Discipline at Work: Wage Premia, Managers, and Shirking in a Large Mexican Retailer

Job Market Paper

Latest version: https://aneyranazarrett.github.io/research.html

Andrea Neyra-Nazarrett*

September 15, 2026

Abstract

Inventory shrinkage—merchandise missing from the store—accounts for about 1.6 percent of sales for United States retailers, on the order of grocers' net profit margin. Yet its drivers are not well understood. Partnering with a large Mexican retailer—more than 300 stores and roughly 25,000 workers at a time—I show that shrinkage varies with the store's organization: with the quality of its management and with its wage premium, and that the two work as complements. Leveraging manager rotations that are not directed by store performance, I estimate manager fixed effects on shrinkage. The managers who reduce it are not better bookkeepers; they monitor: they work more shifts, including nights and weekends, are more co-present with their workers on the floor, and rework the teams they inherit—moving workers up a tier and across areas, and raising turnover in the short run through quits, with no detectable rise in dismissals. On pay and schooling they are indistinguishable from the rest. Shrinkage also falls with the wage premium. But the two levers are not additive: after an arrival by one of these managers, shrinkage falls by 12 percent where the premium is high against 4 percent where it is low—a gap of 8 percentage points. An efficiency-wage model rationalizes the gap. Consistent with it, managers do not reorganize more where

*Harvard University. Email: aneyra@g.harvard.edu. I am deeply grateful to my advisors, Jesse Shapiro, Emily Breza, and Jorge Tamayo, for their guidance and support, and to Augustin Bergeron, Richard Freeman, Anders Jensen, Lawrence Katz, Gabriel Kreindler, and Amanda Pallais for invaluable feedback. I thank members of the DAG and participants in the Harvard Labor and Development Lunch and the Political Economy Lunch for helpful comments. I gratefully acknowledge financial support from the David Rockefeller Center for Latin American Studies (DRCLAS), the Weatherhead Center for International Affairs, and the Mildred Rendl PhD 1954 Graduate Student Fellowship. All analyses use anonymized firm records accessed through a research partnership; store identifiers and exact locations are confidential.

pay is high, but the same presence enforces more: suspensions rise only where there is a more valuable rent to confiscate.

1 Introduction

Inventory shrinkage—merchandise that goes missing from stores—is one of the largest sources of revenue loss in retail. In the United States it reached \$112 billion in 2022, about 1.6 percent of sales (National Retail Federation, 2023), on the order of grocers' net profit margin (Damodaran, 2026); the most recent global survey puts it at a similar 1.8 percent of sales worldwide (Tyco Retail Solutions and PlanetRetail RNG, 2018). Retailers treat shrinkage as a security problem: they lock merchandise behind glass, expand video surveillance, and increase loss-prevention budgets (National Retail Federation, 2023; Selyukh, 2024); some chains have closed stores over it. Yet despite its size, the drivers of shrinkage are not well understood.

The evidence retailers themselves produce points inward. In the industry's own security survey, retailers attribute more than half of their shrink—56 percent—to internal causes, their own staff and processes, against 36 percent to outside actors (National Retail Federation, 2023). The nature of the work says the same: guarding and managing inventory is the quintessential task of a supermarket job—in my setting, it appears in 92 percent of workers' job descriptions. Yet shrinkage is rarely analyzed as a personnel problem.

Two obstacles stand in the way. The first is data: retailers rarely record shrinkage systematically across stores and over time, and where they do, it stays an accounting line, never linked to the workers and managers behind it. The second is a framework. Caring for merchandise is effort no supervisor observes directly, and a lapse costs the worker something only if it is detected—exactly the structure of the canonical shirking model (Shapiro and Stiglitz, 1984). I therefore model shrinkage as the observable residue of on-the-job shirking. Two instruments deter shirking in that model: monitoring—in a store, the manager watching the floor—which sets the probability that a lapse is caught, and a wage premium, which gives the worker a rent to lose if it is. This paper asks whether the two in fact deter shrinkage and, if so, whether they work as complements or substitutes.

This paper answers these questions with administrative records from a large Mexican retailer—more than 300 stores and roughly 25,000 workers in a typical month over 2022–2024. The records measure all three ingredients at once: shrinkage audited weekly in every store; from biometric time clocks, what managers and workers actually do; and each worker's pay relative to her local outside wage.

I present a model that adapts the canonical shirking framework to the shop floor. A worker chooses how much care to supply—the hidden effort that keeps merchandise from going missing—and care is costly. The firm pays a fixed wage by position, so the only formal incentive is dismissal: a worker caught neglecting her job loses the premium she earns over her outside option. Whether she is caught depends on the store's monitoring, which in this setting is the general manager's presence on the floor. The deterrent is therefore a product: the probability that a lapse is detected times the rent lost upon dismissal.

The product structure delivers the paper's three central predictions. First, a better monitor disciplines the hidden margin: shrinkage should fall when a better manager arrives. Second, a larger premium disciplines it too: shrinkage should fall as the rent rises. Third—the prediction at the center of the paper—the two instruments are complements: because they enter multiplicatively, a given manager accomplishes more where the premium is high, and a given premium accomplishes more where someone watches the floor. The model also separates margins by observability: attendance is verifiable and detects itself, so the premium moves it everywhere; shrinkage is punished only when detected, so the rent reaches it only through the monitor. Two further propositions sharpen the account: the sanction is the manager's own act and follows the rent, so she punishes where the job is worth keeping; and a store already watched closely sits at a corner, where a small loss of rent moves nothing.

I build three panels. The first is an employee-month panel covering close to 100,000 employees over 2022–2024, built from the firm's personnel records. For each employee it records demographics, the position held at each point in time, and wages, from which I measure the premium: I match each employee to Mexico's national labor-force survey (ENOE) and define her outside option as the average labor income in her cell of gender, education, age, municipality, and quarter; the wage premium is the difference between the wage the firm pays her and that outside wage. The second is a store-month panel for the firm's more than 300 stores. It records shrinkage in value and physical units, sales and customer counts from high-frequency scanner data, each store's format and organizational structure, and its location—including the municipality, which links the store to its local labor market, and whether it lies in the northern-border zone with its higher minimum-wage floor.

The third is a manager-month panel for the general managers who run each store—a single managerial layer, one or more per store. It records which manager runs which store in each month—the assignment histories that identify arrivals, departures, and rotations—along with her demographics, pay, and tenure, and, from the biometric time clocks, when she is physically in the store: her floor presence, including at night and on weekends, and the overlap of her schedule with her workers' shifts.

I take the model to the data with three designs, one per central prediction. The first tests the manager. I leverage the firm's rotation of general managers across stores—a staffing and development policy, not a response to store performance—to estimate a fixed effect for each manager, classify as high-type those whose fixed effects imply less shrinkage, and follow shrinkage around the arrival of a high-type manager at a store. I also use the time clocks, the movement log, and the personnel records to characterize what the managers who reduce shrinkage do in the store.

The second design tests the premium. Because headquarters sets wages centrally by position, a worker's premium moves when her local outside wage moves, for reasons unrelated to her performance or her store's. I use this variation in a two-way fixed-effects analysis within municipality and time, and I complement it with a statutory shock: Mexico's minimum-wage hikes compress the premium over the legal floor more at the northern border than in the interior, so I compare border against interior stores of the same format around each hike.

The third design tests the complementarity, and it asks the model's question exactly: does the same monitor accomplish more where there is more rent to confiscate? I return to the high-type arrivals and split them by the wage environment the manager walks into—whether the premium in the store's municipality is above or below that municipality's own median over the sample—fixed at the month of arrival, so that nothing that happens after the arrival can reclassify the store. The same event is then read twice: the arrival effect where the rent is large, and where it is small, with the difference between the two as the object of interest. Because the classification is estimated from shrinkage itself, I measure how much of that difference its mechanics alone can produce rather than assume it nets out; what survives is the model's cross-partial. The same records then ask how the gap arises: whether managers act differently where the premium is high, or whether the same actions simply bite harder where there is more to confiscate.

Around the arrival of a high-type manager, shrinkage falls by close to 7 percent, and it is flat in the months before—a level I read as validating the classification. The managers who reduce shrinkage are not better bookkeepers; they monitor: they work more shifts, including nights and weekends, are more co-present with their workers on the floor, and rework the teams they inherit—moving workers up a tier and across areas, and raising turnover in the short run through quits, with no detectable rise in dismissals. They do not add security or supervisory staff, and on pay and schooling they are indistinguishable from the rest.

The premium disciplines too: a 10 percent lower premium is associated with about 1.1 percent more shrinkage, and around the minimum-wage hikes border shrinkage relative to sales is unchanged on impact and rises by about 4 percent as the compression persists. The main finding is what happens when the two levers meet. After a high-type arrival, shrinkage falls by close to 12 percent where the premium is high against roughly 4 percent where it is low—a gap of eight percentage points. Consistent with the model, managers do not reorganize more where pay is high, but the same presence enforces more: suspensions rise only where there is a more valuable rent to confiscate (Figure 12). The two instruments are complements, each worth most where the other is present. The border hikes close the loop from the other side: when the compression landed, the rise in shrinkage concentrated in stores whose manager was seldom on the floor.

The paper's first contribution is the outcome itself. Despite its magnitude, shrinkage has been treated in economics as an accounting line rather than as behavior; to my knowledge, this is the first paper to characterize it systematically and to link it, store by store and manager by manager, to the personnel of the store.

Second, the paper contributes to the evidence that pay premia sustain discipline within firms. The classic reference is the interplant test of Cappelli and Chauvin (1991), who relate the premium a manufacturer pays across plants to disciplinary dismissals; the closest recent work is Emanuel and Harrington (2024), who show that a voluntary firm-wide minimum wage halves worker departures, reduces absenteeism, and raises warehouse productivity. The present study adds variation: pay is set centrally while local outside wages move, and two statutory minimum-wage schedules move the premium by law, so the premium moves both with local labor markets and by statute, across more than three hundred stores, against the same audited outcome.

Third—and closest to this paper—Coviello et al. (2022) establish a wage-by-monitoring complementarity: a minimum-wage increase raises the productivity of bound workers, and by more where monitoring is intensive. This paper differs in how monitoring is measured. Their proxy is a staffing quantity—the store-month supervisor-to-worker ratio, which, as the authors note, "captures the extensive margin of monitoring but not the intensive margin (supervisor effort)"; here the monitor is an individual manager whose arrivals and departures I observe. Rotation converts monitoring into discrete, dated, within-store events rather than gradual changes in a staffing ratio; the time-clock records measure the manager's presence directly—her hours, her store, and her overlap with her workers—the intensive margin a staffing ratio cannot capture; and the manager becomes an object of direct interest, since how managers complement a firm's other incentive instruments is a central question of personnel economics.

Finally, the paper adds an outcome to the literature that recovers manager effects from rotation and asks what good managers do (Bertrand and Schoar, 2003; Lazear et al., 2015; Minni, 2026; Adhvaryu et al., 2026). The novelty lies in the margin and its interpretation: the manager's effect is estimated on a margin no supervisor observes directly, traced to concrete managerial conduct, and shown to depend on the wage environment in which the manager operates.

