition accounting is not mechanism identification

$$\begin{aligned}\bar{y}_1 - \bar{y}_0 = & \underbrace{\int [\mu_1(x) - \mu_0(x)] dF_0(x)}_{\text{change at baseline composition}} \\ & + \underbrace{\int \mu_1(x) d[F_1(x) - F_0(x)]}_{\text{composition change at regime-1 outcomes}} .\end{aligned}$$

- $F_d(x)$: observed type distribution; $\mu_d(x)$: mean outcome by type in regime d ; use a common, well-defined support.
- This is an accounting identity. Within-type changes can contain task shifts, shocks, and unobserved selection as well as effort.
- Changing the reference distribution changes the allocation between terms.

Keep the employment ledger when output disappears

- Distinguish entrants, stayers, movers, and leavers.
- A balanced panel describes survivors, not everyone affected by the rule.
- Do not code a missing review after exit as zero productivity.
- Retention remains an outcome when on-job performance is unobserved.

Two different denominators

Performance among retained workers and retention among all initially assigned workers answer different questions. Neither recovers the missing performance of those who exited.

Separate performance measurement from the pay rule

$$R_{it} = q(y_{it}, \text{quality}_{it}, \text{task difficulty}_{it}) + \nu_{it},$$
$$w_{it} = b_f(R_{it}, \text{grade}_{it}, \text{tenure}_{it}, \text{outside options}_{it}).$$

- R_{it} : recorded performance; q : measurement process; ν_{it} : measurement error or unobserved evaluation inputs.
- b_f maps performance and other conditions into pay. A reform can change q , b_f , or both.
- ν_{it} need not be classical noise: strategic reporting, evaluator bias, and selective observation can matter.

Measure valuable work beyond the rewarded target

- Physical output, quality, time use, ratings, wages, retention, and welfare are distinct outcomes.
- Rewarding a measurable task can divert effort from valuable but poorly measured work.
- Collect an outcome beyond the target the organization rewards.

Build a measurement card

Unit; timing; denominator; producer of the record; incentive to manipulate; missingness rule; claim the measure can support.

Gibbons (1998); Holmström and Milgrom (1991). Week 4 develops the incentive model.

Observe the manager–worker relationship

- Lazear–Shaw–Stanton link worker productivity to supervisors: manager assignments and mobility enter the evidence.
- Hoffman–Tadelis connect survey-based people-management skills to attrition and manager rewards.
- An assessment can contain information, reporting incentives, and selection. Ask who produced it and when.
- Teaching, task allocation, evaluation, and retention are distinct managerial actions; a single manager score may combine them.

Lazear, Shaw, and Stanton (2015), *The Value of Bosses*; Hoffman and Tadelis (2021); Benson and Shaw (2025).

Fixed effects require support and a causal argument

$$Y_{it} = \alpha_i + \gamma_{m(i,t)} + \delta_{k(i,t),t} + \beta D_{u(i,t),t} + \epsilon_{it}.$$

- α_i : worker effect; γ_m : manager effect; $\delta_{k,t}$: job-by-time effect; $D_{u,t}$: rule exposure in unit u .
- Mobility and connected comparisons are needed to separate additive worker and manager effects; perfectly nested IDs cannot do so.
- Reassignment after declining performance can confound manager changes with recovery or common shocks.
- This descriptive specification does not automatically identify causal manager quality, the reform effect, or every production input.

Peers and practices operate through relationships

- **Peer exposure:** visibility, interaction timing, and task dependence matter; a team mean is not the full exposure measure.
- **Practice systems:** training may require supervisor support; incentive pay may require reliable measurement.
- **Bundled interventions:** an effect of a package does not identify the contribution of every component.

Design implication

Record team membership and implementation together with individual outcomes. Check spillovers before treating workers as isolated units.

Mas–Moretti (2009); Ichniowski–Shaw–Prennushi (1997); Bloom et al. (2013).

Within- and between-firm dispersion are accounting objects

$$\text{Var}(\ell_i) = \underbrace{\text{Var}(\mathbb{E}[\ell_i \mid f(i)])}_{\text{between-firm means}} + \underbrace{\mathbb{E}[\text{Var}(\ell_i \mid f(i))]}_{\text{within-firm dispersion}}.$$

- ℓ_i is log wages; define the worker-weighted population. Equal firm weights produce a different statistic.
- The identity does not separate productivity, rents, sorting, discrimination, or amenities.
- Internal recruiting, training, promotion, and retention can shape who ends up at which firm.
- Broad administrative data and detailed personnel records illuminate different parts of the allocation process.

Song et al. (2019), *Firming Up Inequality*.

Choose a comparison that matches the question

Design	What to establish	Main unresolved threat
Descriptive personnel panel	Job taxonomy and transitions	Chosen allocation
Within-firm rule change	Timing, exposure, concurrent changes	Unobserved shocks and selection
Manager/team transitions	Assignment spells, mobility support	Endogenous reassignment
Randomized rule or offer	Assignment unit and follow-up	Spillovers and portability
Choice/survey-linked design	Menu, stakes, rubric, response	Selection and measurement

There is no ranking by dataset prestige. Each design supplies a particular comparison and leaves particular claims unresolved.

Examples: Baker et al. (1994); Lazear (2000); Lazear et al. (2015); Bloom et al. (2013); Mas-Pallais (2017).

Data resources offer different objects and access routes

Resource	Available route or object	Boundary to preserve
WMS	Questionnaires and data pathways	Identified/panel access has separate terms; downloads may be under maintenance
MOPS	Public tables and documentation	Aggregate tables are not public personnel microdata
LPP	Linked employer/employee perspectives	Application and controlled access
O*NET	Public versioned occupation descriptors	Occupation averages are not employee outcomes
Firm records	Histories under a research agreement	No partnership or unrestricted access assumed

World Management Survey; U.S. Census MOPS; IAB Research Data Centre LPP; O*NET. Access distinctions recorded in the course resource map.

Build histories before running a regression

- 1 Specify stable worker IDs, employment spells, assignment dates, pay components, outcomes, eligibility, and exposure.
- 2 Join each outcome to the manager and job held *at that time*; preserve histories rather than the latest label.
- 3 Reconcile payroll periods, calendar dates, and event time. Test unique keys and merge cardinality.
- 4 Record when each field became reliable; a new dashboard can increase measured activity without changing production.
- 5 Keep a missingness and exit ledger; define supported comparison cells.

Management measures also require a rubric and protocol: Bloom and Van Reenen (2007).

A mechanism can travel even when its magnitude does not

Write a portability statement

Production technology; workers and jobs observed; outside options; organizational rule; compliance and implementation; missing outcomes.

- A setting with standardized tasks and visible coworkers can reveal peer pressure without representing all work.
- Many workers in one firm do not create many independent organizations.
- A multi-firm sample widens coverage but can still select participating firms and observed workers.
- State which setting features would strengthen, weaken, or reverse the proposed mechanism.

Frontier cases motivate specific new observations

Starting case	Proposed extension, not an established gap
Hiring discretion	Link screening to later tasks, training, and exits; distinguish selection from organizational support.
Career spillovers	Observe vacancies and internal applications to distinguish displacement from new opportunities.
AI at work	Record changes in delegation, training, and task division after tool introduction.
Talent hoarding	Observe blocked applications as well as completed internal moves.

A workplace rollout does not estimate economy-wide AI employment effects.

Hoffman–Kahn–Li (2018); Bianchi et al. (2023); Brynjolfsson–Li–Raymond (2025); Haegele (2026).

Trace incidence for firms and workers separately

Firm outcomes

- Revenue and quality
- Implementation costs
- Recruiting and turnover costs
- Coordination and future capacity

Worker outcomes

- Earnings and effort costs
- Schedules, safety, and autonomy
- Learning and promotion
- Outside options and future value

Low quits can reflect attractive work, weak outside options, or both. Neither a quit reduction nor a profit increase alone is a welfare calculation.

Research Lab: audit a real organizational experiment

- **Anchor:** Bloom, Han, and Liang (2024), *Hybrid Working from Home Improves Retention without Damaging Performance*.
- **Reproduce:** reconstruct a bounded, documented retention contrast.
- **Diagnose:** separate quit, satisfaction, review, promotion, and task-output outcomes.
- **Transfer:** use artificial personnel records to distinguish selection, performance, and measurement.

This week's question

What population, exposure, denominator, and missingness process support each claim? Week 9 develops the full work-design debate.

Article: [doi:10.1038/s41586-024-07500-2](https://doi.org/10.1038/s41586-024-07500-2). Data/code pointer: [doi:10.7910/DVN/6X4ZZL](https://doi.org/10.7910/DVN/6X4ZZL).