2 Data and Institutional Background

2.1 The Firm and the Data

I partner with a large Mexican retailer with a nationwide presence. The firm operates more than three hundred stores across the country and employs around twenty-five thousand workers in any given month; the panels cover roughly one hundred thousand distinct workers over the full period. Central headquarters standardizes stores tightly within format, setting the physical layout, the wage attached to each position, the human-resources and personnel policies, and the operating routines that govern each location. Table 1 summarizes the firm's scale and the structure of the data, which span January 2022 through June 2024.

Table 1: Company summary

Category	**Description**
Country	Mexico
Sector	Retail
Number of workers	around 100,000
Number of stores	300+
Stores outside border zone	200+
Stores inside border zone	100+
Period	January 2022 – June 2024

Data access provided through a research partnership. All analyses use anonymized records; store identifiers and exact locations are confidential.

Using the firm's administrative records and the ENOE, I build three linked panels. The first is the employee–month panel, from the personnel and scheduling system and the time clocks: demographics, position, and the description, tasks, and actions of each job—guarding and managing inventory appears in 92 percent of these job descriptions; wage and wage premium; shifts and hours, with each entry and break stamped to the minute; absences with their reason; suspensions; and every move—promotion, transfer, separation—with the reason it occurred. The biometric time clock is the same instrument for everyone in the store: it stamps the start and the end of every worked shift to the minute, for frontline workers and general managers alike, so the records place each person in the store day by day—nights and weekends included—and any two employees' schedules can be laid over one another. The personnel-movement record behind these variables is a log of events, not a snapshot: it registers every hire, every separation with its recorded cause—quits, dismissals, and severances enter separately—every within-store reassignment across positions, tiers, and occupational areas, and every promotion and demotion, each stamped with the month it occurred. The same system carries the disciplinary record: every formal suspension is logged with its date, so the sanctions applied in a store can be counted month by month. The second is the store–month panel: scanner-level sales aggregated to the month, the weekly inventory audit's shrinkage in value and physical units, format, and location. The third is the manager–store panel, from the same personnel system, which follows each general manager across stores over time.

2.2 Store Scale and Structure

Table 2 reports the distribution of sales and headcount at the store level. Each store has three tiers: a managerial layer of one or more general managers (up to three in the largest formats), a middle tier of area sub-managers and supervisors, and the general workforce. Every position in the firm's catalog maps to a tier and to an occupational area—client-facing roles in sales, service, and food; back-office roles in logistics and administration; security; and the chain that receives, tags, and protects merchandise—so the composition of a store's workforce is observed position by position in every month.

Table 2: Store sales and activity (store–month)

Variable	Mean	Median	Min	Max
Sales (USD)	682,726	550,828	136,785	2,087,354
Sales (units)	426,469	347,440	95,543	1,197,007
Clients (count)	47,123	42,156	3,608	118,945
Workers per store	67.3	54	1	262

Store–month observations (more than 300 stores, January 2022–June 2024). Monetary values are in USD, converted at MXN 18.5/USD, the approximate average interbank rate over the period. Inventory shrinkage is reported separately in Table 4.

2.3 Store Managers

The general manager is the pivotal personnel actor in each store. She hires and dismisses workers, assigns and reassigns them across occupational areas, promotes and moves them within the store, sets the schedule, and oversees inventory—and she is the store's monitor, supervising through her presence on the floor and the hours she overlaps with frontline workers. Because pay is fixed centrally by position, none of these levers runs through compensation: the manager can neither raise nor cut a worker's wage, so her influence operates through selection, allocation, scheduling, and supervision. The firm rotates general managers across stores routinely. For each general manager the records carry the same detail as for any employee: her demographics, schooling, pay, and tenure; her complete assignment history across stores; and her own time-clock entries, which locate her on the floor shift by shift.

2.4 Wage Premia and Shirking Measures

Two constructs anchor the analysis: the wage premium, which measures the rent the firm pays over the outside option, and a set of shirking measures, of which inventory shrinkage is the central one.

Wage premia. Following the efficiency-wage logic of Cappelli and Chauvin (1991), I define the wage premium as the difference between a worker's wage inside the firm and the wage the same worker would command outside it,

$$W_{it}^{p} \;=\; w_{it} - w_{it}^{m}, \tag{1}$$

where w_{it} is worker i's firm wage in month t and w_{it}^{m} is her outside wage. I measure the outside wage by matching each worker to a reference profile in the ENOE, the national labor-force survey administered by INEGI, which covers both formal and informal workers: the average labor income of employed respondents in her full cell—same gender, education, age, municipality, and quarter—which covers about 86 percent of worker-months. Where the full cell is empty, the match relaxes one dimension at a time, education last, and never drops the quarter or the state; 99 percent of worker-months obtain an outside wage.[1] The ENOE also reports each municipality's unemployment rate by quarter, which allows the outside option to be scaled by the probability of finding a new job. Table 3 reports the firm wage, the outside wage, and the premium in levels and as a multiple of it.

[1]Relaxing gender, then age, then municipality accounts for a further 0.7, 0.5, and 1.2 percent of matches (so about one percent are matched at the state rather than the municipality level); education—the most predictive dimension given the firm's own wage structure, and therefore relaxed last—accounts for the remaining 12 percent.

Table 3: Wages and the wage premium (worker–month)

	Mean	Median
Firm wage (USD)	476	428
Outside wage, ENOE (USD)	382	364
Wage premium (USD)	93	69
Wage premium (multiple)	1.33	1.20

Worker–month observations. The firm wage is monthly contractual pay; the outside wage is the ENOE cell wage matched on gender, education, age, municipality, and quarter (the cascade for empty cells never drops the quarter or the state). The wage premium in levels is the difference between the two; the multiple is the ratio of the firm wage to the outside wage. USD at MXN 18.5/USD.

Shirking. To gauge shirking and effort I use four personnel outcomes: turnover, the monthly rate at which workers leave the store (the records carry the reason of each separation); absenteeism, in days and as an indicator for any absence, separating non-justified from justified; and, from the scheduling system, monthly shifts and daily worked hours. As in other large retailers, tenure is short and absence and turnover are high: median tenure is 26 months. The paper's central measure is inventory shrinkage—the audited value of merchandise the firm records as owned but cannot find on the shelf.[2] The firm's weekly inventory audit measures it in both value and physical units: it compares the book stock the firm's system expects a store to hold against a physical count of what is on the shelf, and records the gap. The firm counts at three frequencies—a daily reconciliation of book against physical stock and fuller counts each week and each month—so losses are caught close to when they occur; at the close of each month it reconciles the flow identity (inflows = sales + shrinkage), so recorded losses tie out against measured stock movement. Who takes the count changes from week to week; it is not a fixed person. In this firm shrinkage averages 2.53 percent of sales.

Interpreting shrinkage as shirking rests on an assumption about where the losses come from. Shrinkage responds to two behaviors of the workforce: the attention workers devote to preventing losses on the floor—watching merchandise, handling it with care, flagging problems as they appear—and the losses workers themselves originate. Effort moves both margins in the same direction: high effort prevents shrinkage, while low effort stops preventing it or, at the extreme, causes it. The data support this reading: store shrinkage covaries most with the workforce and organization—stores whose workforces are less experienced, less

[2] Merchandise that is spoiled or damaged but still identifiable is logged in a separate register and is not counted as shrinkage, so the measure captures stock that is simply missing.

stable, and more often absent without justification carry more shrinkage—while customer traffic enters weakly and local robbery, homicide, and perceived insecurity cluster near zero (Appendix L). Table 4 reports summary statistics for shrinkage and the remaining shirking measures at the store and worker levels.

Table 4: Shirking measures at the store and worker levels

Variable	Mean	Median	Min	Max
Store level (store–month)				
Inventory shrinkage				
Shrinkage-to-sales (%)	2.53	2.45	0.10	10.52
Shrinkage (USD)	16,931	12,450	807	67,329
Shrinkage (units)	6,440	5,012	47	21,497
Shrinkage per worker (USD)	242	230	16	1,219
Absence (store total, days per month)				
Absence days, total	126.3	101	0	626
Non-justified absence days	58.2	46	0	313
Turnover (% of workforce per month)				
Turnover rate, total	6.20	5.48	0	31.25
Worker level (worker–month)				
Absence (days per month)				
Absence days, total	1.88	0	0	31
Non-justified absence days	0.86	0	0	28
Absence (indicator)				
Any absence (= 1)	0.462	0	0	1
Any non-justified (= 1)	0.361	0	0	1
Effort				
Monthly shifts	21.33	25	1	56
Daily worked hours	8.26	7.92	0	24

Store statistics are computed at the store–month level (more than 300 stores); worker statistics at the worker–month level, January 2022 through June 2024. USD at MXN 18.5/USD. At the store level, absence days are summed across the store's workers, and turnover rates are separations per hundred workers. At the worker level, absence days are per-worker monthly counts and the absence indicators equal one if the worker had at least one absence of that type. *Non-justified* absence is absence without an approved justification; vacation days are reported separately and excluded here. Effort margins (shifts, hours) are recorded by the scheduling system.

3 Theoretical Framework

I model the effort choice of a worker whose only incentive is the threat of losing a job that pays a premium over the local market wage. The model adapts Shapiro and Stiglitz (1984) and is the pure efficiency wage case of Coviello et al. (2022): fixed salaries, no pay-for-performance, an outside option equal to the market wage rather than unemployment, and dismissal as the firm's only formal incentive. The expected penalty for shirking is a *product*: the probability that a lapse is detected, which the manager sets, times the rent the worker forfeits if sanctioned, which the premium sets. The results follow the order of the paper—the monitor, the premium, their product, and who enforces it—and each proposition names the section that tests it.

3.1 Environment

Worker i of type x_i is employed at store j under manager m. Each period she chooses a *care level* $e_i \in [0,1]$—the hidden effort that keeps merchandise from going missing, including keeping time on the floor once she is there—and an *absence level* $a_i \in [0,1]$, the share of assigned shifts she shows up for. Care has consequences only if someone detects the losses it produces; absence is recorded by the time-clock and detected mechanically, whoever runs the store. Store shrinkage is $S_j = f(1-\bar{e}_j)$, with $\bar{e}_j$ average care and f strictly increasing. Costs are $c(x_i, e)$, with $c_e > 0$ and $c_{ee} > 0$, and $\kappa(a)$, increasing and convex.

The worker receives a fixed wage

$$w_i \;=\; w^m + W_i^p, \tag{2}$$

where w^m is the local market wage she can obtain immediately if she leaves[3] and $W_i^p \geq 0$ is the *wage premium* (Section 2, Equation 1). The wage does not depend on realized shrinkage. The model is stationary: the propositions are comparative statics in the standing premium, and each worker-month observation maps to the model evaluated at its current premium, treated as permanent.

The job ends through two hazards, one per margin. On the care margin, manager m monitors with intensity $\mu_m \geq 0$; a lapse by a worker exerting care e is detected with probability $q(\mu_m)h(e)$ and, in the baseline, sanctioned automatically, so she keeps her job

[3]Immediate reemployment is a normalization. With a search spell in informal work at $w_u < w^m$ and reemployment at rate λ, everything below holds with the premium augmented by $\delta \equiv (w^m - w_u)\, r/(\lambda + r)$, about three percent of the median premium under ENOE spells of one quarter.

with probability

$$\pi(e;\mu_m) \;=\; 1-b-q(\mu_m)\,h(e), \qquad b\in(0,1),\; q>0,\; q'>0,\; h\geq 0,\; h'<0,\; h''\geq 0, \tag{3}$$

with b an exogenous separation rate, q the detection rate the manager generates, and h the worker's exposure to detection. Hence

$$\pi_e \;=\; q(\mu_m)\big(-h'(e)\big) \;>\; 0, \qquad \pi_{e\mu} \;=\; q'(\mu_m)\big(-h'(e)\big) \;>\; 0: \tag{4}$$

more care lowers the dismissal risk, and closer monitoring rewards care more strongly. The parametric form is not innocuous (Section 3.5). On the absence margin, retention is $\pi^A(a)$ with $\pi^{A\prime}>0$ and no μ_m. Treating the two hazards as mutually exclusive within a period,

$$\Pi(e,a;\mu_m) \;=\; \pi(e;\mu_m)+\pi^A(a)-1, \tag{5}$$

which must lie in $[0,1]$. The additive form is substantive: a multiplicative one would place μ_m inside the absence condition.

3.2 The Worker's Problem

Time is discrete. Let $V^E(x_i)$ be the value of employment and V^A that of the outside option, with flow value $u^A=\frac{r}{1+r}V^A=w^m$ since the worker can always find work at w^m. Writing the employment Bellman equation and subtracting V^A, the *employment surplus* $V(x_i)\equiv V^E(x_i)-V^A$ satisfies

$$\boxed{V(x_i)=\max_{e,\,a\,\in\,[0,1]}\left[W_i^p-c(x_i,e)-\kappa(a)+\frac{\Pi(e,a;\mu_m)}{1+r}\,V(x_i)\right].} \tag{6}$$

The market wage drops out: only the premium determines incentives. At an interior optimum the first-order conditions are

$$c_e(x_i,e^*) \;=\; \frac{\pi_e(e^*;\mu_m)}{1+r}\,V(x_i), \tag{7}$$

$$\kappa'(a^*) \;=\; \frac{\pi^{A\prime}(a^*)}{1+r}\,V(x_i), \tag{8}$$

and the surplus at the optimum is

$$V(x_i) \;=\; \frac{(1+r)\big[W_i^p-c(x_i,e^*)-\kappa(a^*)\big]}{1+r-\Pi(e^*,a^*;\mu_m)}. \tag{9}$$

Both conditions price the surplus the worker stands to lose; only the care condition runs through the manager's monitoring.

Definition 1 (Stationary partial equilibrium of the worker)**.** *Given* (x_i, W_i^p, μ_m)*, a triple* (V, e^*, a^*) *such that (a)* $0 \leq \Pi(e, a; \mu_m) \leq 1$ *for all* (e, a)*; (b)* (e^*, a^*) *maximizes the right-hand side of* (6) *given* V*, with Kuhn–Tucker conditions* $g_e \equiv -c_e + \pi_e V/(1+r) \leq 0$ *if* $e^* = 0$*,* $= 0$ *if interior,* ≥ 0 *if* $e^* = 1$*, and likewise* $g_a \equiv -\kappa' + \pi^{A\prime} V/(1+r)$*; (c)* V *is the fixed point* (9)*; (d)* $V \geq 0$*.*

Nothing here is a firm's or a market's equilibrium: (W_i^p, μ_m) are exogenous. Write $\rho \equiv 1 + r - \Pi(e^*, a^*; \mu_m)$, $D \equiv c_{ee} - \pi_{ee} V/(1 + r)$, $D_A \equiv \kappa'' - \pi^{A\prime\prime} V/(1 + r)$, and let $R = W \times [\mu_{LT}, \mu_{HT}]$ be a rectangle of premiums and monitoring intensities on which (e^*, a^*) is interior, with c and h three times and κ, π^A twice continuously differentiable.

Lemma 1. *Under* (3) *and* $D_A > 0$*: (i)* V *exists and is unique, the care problem is single-peaked, and* $D > 0$*; (ii) the care condition is* $M(e^*) = q(\mu_m)\, V/(1 + r)$ *with* $M(e) \equiv c_e(x_i, e)/(-h'(e))$ *and* $M' > 0$*, so care is a strictly increasing transform of the expected-penalty product; (iii)* V *is strictly increasing in* W_i^p*, and* $V > 0$ *at an interior optimum requires* $W_i^p > c(x_i, e^*) + \kappa(a^*)$*.*

Proofs are in Appendix A.

3.3 Linking the Model to the Setting

The monitor. The argument of (7) is the monitoring the worker actually faces in store j in period t, $\mu_{mjt} = \bar{\mu}_m + u_{mjt}$, with $\bar{\mu}_m$ portable. The arrival designs classify managers by their AKM shrinkage fixed effect (Section 4), high-type (HT) below the median of their connected set; Section 5 grounds the label in conduct: high-type managers work more shifts, overlap more of their workers' hours, and their workers keep better time while they are on the floor. Monitoring is not an apparatus the manager purchases but her presence on the floor, which proxies μ_{mjt} itself; nothing in her problem re-optimizes it against the local rent, so her actions should not vary with the premium while the workers' response does (Section 8.3). Where label and realized presence disagree as moderators (Section 9), the model sides with the realized value.

The premium. In Shapiro and Stiglitz (1984) the punishment for shirking is unemployment; here the market always offers w^m, what dismissal destroys is the premium, and (6) makes the premium the whole of the incentive. Workers earn on average about 1.33 times their outside wage (Section 2). Because the contract wage is set centrally by position and moves rarely, movements in the local outside wage pass through one-for-one to the premium, the variation the two-way fixed-effects design of Section 6 exploits.[4] A statutory minimum-wage hike (Section 7) moves the premium in two phases with opposite signs: on impact it raises the contract wage of the workers the floor covers, the case Coviello et al. (2022) study; as the local outside wage catches up against a scale that does not move, the premium compresses, the case they set aside as demotivating. Proposition 2 signs both; the stationary model does not date the second.

The no-shirking condition. In the binary case of Shapiro and Stiglitz (1984)—$e \in \{0, \bar{e}\}$, absence suppressed, retention $1-b$ with effort and $1-b-q$ without—the surplus of a worker who exerts effort is $V_H = (1+r)(W_i^p - \bar{e})/(r+b)$, a one-period deviation saves $\bar{e}$ and exposes V_H to detection risk q, so effort is optimal if and only if

$$W_i^p \;\geq\; \bar{e} + \frac{(r+b)\,\bar{e}}{q} \;\equiv\; \widehat{wp}, \tag{10}$$

the effort cost, which a rent gross of effort must cover, plus the no-shirking mark-up. Read as the firm's menu, the condition makes the two instruments *substitutes*: the same care can be sustained with more monitoring and less premium. That is the null the arrival design of Section 8 rejects in levels.

3.4 Results

Proposition 1 (The monitor disciplines the hidden margin)**.** *Let (e^*, a^*) be interior. Under higher monitoring intensity μ_m:*

(i) Care is higher at any given premium, $\partial e^*/\partial \mu_m > 0$, and shrinkage lower.

(ii) Absence does not fall. Each margin's response to the manager is its response to the premium rescaled by a detection stake,

$$\begin{aligned} \frac{\partial a^*}{\partial \mu} &= -\frac{q'(\mu)\,h(e^*)\,V}{1+r} \cdot \frac{\partial a^*}{\partial W^p} \;\leq\; 0, \\ \frac{\partial e^*}{\partial \mu} &= \frac{q'(\mu)\,V}{1+r} \cdot \frac{r+b+1-\pi^A(a^*)}{q(\mu)} \cdot \frac{\partial e^*}{\partial W^p} \;>\; 0, \end{aligned} \tag{11}$$

so a stricter monitor moves absence only through the residual exposure $h(e^*)$ of a

[4] The treatment of monitoring parallels Coviello et al. (2022): detection depends on effort when monitored, termination follows detection, monitoring is exogenous. Their μ is the share of workers monitored in a store; here μ_m is the intensity embodied in the manager.

compliant worker, zero in the benchmark $h(e^*) = 0$*, while her effect on care runs through the slope* $-h'(e^*)$ *and is bounded away from zero.*

(iii) With an unanticipated shock $\varepsilon \sim G$ *to the outside option, realized after actions, quits rise with* μ_m *in proportion to* $g(V)\, q'(\mu) h(e^*) V/\rho$*: only among workers who still carry exposure. The dismissal rate of a compliant worker,* $q(\mu)h(e^*)$*, stays near zero.*

Part (i) uses the form (3), not only the signs (4): a stricter monitor also trims the surplus of a worker with exposure, and under (3) the deterrence channel dominates.[5] Sections 4 and 5 test the proposition: shrinkage falls when a high-type manager arrives and climbs back when she leaves; turnover rises through quits with dismissals flat; and the test of (ii) is the arrival design on non-justified absence (Section 8, Appendix H), not the comovement of manager effects across outcomes, which validates the label and is silent on the mechanism.

Proposition 2 (The premium disciplines both margins)**.** *At an interior optimum:*

(i) Care rises, $\partial e^*/\partial W^p = \pi_e/(\rho D) > 0$*, and shrinkage falls.*

(ii) Absence falls, $\partial a^*/\partial W^p = \pi^{A\prime}/(\rho D_A) > 0$*, and time on the job rises.*

(iii) The dismissal hazard $1 - \Pi$ *falls, and with the shock of Proposition 1(iii) quits fall as well.*

Sections 6 and 7 test it: shrinkage, absence, and turnover fall and shifts and hours rise with the premium, and the statutory compression raises shrinkage. Two assumptions sign what follows.

Assumption 1 (Compliant-care detection)**.** *(a) The care hazard is* (3)*. (b) At every optimum in R, the residual exposure of compliant care is small relative to its slope:*

$$h(e^*)\, \pi^{A\prime}(a^*)\, \kappa'(a^*)\, D \;<\; \big(-h'(e^*)\big)\, c_e(x_i, e^*)\, \big(r + b + 1 - \pi^A(a^*)\big)\, D_A. \qquad (12)$$

Assumption 2 (Curvature regularity)**.** $M(e) = c_e(x_i, e)/(-h'(e))$ *is log-concave,* $(\log M)'' \leq 0$.

Part (b) is "$h \approx 0$ at compliant care" as an inequality; the benchmark $h(e^*) = 0$ satisfies it strictly. Assumption 2 disciplines the third-order terms the cross-partial of an optimum drags in and holds whenever c_e and $-h'$ have constant elasticities.

[5] Under $\pi_e \geq 0$, $\pi_{e\mu} \geq 0$ alone, $\partial e^*/\partial \mu = \frac{V}{(1+r)D}[\pi_{e\mu} + \pi_e \pi_\mu/\rho]$ carries the unrestricted sign of π_μ: with $c = e^2/2$, $r = 0.05$, $W^p = 0.05$, $\pi = 0.8 + \mu e - 0.4\mu^2$, at $\mu = 0.9$ the solution is $e^* \approx 0.084$, $\pi^* \approx 0.55$, and $\partial e^*/\partial \mu \approx -0.014$. The form (3) is a parametric member of the monitoring order of Coviello et al. (2022, Definition 2), which rules such cases out.