Reproduce: Mode B reconstructs published arithmetic

Published rounded quit rate	
Control	7.2%
Treatment	4.8%

$$\begin{aligned}0.048 - 0.072 &= -0.024, \\ \text{difference} &= -2.4 \text{ pp}, \\ \text{risk ratio} &= 2/3, \\ \text{relative reduction} &= 1/3.\end{aligned}$$

Published-statistic reconstruction; not microdata replication

The rates are not newly estimated. Rounded rates do not supply empirical standard errors, confidence intervals, or group counts.

Bloom, Han, and Liang (2024), main-text retention result. Offline lab mode B.

Mode A: a bounded official-release re-estimation

- The local lab re-estimates retention and the first performance review from the official V1.1 release under its CC BY 4.0 terms.
- A Python translation follows the inspected release contrast. The original Stata/R programs were not executed; this is a bounded re-estimation.
- Mode B remains independently runnable offline. The status records distinguish the two modes and their inputs.
- The adapter preserves anonymized split files without reconstructing identities or inventing cross-file keys.

Assignment-based comparison

The treatment is the offer of the schedule. Birthday-parity categories are not two independent workplace clusters. Actual use requires a separate causal argument.

Diagnose: each outcome needs its own ledger

Outcome family	Audit question
Quit	Which initially assigned workers remain in the denominator?
Satisfaction	Who responds, when, and under what response incentives?
Review	Who is reviewed, by whom, and what happens after exit?
Promotion	Who is eligible and at risk over the stated horizon?
Task output	Which jobs record output, with what hours and quality?

No statistically detected difference does not establish equivalence within a stated margin. Inspect the original equivalence test before paraphrasing its claim. One significant subgroup and one insignificant subgroup do not establish a difference between them.

Transfer: compare three synthetic mechanisms

- 1 **Within-worker change:** performance rises with unchanged workforce composition.
- 2 **Replacement:** composition improves while type-specific production functions remain unchanged.
- 3 **Measurement:** recorded task counts rise while quality or the task mix changes.

Compare the raw mean, a composition-standardized mean on supported cells, and a fixed-incumbent comparison. Keep exits in the employment ledger without inventing their on-job output.

Synthetic teaching example — not an empirical estimate

Similar headline gains can have different mechanisms. Repeated worker IDs do not by themselves make a stayer comparison causal.

Methods Box: inference follows the organization

- **Assignment unit:** team-level randomization does not become worker-level randomization because the file has worker rows.
- **Dependence:** common managers, production shocks, and repeated workers shape the comparisons and inference.
- **Support:** inspect transitions, connectedness, and pre-move outcomes before interpreting manager effects.
- **Exposure:** distinguish eligibility, assignment, take-up, and spillovers.
- **Limits:** controls and clustered standard errors cannot repair a missing causal comparison.

Challenge: redesign the ledger for The Value of Bosses

- 1 Map each worker's manager spell to the dates of observed production.
- 2 Identify the comparisons supported by worker–manager transitions.
- 3 Ask why the transition occurred: what if reassignment followed a missed target?
- 4 Name a better source of variation and the records needed to assess it.
- 5 State firm gains and worker consequences that output alone would miss.

The outcome ledger is reusable. The experiment's identification argument does not transfer automatically to nonrandom manager changes.

Lazear, Shaw, and Stanton (2015), *The Value of Bosses*.

Read for a specific role in the argument

Reading role	Assignment
Required foundations	Lazear–Shaw (2007): personnel questions as economic choices.
Required field map	Hoffman–Stanton (2025): selected introduction, productivity, and external-validity sections.
Required personnel lens	Baker–Gibbs–Holmström (1994): what records reveal before causal modeling.
Required lab	Bloom–Han–Liang (2024): design, outcome definitions, and data availability.
Rotating discussion	One of Lazear–Shaw–Stanton (2015) or Haegele (2026); compare across groups.
Managers' pathway	Benson–Shaw (2025): distinguish managing people and projects.

Full titles and clickable scholarly references appear in the chapter's reading ladder.

Build a design memo and carry it into Week 2

One-page research design

Decision-maker; organizational rule; worker response; firm and worker outcomes; data; comparison; principal threat; one result that would falsify the proposed mechanism.

- Explain the assignment total of 21 without treating the matrix as observed data.
- Give three mechanisms for higher recorded output and one discriminating measurement for each.
- Finish a portability statement for one empirical paper.

Week 2: whom does the firm see, how does it screen, and who enters? Link recruitment to later assignment, learning, performance, and retention.