Proposition 3 (The premium and the monitor are complements)**.** *Let $e^*(W^p, \mu)$ solve* (7) *on R under Assumption 1(a). Then:*

(i) ***(One product does all the work.)*** *Optimal care satisfies*

$$M(e^*) \;=\; q(\mu)\,\frac{V}{1+r}, \qquad M' > 0, \tag{13}$$

with slopes

$$\frac{\partial e^*}{\partial W^p} = \frac{\pi_e(e^*;\mu)}{\rho\, D} \;>\; 0 \qquad \textit{and} \qquad \frac{\partial e^*}{\partial \mu} = \frac{q'(\mu)\big(-h'(e^*)\big)\big(r+b+1-\pi^A(a^*)\big)\, V}{(1+r)\,\rho\, D} \;>\; 0. \tag{14}$$

(ii) ***(Local complementarity.)*** *Under Assumptions 1–2,*

$$\frac{\partial^2 e^*}{\partial W^p\,\partial\mu} \;>\; 0 \qquad \textit{on } R. \tag{15}$$

(iii) ***(The premium amplifies the monitor.)*** *For every $W^p \in W$, $\Delta_\mu e^*(W^p) \equiv e^*(W^p, \mu_{HT}) - e^*(W^p, \mu_{LT})$ is strictly positive and strictly increasing in W^p: for $W^p_L < W^p_H$ in W,*

$$\Delta_\mu e^*(W^p_H) - \Delta_\mu e^*(W^p_L) \;=\; \int_{W^p_L}^{W^p_H}\int_{\mu_{LT}}^{\mu_{HT}} \frac{\partial^2 e^*}{\partial W^p\,\partial\mu}\, d\mu\, dW^p \;>\; 0. \tag{16}$$

Propositions 1(i), 2(i), and 3 are three readings of (13): the derivative of the product in the detection rate, in the rent, and their cross-partial. The premium's proportional grip on the surplus is $\partial \log V/\partial W^p = 1/(W^p - c - \kappa)$, and a stricter monitor raises the cost of the actions the worker chooses, so the same peso moves the surplus by more in proportion; the sign is established under Assumptions 1(b) and 2, not by the product alone. Part (iii) is what Section 8 tests cell by cell—the reinforcement Coviello et al. (2022) find from the opposite direction.

Enforcement. In the store the sanction is the manager's act—a suspension or a dismissal she must file, argue, and staff around—and it costs her.[6] What it buys her is deterrence, which runs only through what the sanction confiscates: a worker with no rent to lose is indifferent to losing the job. Suppose she sanctions a detected lapse if and only if the rent it

[6]A dismissal confiscates the rent V; a suspension without pay confiscates a share of it. The threshold logic is the same, and suspensions are the sanction the records observe (Section 8.3).

confiscates is at least $\bar{v} > 0$. Writing $s \equiv \mathbf{1}[V \geq \bar{v}]$, with V the surplus enforcement sustains, the care hazard becomes $\pi = 1 - b - s\, q(\mu)h(e)$: the detection technology is the manager's presence, but the threat behind it is exercised only where there is rent to confiscate.[7]

Proposition 4 (Enforcement follows the rent)**.** *Under the sanction rule and Assumption 1(a):*

(i) ***(Credibility threshold.)*** *There is a premium threshold $\overline{W}(\mu)$, defined by $V(\overline{W}, \mu) = \bar{v}$, such that enforcement is credible if and only if $W^p \geq \overline{W}(\mu)$. Below it $\pi_e = 0$: care sits at its lower corner and responds neither to the premium nor to the monitor.*

(ii) ***(Sanctions on impact.)*** *When a stricter monitor replaces a laxer one, $\mu_L \to \mu_H$, at the care level workers chose under the old monitor the sanction rate rises by $\big[q(\mu_H) - q(\mu_L)\big]\, h(e_L^*)$ where enforcement is credible under the new monitor, $W^p \geq \overline{W}(\mu_H)$, and by nothing where it is not.*

(iii) ***(Extensive margin.)*** *Above the threshold Proposition 3 holds unchanged; across it, the arrival effect on care rises from zero to $\Delta_\mu e^*(W^p) > 0$. The share of a store's workers above the threshold rises with the store's premium, so the store-level arrival effect rises with the premium through both margins.*

Proposition 4 is the mechanism Section 8.3 tests. The arrival of the same manager should show the same presence in every wage environment, more sanctions where the premium is high and none where it is low, and a larger fall in shrinkage where the premium is high. The signed object is the differential in sanctions on impact; the stationary sanction rate $q(\mu_H)h(e_H^*)$ is not signed. The same rule explains why no premium rescues a store from a weak monitor: with $q(\mu)$ near zero the sanction rate is near zero at any rent (Section 8, low-type panel).

Corollary 1 (Aggregation)**.** *Let store j employ N_j workers with premiums $W_{ij}^p = \overline{W^p}_j + \eta_i$, the joint distribution of (x_i, η_i) fixed, a common μ_j, every (W_{ij}^p, μ_j) in R, and $S_j = \alpha_j + \frac{1}{N_j}\sum_i \ell(1 - e_{ij}^*) > 0$ with ℓ linear and increasing. Under the hypotheses of Proposition 3, $\partial S_j/\partial \overline{W^p}_j < 0$, $\partial S_j/\partial \mu_j < 0$, $\partial^2 S_j/\partial \overline{W^p}_j \partial \mu_j < 0$, and the same signs hold for* $\log S_j$*.*

With convex losses, $f'' > 0$, the corollary requires $f''\, \bar{e}_{W^p}\, \bar{e}_\mu < f'\, \bar{e}_{W^p\mu}$; it also holds composition fixed, which the team evidence of Section 5 licenses.

[7] The rule is a commitment evaluated at the surplus under enforcement. Evaluated lapse by lapse it leaves a band of premiums, between the surplus a non-enforcing manager's workers hold and $\bar{v}$, where the only consistent policy sanctions with some probability and the threshold of part (i) becomes a ramp.

Corners. Section 9 reads a pattern the interior results do not predict: a marginal compression of the premium raises shrinkage only where the manager is not on the floor. On R, by (15), an interior compression removes *more* care where monitoring is high. The insured store is a corner. By Definition 1(b) and Lemma 1(ii), care is at full compliance, $e^* = 1$, if and only if

$$\begin{aligned} \Gamma(W^p,\mu) \;&\equiv\; \frac{q(\mu)\,V^{(1)}(W^p,\mu)}{1+r} - M(1) \;\geq\; 0, \\ V^{(1)}(W^p,\mu) \;&\equiv\; \frac{(1+r)\big[W^p - c(x_i,1) - \kappa(a^*)\big]}{r + b + q(\mu)h(1) + 1 - \pi^A(a^*)}, \end{aligned} \tag{17}$$

with $V^{(1)}$ the surplus at full care and a^* solving (8) at $V = V^{(1)}$.

Proposition 5 (Saturation and marginal insurance)**.** *Under Assumption 1(a), with $D_A > 0$ and $V^{(1)} > 0$:*

(i) Γ is strictly increasing in W^p and in μ, so there is a threshold $\bar\mu(W^p)$, decreasing in W^p, with $e^*(W^p,\mu) = 1$ if and only if $\mu \geq \bar\mu(W^p)$.

(ii) ***(Marginal insurance.)*** *Fix W^p and $\mu_L < \bar\mu(W^p) < \mu_H$. For every sufficiently small compression $dW^p < 0$,*

$$\begin{aligned} e^*(W^p + dW^p, \mu_H) - e^*(W^p,\mu_H) \;&=\; 0, \\ e^*(W^p + dW^p, \mu_L) - e^*(W^p,\mu_L) \;&=\; \frac{\pi_e}{\rho D}\,dW^p + o(dW^p) \;<\; 0. \end{aligned} \tag{18}$$

(iii) If instead both (W^p,μ_L) and (W^p,μ_H) lie in R under Assumptions 1–2, the ordering reverses: the same compression removes more care under the strict monitor.

Section 9 tests Proposition 5. It is not the marginal form of the complementarity—part (iii) orders the stores the other way—but a statement about where each store sits relative to the corner: a compression damages the unwatched store because its care margin is interior and spares the watched one because a small loss of rent leaves the slack Γ positive. Part (i) says which stores should be saturated, those with a high premium *and* a strict monitor, and a compression large enough to push Γ below zero would end the insurance. The same logic on the absence margin, $g_a(1) \geq 0$, insures the attendance of a worker at full compliance.

3.5 Discussion

Workers are risk neutral throughout; Appendix A.2 records where the data refine the model. Three comments.

The detection technology is load-bearing. The form (3) delivers concavity in care by construction and carries the property the two-margin results rest on: a compliant worker has little to fear from a watchful manager, so monitoring changes the return to care, not the survival odds of the compliant; without it the monitoring comparative static can reverse. Assumption 1(b) is stated at the optimum: by Equation (30) of Appendix A.1 it is equivalent to a positive induced-retention response, $\beta \equiv \pi_e \, \partial e^*/\partial\mu + \pi^{A\prime} \, \partial a^*/\partial\mu > 0$, which under linear exposure $h = h_0(1-e)$ holds whenever compliant care is close to full or the absence hazard is flat at the optimum. Exact invariance of absence is a limit, not a specification: $h \equiv 0$ would erase the care margin too, so the model delivers the bound in (11), which vanishes at fixed slope only as $e^* \to 1$, the saturation region of Proposition 5. $D_A > 0$ is assumed, with sufficient conditions parallel to Assumptions 1–2 of Coviello et al. (2022).

What the stationary model does not predict. Timing: the model compares steady states and does not date the compression's lag (Sections 7 and 9); a partial-adjustment extension, $e_{t+1} = (1-\lambda)e_t + \lambda \, e^*(W_t^p, \mu_t)$, delivers it, and its impact period is where Proposition 4(ii) lives. Team production: the hazard $q(\mu)h(e_i)$ assumes detection identifies individual negligence; were losses attributable only to the store, care would be a public good within the team. Corollary 1 holds N_j fixed.

The low-type pattern. A low-type arrival raises shrinkage in both wage environments with no significant difference between them (Section 8). The model predicts the sign, by (14), not the absence of scaling: on the interior, (15) makes the damage from a weak monitor larger where the premium is high. The flat pattern is descriptive, consistent with imprecision or with the low-premium store sitting at the lower corner under the weak monitor. What the model does license is the weaker statement: a premium disciplines the shrinkage margin only through a monitor who enforces (Proposition 4).

4 Manager Effects on Shrinkage: AKM Decomposition

I study the general manager's effect on inventory shrinkage using the firm's rotation of managers across stores. I proceed in three steps. First, I estimate a two-way fixed-effects model in the tradition of Abowd et al. (1999) (AKM), leveraging manager rotations that are unrelated to store shrinkage. Second, I classify managers as high- or low-type by their estimated fixed effects. Third, I take the classification to the arrival-and-departure design of Section 5, which tests it against store transitions and asks what high-type managers do.

The model is

$$Y_{jt} \; = \; \theta_{m(j,t)} \; + \; \psi_j \; + \; \delta_t \; + \; \varepsilon_{jt}, \tag{19}$$

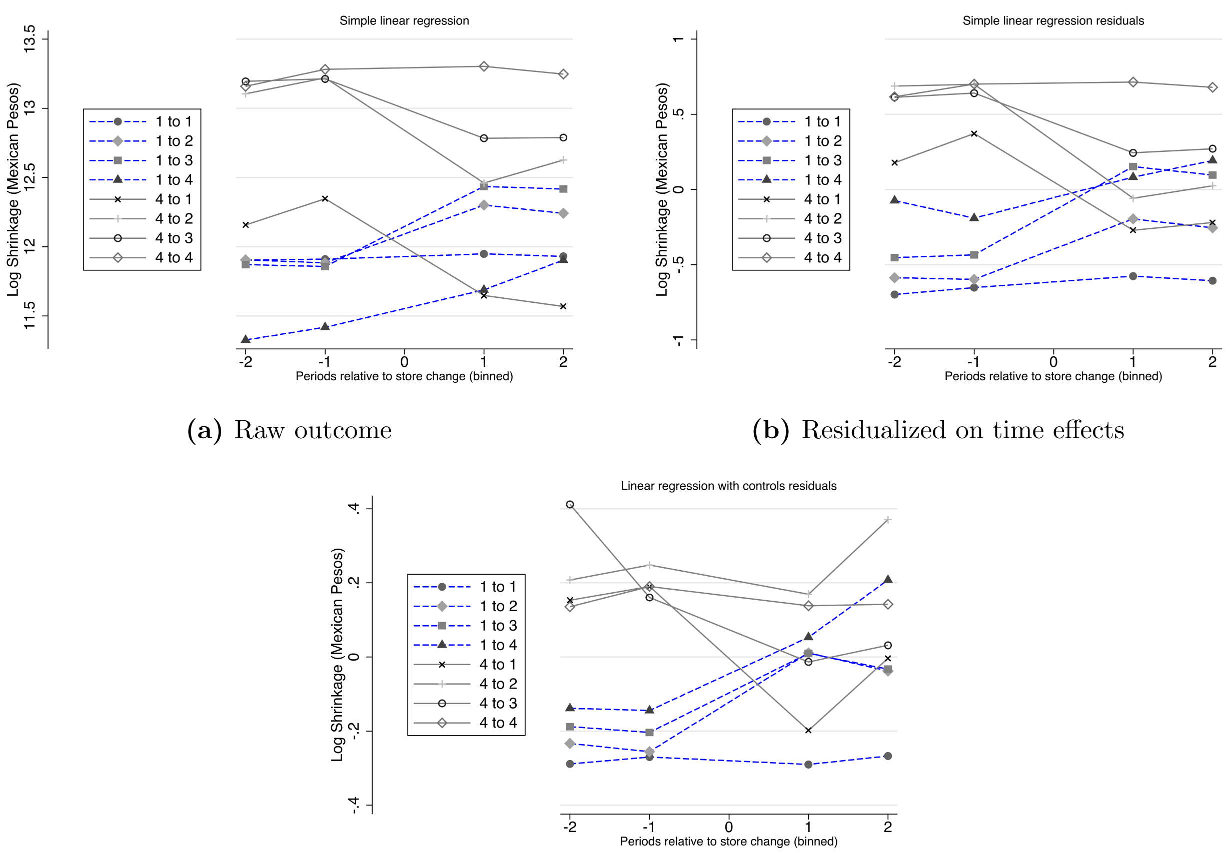
where Y_{jt} is log shrinkage value in store j and month t, $\theta_{m(j,t)}$ is a fixed effect for the general manager running the store, ψ_j absorbs permanent store characteristics, and δ_t absorbs seasonality and firm-wide shocks. Following Card et al. (2013), the error decomposes as

$$\varepsilon_{jt} \; = \; \eta_{m(j,t),j} \; + \; \xi_{jt} \; + \; \nu_{jt}, \tag{20}$$

a manager–store match component, a drifting store component, and transitory noise. Identification requires that manager–store assignment be mean-independent of ε_{jt}: managers may sort on the permanent components, which the fixed effects absorb, but not on the match or on transitory shocks (Appendix B states the moment conditions).

Three facts show that manager assignment is unrelated to store shrinkage, in levels and in changes. First, although the firm audits shrinkage weekly, it is not a key performance indicator in manager assignment: human resources does not track it systematically, and, by the firm's own descriptions of manager reassignments, placement is not directed by the performance of the destination store—in particular, the firm does not send its best managers to the stores with the most shrinkage. Second, the movement records themselves show coordinated rotation rather than targeted placement. Table 14 (Appendix B) classifies the 522 between-store moves: nearly half are links in a same-month chain—the origin store receives another moving manager the month its own departs—one in five is a direct swap between two stores, and one in three fills a vacant store; and the moves cluster in waves—a handful of months absorb the largest share—consistent with coordinated administrative rotation rather than case-by-case reaction. The distance the moves travel in shrinkage is zero: the destination's pre-move shrinkage exceeds the origin's by 0.026 log points ($p = 0.27$).

Third, the moves pass the two direct tests of exogenous mobility. The first is the event-study test of Card et al. (2013). I rank stores by quartile of their average shrinkage, isolate the managers who move across quartiles, and follow *each manager* across her move: the outcome at every date is the shrinkage of whatever store she runs at that date, so at the move the store changes and the manager does not. If assignment were unrelated to the error term, a mover's outcome should jump to the destination store's level on impact—the jump measures the *store*, not the manager—with no differential trend before the move and with upward and downward moves mirroring each other. Figure 1 shows exactly that pattern: movers from bottom- to top-quartile stores step up to the destination level on impact and the reverse moves fall symmetrically; within-quartile moves are flat, so the step is not an artifact of moving itself; and trajectories are flat within origin quartile before the move, so movers were not selected on a diverging shrinkage path.



(a) Raw outcome

(b) Residualized on time effects

(c) Residualized on time effects, tenure, and format controls

Figure 1: Event study around manager moves, by origin and destination store quartile (log shrinkage value)

Notes: Stores are ranked within connected set by quartile of average log shrinkage value. Lines labeled "i to j" trace movers from a quartile-i origin to a quartile-j destination in binned periods relative to the move (±1: one to four months; ±2: more than four months). Dashed lines: origin quartile 1; solid lines: origin quartile 4.

The second test uses the symmetry that mean-independence imposes on the match component of Equation (20). If managers moved to stores where their particular match is favorable, moving up and moving down the store-quality ladder would *both* look advantageous, and the gains from upward moves would exceed the losses from downward ones. Figure 2 plots the average residual change for each upward quartile pair against the corresponding downward pair; symmetry implies the −45-degree line. The points cluster around that line, and no pair deviates in the direction implied by match-based sorting.

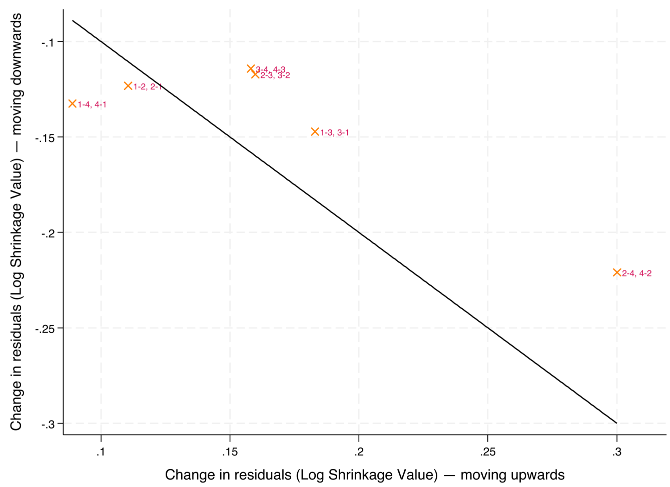


Figure 2: Symmetry of residual changes for upward and downward manager moves (log shrinkage value)

Notes: Each point is a pair of quartile transitions: the horizontal axis reports the average change in residual log shrinkage for managers moving up, the vertical axis the average change for the reverse move. Residuals are from a regression on time effects within connected set. The solid line is $y = -x$, the locus implied by symmetric gains and losses.

I also address limited-mobility bias. A majority of general managers are movers: 55.0 percent run at least two different stores in the sample, 25.4 percent run three or more, and 91.7 percent of stores receive at least one mover (Table 13, Appendix B)—high relative to the matched employer–employee panels on which AKM methods were developed, where mover shares of 12, 25, and 35 percent are typical (Andrews et al., 2012; Card et al., 2013; Alvarez et al., 2018). Manager and store effects are identified within connected sets of stores linked by moves; the largest covers about three-quarters of the panel, and every decomposition below normalizes within its set. The remaining concern is statistical rather than economic: fixed effects estimated from finite mover counts are noisy, and the noise inflates plug-in variance components. I address it with the leave-out estimator of Kline et al. (2020) (KSS).

Table 5: AKM variance decomposition with limited-mobility-bias corrections (log shrinkage value)

	Baseline	*Andrews*	*KSS*
$\widehat{\mathrm{Var}}(\theta)$	0.049	0.009	0.016
$\widehat{\mathrm{Var}}(\psi)$	0.287	0.243	0.251
$\widehat{\mathrm{Cov}}(\theta,\psi)$	−0.014	0.002	−0.002
$\widehat{\mathrm{Corr}}(\theta,\psi)$	−0.120	0.044	−0.038

Largest connected set of the general-manager panel. θ = manager fixed effect, ψ = store fixed effect; observation-weighted variances and covariances. *Baseline*: OLS variance of the estimated fixed effects. *Andrews*: homoskedastic correction of Andrews et al. (2008). *KSS*: leave-out estimator of Kline et al. (2020). The ratio measures (shrinkage over sales and per worker) behave the same way; Table 15 in Appendix B reports them.

Table 5 reports the variance decomposition with and without the correction. In the plug-in decomposition, manager fixed effects account for 49.8 percent of the variance of log shrinkage value and store fixed effects for 42.4 percent—covariance shares, $\mathrm{Cov}(y,\cdot)/\mathrm{Var}(y)$, which allocate the sorting covariance and so need not mirror the raw variance components the table reports. Every manager claim in the paper rests on the *KSS* column; the plug-in shares are quoted only to show what the correction changes.

The leave-out correction cuts the manager component to roughly a third of its plug-in value—$\widehat{\mathrm{Var}}(\theta)$ falls from 0.049 to 0.016—while the store component barely moves (0.287 to 0.251), and the corrected sorting correlation is −0.04, indistinguishable from zero. The corrected component remains economically large: a manager one standard deviation better is associated with about 13 percent lower shrinkage ($\sqrt{0.016}$), holding the store and the month fixed. Appendix B plots the estimated fixed effects themselves; those plots convolve true effects with sampling noise, which is why the inference rests on the corrected components of Table 5 rather than on raw dispersion.

The classification is also insensitive to the estimation error that the correction quantifies. Managers are labeled by a within-cell median split, so the label depends only on ranks, and shrinking each fixed effect toward its cell mean in proportion to its noise reorders very few of them: an empirical-Bayes correction under the plug-in reliability reclassifies fewer than one percent of managers, and a correction under the reliability implied by Table 5 itself—roughly one third—reclassifies under two percent. Re-estimating every design that uses the label under the corrected classifications leaves the results unchanged: the arrival effect, the enforcement response, the presence response, and the complementarity differential retain 88 to 93 percent of their magnitudes at the same significance levels.

The decomposition assigns the manager a first-order role in shrinkage, of the same order as every permanent characteristic of the store combined. It is an accounting result, built on fixed effects; the sharpest confirmation of the economic content of the manager component is in the next section, which leverages the firm's rotation of managers directly and asks whether shrinkage moves when the manager moves, in an arrival-and-departure design.

5 What Do Good Managers Do?

Section 4 established that the general manager has an effect on the store's inventory shrinkage. This section exploits the high-frequency data on managers' actions—the minute-level time clock, the personnel record, and the store's operational panel—to ask what the managers whose fixed effect reduces shrinkage do differently inside the store.

Classifying managers. I classify a general manager as high-type (HT) if her AKM fixed effect for log shrinkage value—estimated in Section 4—lies below the median of the manager-effect distribution within her connected set, and as low-type (LT) otherwise. A lower fixed effect means less shrinkage under her management, so high-type managers are the better half of the distribution.

Design. One event design, applied in four forms, carries the whole section: the arrival and the departure of each manager type. An *arrival* occurs the first time a store comes under a manager of the given type for two consecutive months, and the first of those months is the event date; the two-month rule screens out temporary covers and end-of-month payroll overlaps. A *departure* is the mirror image: the first month the store is no longer under that type, immediately following a spell of it. Each design keeps only the stores that experience its event, so identification comes from the timing of events across stores, never from a comparison with stores that are never treated. For a store-month outcome Y_{jt} I estimate

$$Y_{jt} \;=\; \sum_{\substack{k=-8\\ k\neq -1}}^{8} \beta_k\,\mathbf{1}[t-E_j=k] \;+\; \varphi_j \;+\; \lambda_{m(t)} \;+\; \xi_{y(t)} \;+\; \varepsilon_{jt}, \tag{21}$$

where E_j is store j's event date, φ_j is a store fixed effect, $\lambda_{m(t)}$ and $\xi_{y(t)}$ are calendar-month and year effects, and the endpoint dummies bin event times beyond ± 8 months. Standard errors are clustered at the store. A companion difference-in-differences replaces the event-time dummies with a single $\text{Post}_{[0,8]}$ indicator for the first nine months after the event; that coefficient is the one reported in every table cell. Departures are far scarcer than arrivals, so the departure columns are estimated with less precision and read as suggestive.[8]

[8]The shrinkage event studies take the month before the event ($k=-1$) as the reference. For the presence and team outcomes the reference is the average of event months -8 to -2, with the handover month $k=-1$ omitted: in that month the outgoing manager is already gone and the incoming one not yet in place, so manager-based measures are mechanically depressed. Most events enter the sample already under the new manager type, so the individual pre-event months are identified off few stores and bounce accordingly; pooling the pre months into bins yields flat pre paths (equality of pooled pre bins is not rejected for any presence or inventory outcome).

Throughout, each table reports the same four rows—HT arrival, HT departure, then LT arrival, LT departure—and each event-study figure the same four panels: arrivals on the left, departures on the right, the high type above the low type.

Three findings organize the section, and each sharpens the last. First, the classification is not an accounting artifact: the managers it ranks best on shrinkage are the same managers who run stores with less turnover and less non-justified absence—outcomes recorded by payroll and the time clock, beyond any inventory process. Second, what a good manager changes is her own supervision, not the store: she is on the floor more days a month, nights and weekends included, and more of her workers' hours pass under her watch; the team she reworks through quits, replacement hires, and internal upward moves—no detectable rise in dismissals, headcount flat. Third, none of it shows in the personnel file: on pay and schooling, high- and low-type managers are indistinguishable—the firm cannot see its best managers in the records it prices.

5.1 Validating the Classification

If the classification measures an attribute of the person, it makes four predictions. Shrinkage should fall when a high-type manager arrives, rise when she leaves, rise when a low-type manager arrives, and show nothing sharp when a low-type manager leaves. Table 6 reports the four designs for log shrinkage value, and the data deliver all four. One caution governs the reading of this subsection: because managers are classified from the shrinkage fixed effect itself, these estimates are not independent measures of managerial quality. They are a *test*—shrinkage should move when, in which direction, and only when the type of the manager in place changes—and I read them as validation of the classification, never as fresh causal effects. Shrinkage falls 6.8 percent over the nine months after a high-type manager arrives and rises 6.5 percent after she leaves, reversing essentially the whole arrival reduction.[9] The low-type designs mirror it with less force: shrinkage rises 4.9 percent when a low-type manager arrives and moves insignificantly when she departs. Appendix C reports the same four designs for shrinkage-to-sales, with the same pattern.

[9] Effects on log outcomes are reported as $100 \times \hat{\beta}$. For the larger effects this slightly overstates the percent change: the 12.6 of Section 8, for instance, is an 11.9 percent fall.

Table 6: Shrinkage on manager arrivals and departures

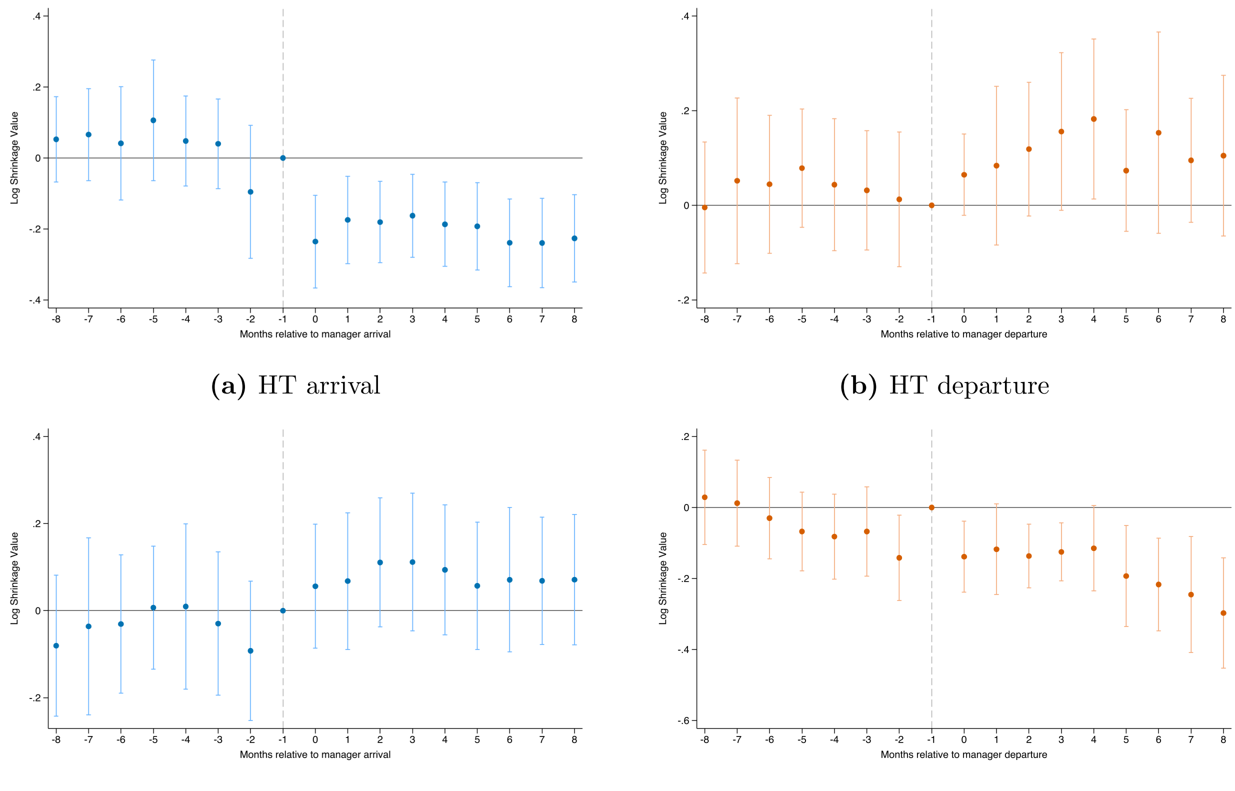
	Log shrinkage value
Arrival of a high-type manager	-0.0683^{***}
	(0.0146)
Departure of a high-type manager	0.0653^{*}
	(0.0336)
Arrival of a low-type manager	0.0489^{***}
	(0.0176)
Departure of a low-type manager	-0.0343
	(0.0305)
Mean	12.43
Median	12.35

Each cell is the $\text{Post}_{[0,8]}$ coefficient from Equation (21) collapsed to a single post indicator, with store, calendar-month, and year fixed effects and standard errors clustered at the store, on the winsorized log value of inventory losses. Each design keeps the stores that experience its event. The departure designs rest on far fewer events than the arrivals and are read as suggestive. Mean and Median are of the log outcome. $^{***}p<0.01$, $^{**}p<0.05$, $^{*}p<0.1$.

Figure 3 traces the four paths, and the high-type arrival is the sharpest object in the figure: the eight months before arrival are flat and indistinguishable from zero, and shrinkage drops by about 0.24 log points—roughly a fifth—in the month the manager arrives and stays down. That shape is what the classification predicts if it tags an attribute of the person. A fixed store characteristic cannot produce an on-impact break, and a mean-reverting shock cannot stay flat beforehand and then fall at the change. The high-type departure builds in the opposite direction on a flat pre-departure path; the low-type panels show the rise on arrival and nothing sharp on departure.

If anything, the table understates the size of the break. Under staggered timing the pooled coefficient of Table 6 is a conservative summary: stores treated earlier serve as controls for stores treated later while their shrinkage is still depressed, which attenuates the estimate (Goodman-Bacon, 2021). Restricting the same regression to the event window puts the fall in shrinkage after a high-type arrival at 19 percent (not tabulated), in line with the depth of the event-study path.

Log shrinkage value



(a) HT arrival

(b) HT departure

(c) LT arrival

(d) LT departure

Figure 3: Event studies: log shrinkage value around manager arrivals and departures

Notes: Coefficients β_k from Equation (21); the month before the event ($k = -1$) is the omitted reference. Each panel is one of the four designs of Table 6, on its own event sample. Bands are 95 percent confidence intervals; standard errors clustered at the store.

Beyond shrinkage. One worry remains: does the classification reflect better recording or a lasting change in behavior? A fixed effect estimated from shrinkage alone might reward a manager who *records* inventory better rather than one who changes conduct in the store. The two look identical in the shrinkage data, but they separate everywhere else, so I test the classification against outcomes it was not built from. I re-estimate the decomposition of Equation (19) separately for five additional outcomes—two alternative shrinkage definitions, turnover, and two non-justified-absence measures, the latter three recorded by payroll and the time clock, outside any manager's inventory paperwork—and compare each manager's fixed effects across outcomes. Figure 4 is the scatterplot matrix. The three shrinkage-based effects are strongly correlated (0.63–0.85), and the shrinkage-value effect also comoves with the manager effects in turnover (0.38) and non-justified absence (0.65–0.82); high-type managers carry better fixed effects in every other outcome—strongly so for the shrinkage normalizations and turnover ($p < 0.001$), more weakly for the two absence measures ($p = 0.06$–0.09). A recording skill does not lower absenteeism. The classification captures a broad dimension of managerial quality, in line with the manager-specific productivity effects of Lazear et al. (2015) and Bertrand and Schoar (2003). The arrival effect is also graded rather than binary—it scales with how much the arriving manager improves on the incumbent and vanishes for arrivals that are not upgrades—a dose response and novelty placebo reported in Appendix D.2.

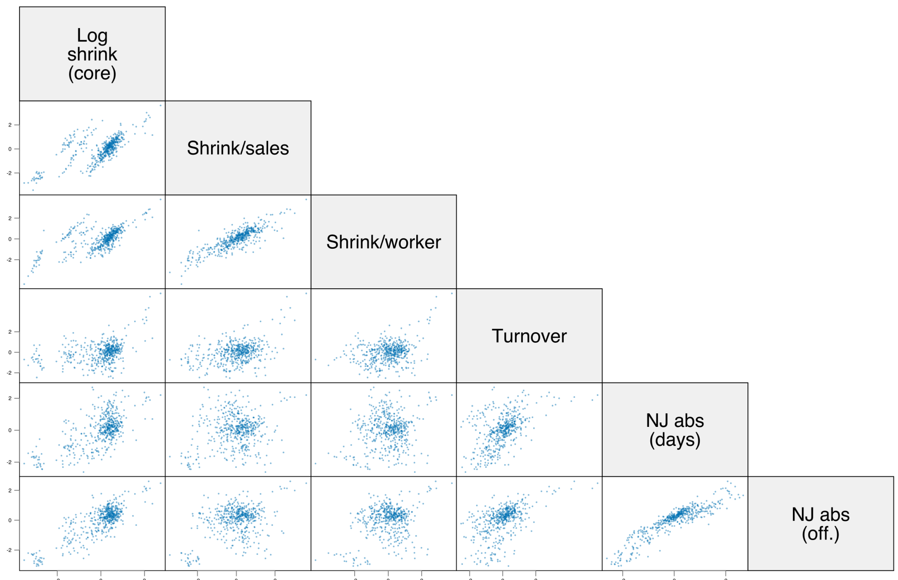


Figure 4: The shrinkage classification captures general managerial quality

Notes: Scatterplot matrix of each manager's AKM fixed effects across the six outcomes, each estimated from a separate decomposition (Equation 19) and standardized to standard-deviation units; the core measure, log shrinkage value, is the first row and column. Higher values mean more shrinkage, turnover, or absence, so the positive comovement in every panel shows that a manager who controls shrinkage well also runs a store with less turnover and absence. Pairwise Pearson correlations range from 0.38 (turnover) to 0.85 (across shrinkage definitions). Appendix D.7 presents the same comovement pairwise against the core effect.

Nor is the classification picking up something obvious and observable. Comparing the two types directly—one observation per manager-store spell, testing the difference in means trait by trait—high- and low-type managers are statistically indistinguishable on the observables a hiring decision prices: pay (a difference of 1.5 log points, $p = 0.41$) and schooling (0.16 years, $p = 0.39$). They differ only marginally along career-path dimensions a firm could not read as competence in advance (Appendix D.6 reports every comparison). On paper, the good manager is ordinary: what the fixed effect captures is not a credential the firm already prices, which is precisely why identifying good managers from personnel records alone is hard.

The outcomes in the rest of the section—presence and the team—are not built from the shrinkage fixed effect, so the designs there carry no such circularity. With the classification validated, the rest of the section turns to the economic content: what the good manager actually changes in the store.

5.2 Co-Presence: Management as Monitoring

What does the good manager do with the store once she has it? The one input she most directly controls is her own time on the floor, so I start there. I measure it from the minute-level time-clock records. From these I construct, at the store-month level, the manager's own shifts—total, weekend, and night—and *co-presence*: the share of frontline worker shift-hours worked while the store's general manager is simultaneously on the clock, computed by overlapping each worker's shift with the manager's daily presence interval.

Table 7 reports the four designs for the five measures, and the high-type arrival column tells a simple story: the manager is there more. She works 1.5 more shifts per month on a base of 23.6, with weekend shifts up 0.39 on a base of 7.0 and night shifts up 0.66 on a base of 6.0—increases of five to eleven percent. She scales her existing schedule up rather than recomposing it: her weekend share and her arrival hour barely move, and the store's night mix is flat (Appendix D). And because she is present more while the store operates no differently, the frontline ends up supervised more directly: co-presence rises 3.2 percentage points on a base near 57 percent. None of this happens when a low-type manager arrives: she moves none of the four measures. The departure columns all shrink toward zero and are too imprecise to sign—consistent with the presence margin switching off when she goes, though not decisive on its own; the sharper reversal is on shrinkage itself (Table 6).

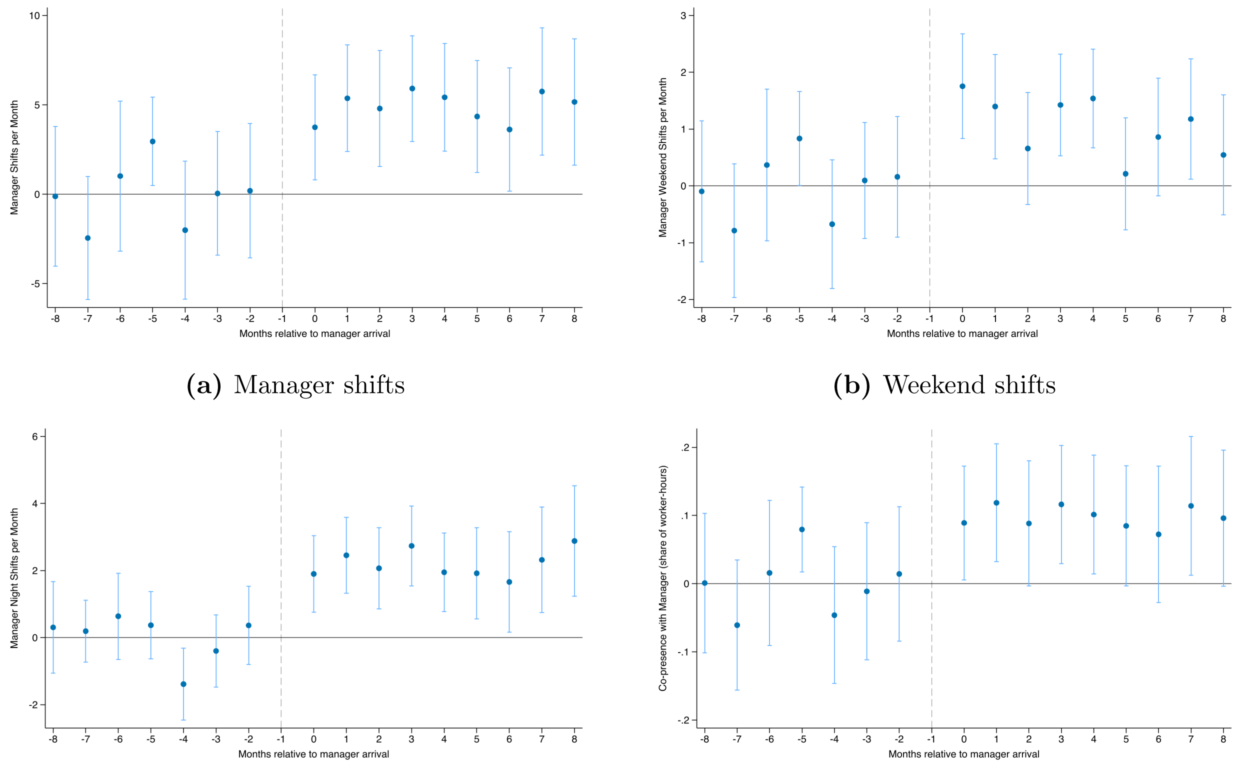
This is why I read the manager's presence as the monitoring technology itself. High-type managers do not buy more monitoring: they add no loss-prevention, supervisory, or security staff (Appendix D). They *are* present more, and their workers keep better time when they are (Appendix D.5); the added presence is broad rather than targeted at particular groups (Appendix D.4). Presence is the empirical content this paper gives the monitoring term μ_m: the manager makes the dismissal threat credible by being there to notice, connecting the result to the classic evidence that monitoring deters opportunism (Nagin et al., 2002) and that workers lift effort when observed (Mas and Moretti, 2009).

Table 7: Managerial presence and manager–worker co-presence

	(1) Manager shifts/month	(2) Manager weekend shifts	(3) Manager night shifts	(4) Co-presence (worker-hrs)
HT manager arrival	1.5380***	0.3923**	0.6565**	0.0324***
	(0.4739)	(0.1583)	(0.2741)	(0.0121)
HT manager departure	−0.3205	−0.1342	−0.5222	−0.0238
	(1.6438)	(0.5279)	(0.6668)	(0.0385)
LT manager arrival	0.3250	0.1401	0.4689	0.0039
	(0.7416)	(0.2325)	(0.3773)	(0.0160)
LT manager departure	0.2035	−0.0832	0.3389	−0.0038
	(1.0316)	(0.3287)	(0.5016)	(0.0276)
Mean	23.58	6.99	6.04	0.575
Median	26.00	8.00	5.00	0.654

Each cell is the $\text{Post}_{[0,8]}$ coefficient from the design of Table 6 for the row's event, estimated on store-month measures built from shift-level time-clock records. "Manager shifts" counts the general manager's worked shifts in the month; "Manager weekend (night) shifts" the subset on Saturday–Sunday (night hours). "Co-presence" is the mean share of a worker shift's hours overlapping the manager's daily presence interval. The departure designs rest on far fewer events and are read as suggestive. Mean and Median are of the dependent variable. $^{***}p<0.01$, $^{**}p<0.05$, $^{*}p<0.1$.

Figure 5 traces each measure around the high-type arrival: flat before, up on impact, holding through the window. The departure and low-type designs are in Appendix D.



(a) Manager shifts

(b) Weekend shifts

(c) Night shifts

(d) Co-presence

Figure 5: Event studies around the high-type arrival: presence

Notes: Event studies around the high-type manager arrival, one panel per outcome, from Equation (21). Coefficients are relative to the average of event months −8 to −2; the handover month ($k = -1$) is omitted. The sustained post-event levels sit above the pooled coefficients of the tables because the pooled comparison keeps months beyond the event window inside the comparison group. Bands are 95 percent confidence intervals, clustered at the store. The departure and low-type designs are in Appendix D.

5.3 The Team: Recomposition Without a Purge

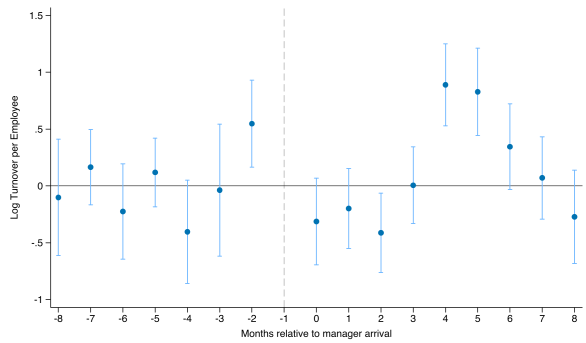
Does the good manager change *who* works in the store, or how? Table 8 keeps the four margins that answer it. Turnover rises 17 percent over the nine months after a high-type arrival, yet team size does not move: the churn is a recomposition, not a purge—the rise is quits (+0.17) and one-for-one replacement hires (+0.21), with no detectable rise in dismissals (+0.003, standard error 0.079; Appendix Table 19 reports the full decomposition). What she does with the workers she keeps is modest and internal: reassignments to a higher tier rise by 0.19 per hundred workers on a base of 1.0 and reassignments across functional areas by 0.21 on a base of 1.4, while the composition of the team—its diversity of age, tenure, education, gender, tier, and pay—is untouched (Appendix D), and the cross-area detail by functional area is in Appendix D.3. A low-type arrival moves none of these margins, and the departure designs are null throughout, with one exception consistent with the same picture: upward and cross-area reassignments rise after a low-type manager leaves, when better management arrives.

Table 8: The team on manager arrivals and departures

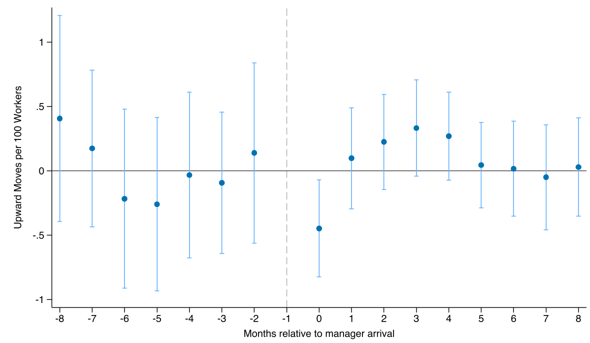
	(1) Log turnover (p.e.)	(2) Upward moves (per 100)	(3) Cross-area moves (per 100)	(4) Log team size
HT manager arrival	0.1738**	0.1920***	0.2076**	0.0074
	(0.0724)	(0.0651)	(0.0920)	(0.0108)
HT manager departure	0.0519	0.1260	0.2021	−0.0081
	(0.1896)	(0.1608)	(0.1993)	(0.0248)
LT manager arrival	0.0245	−0.0048	0.0286	−0.0089
	(0.0610)	(0.0702)	(0.1134)	(0.0091)
LT manager departure	0.0279	0.3125**	0.5081***	0.0065
	(0.1486)	(0.1423)	(0.1896)	(0.0193)
Mean	−3.32	1.01	1.44	4.12
Median	−2.80	0.00	0.79	4.03

Each cell is the $\text{Post}_{[0,8]}$ coefficient from the design of Table 6 for the row's event. Turnover is the log of the monthly separation rate per employee; upward moves are within-store reassignments to a higher tier and cross-area moves are reassignments across functional areas, both per hundred workers; the full decomposition into quits, dismissals, and hires is in Appendix Table 19. The departure designs rest on far fewer events and are read as suggestive. Mean and Median are of the dependent variable. $^{***}p<0.01$, $^{**}p<0.05$, $^{*}p<0.1$.

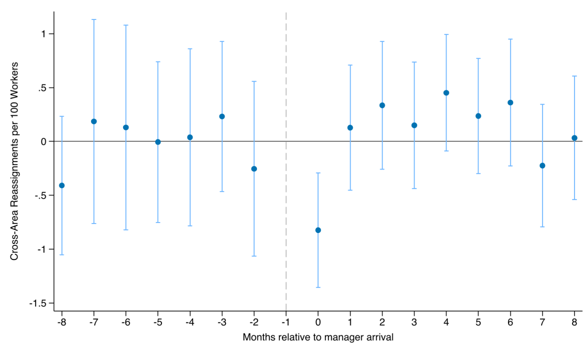
Figure 6 traces the four margins around the high-type arrival: turnover spikes in the transition months and fades, the reallocation margins are flat to modestly positive, and team size stays on its pre-event level throughout; the remaining designs are in Appendix D.



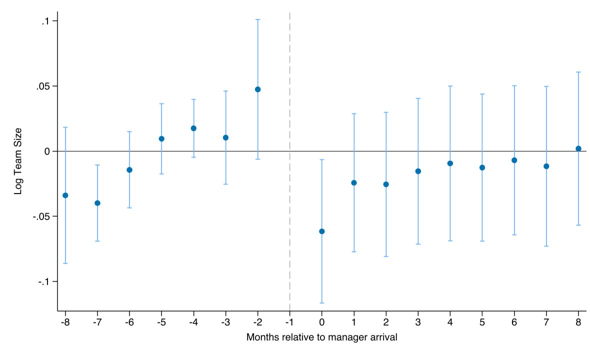
(a) Log turnover



(b) Upward moves



(c) Cross-area moves



(d) Log team size

Figure 6: Event studies around the high-type arrival: the team

Notes: One panel per outcome, from Equation (21); conventions, reference period, and bands as in Figure 5.

Section 8 brings this monitor together with the wage premium and tests whether the two reinforce each other.

6 Wage Premia and Shirking: Two-Way Fixed-Effects Evidence

Sections 4 and 5 measured one factor of the model's deterrent, the monitor. This section turns to the other, the rent. The model in Section 3 predicts that shirking falls as the wage premium rises; this section tests that prediction on the shirking measures of Section 2.

Because the firm sets pay by position, the premium a store pays moves with the outside wage of its municipality, which varies widely across Mexico and over time (Appendix Figure 18). I use this variation in a two-way fixed-effects model at the store level, the level of the paper's central outcome:

$$\log Y_{it} \;=\; \delta \, \log W^{p}_{it} \;+\; Z_{it}\,\theta \;+\; \eta_{m(i)} \;+\; \tau_t \;+\; \nu_{it}, \tag{22}$$

where Y_{it} is a shirking or effort outcome of store i in quarter t: inventory shrinkage, absenteeism (total and non-justified), turnover, monthly shifts, or daily worked hours; W^{p}_{it} is the store wage premium, the average of the worker premium in Equation 1 across the store's workers; $\eta_{m(i)}$ is a municipality fixed effect; τ_t is a year-by-quarter fixed effect; and Z_{it} optionally includes store format and log sales. Standard errors are clustered by municipality. The fixed effects absorb permanent differences across municipalities and shocks common to all stores in a quarter, so δ rests on movements in the local outside wage rather than on its level.[10] The model predicts a negative δ for the shirking outcomes and a positive δ for shifts and hours.

Tables 9 and 10 report the estimates. Shrinkage, absenteeism, and turnover fall with the premium; monthly shifts and daily hours rise. Every gradient keeps its sign and significance when I add store format and log sales and when I replace the year-by-quarter fixed effects with separate year and quarter effects (Table 9, columns 2–4; Appendix Table 29).

[10] δ captures how shirking covaries with the rent under two conditions: the firm's pay-setting does not respond to a store's own discipline—ruled out by the position-level wage policy—and outside-wage movements are uncorrelated with local determinants of shirking beyond what municipality and time fixed effects absorb. The second is the demanding one: a local boom that lifts outside wages could also shift crime, staffing, or the mix of hires. Municipality and year-by-quarter fixed effects absorb permanent local differences and common national shocks, and Section 7 re-estimates the same relation off statutory minimum-wage hikes that move the premium by law, on a known date, for reasons unrelated to any store's conditions.

Table 9: Store-level TWFE: inventory shrinkage (Log–Log)

Wage premium (Log)	(1)	(2)	(3)	(4)
Coefficient	-0.113^{***}	-0.114^{***}	-0.045^{**}	-0.111^{***}
SE	(0.032)	(0.033)	(0.021)	(0.031)
Store format	No	Yes	Yes	No
Log sales	No	No	Yes	No
Municipality FE	Yes	Yes	Yes	Yes
Year×Qtr FE	Yes	Yes	Yes	No
Year + Quarter FE	No	No	No	Yes
Mean	9.49	9.49	9.49	9.49
Median	9.43	9.43	9.43	9.43

Store-quarter level. The dependent variable is the log value of inventory losses in USD; Mean and Median are of the log outcome. Standard errors clustered at the municipality level. $^{***}p<0.01$, $^{**}p<0.05$, $^{*}p<0.1$.

Table 10: Store-level TWFE: attendance, turnover, and effort (log wage premium)

Wage premium (Log)	Absenteeism	Non-justified absenteeism	Turnover	Monthly shifts	Daily hours
Coefficient	-0.036^{***}	-0.129^{***}	-0.158^{***}	$+0.032^{***}$	$+0.012^{***}$
SE	(0.010)	(0.023)	(0.025)	(0.009)	(0.002)
Store format	No	No	No	No	No
Log sales	No	No	No	No	No
Municipality FE	Yes	Yes	Yes	Yes	Yes
Year×Qtr FE	Yes	Yes	Yes	Yes	Yes
Mean	0.68	0.00	2.55	2.98	2.10
Median	0.71	0.06	2.69	2.99	2.10

Store-quarter level. Each column is a separate regression of the log outcome on the log store wage premium, in the baseline specification of column 1 of Table 9; Appendix Table 29 adds store format, log sales, and separate year and quarter effects. Turnover is a rate per hundred workers. Mean and Median are of the log outcome. Standard errors clustered at the municipality level. $^{***}p<0.01$, $^{**}p<0.05$, $^{*}p<0.1$.

The paper runs this exercise at two levels. The store-level version above compares stores that pay different premia within a municipality and quarter. The worker-level version estimates the same equation on the worker–quarter panel, with a worker fixed effect in place of the municipality effect and a store fixed effect added, so δ is identified from changes in a worker's own premium as the outside wage of her cell moves; the discrete outcomes (any absence, any non-justified absence, separation) are estimated with a logit. Inventory shrinkage is a store outcome and does not enter this exercise.

The two levels tell the same story (Appendix E, Tables 30–32). Within the same worker, a larger premium goes with fewer absence days, total and non-justified, with lower odds of any absence, of any non-justified absence, and of separation, and with more monthly shifts and longer daily hours; every estimate keeps its sign and significance under controls for managerial tier, tenure, and age and under the choice of time fixed effects. Table 32 sets the two levels side by side for the four margins observed at both: the signs agree on every margin, and the magnitudes are close for absence days, shifts, and hours, while the non-justified absence gradient is larger at the store level than within the worker. That pattern is what one expects if the store coefficient adds a composition component to the individual response: the store premium is an average over the store's positions and workers, and stores that pay a larger premium also differ in size and mix, so part of the store gradient can reflect who works there rather than how each worker behaves. The worker fixed effect removes every permanent attribute of the worker and of her store, and the response survives. The store-level estimates can therefore be read as the aggregate of individual responses to the premium, with the worker-level exercise as the check that the aggregate is not sorting.

The shrinkage gradient is also robust to normalizing shrinkage by sales, in value and in physical units (Appendix Table 28).

7 Minimum-Wage Hikes at the Border: Moving the Premium by Statute

Section 6 estimated the gradient between shrinkage and the wage premium off movements in the local outside wage. Its residual threat is a local shock that moves outside wages and shrinkage together. This section moves the outside wage by statute instead. A minimum-wage hike raises the wage the law guarantees a worker elsewhere by an amount and on a date published in advance, for reasons unrelated to any store, while the firm's pay scale, set centrally by position, does not move with it. What the hike moves is the premium the firm pays over the legal floor, and it moves it more in one part of the firm than in the other.

The exercise is an event study around the January hike. Mexico keeps two statutory floors, a general one and a higher one along the northern border, and each January both rise by the same percentage, so the step in pesos is about 1.5 times larger at the border. Because the firm's scale is the same on both sides, the border store's premium over its floor compresses by more than the interior store's. The border stores are the treated group, the interior stores of the same format are the control group, and the event is the January in which the new floor takes effect. I follow both groups from six months before to six months after each hike and ask whether border shrinkage rises relative to interior shrinkage once the floor moves.

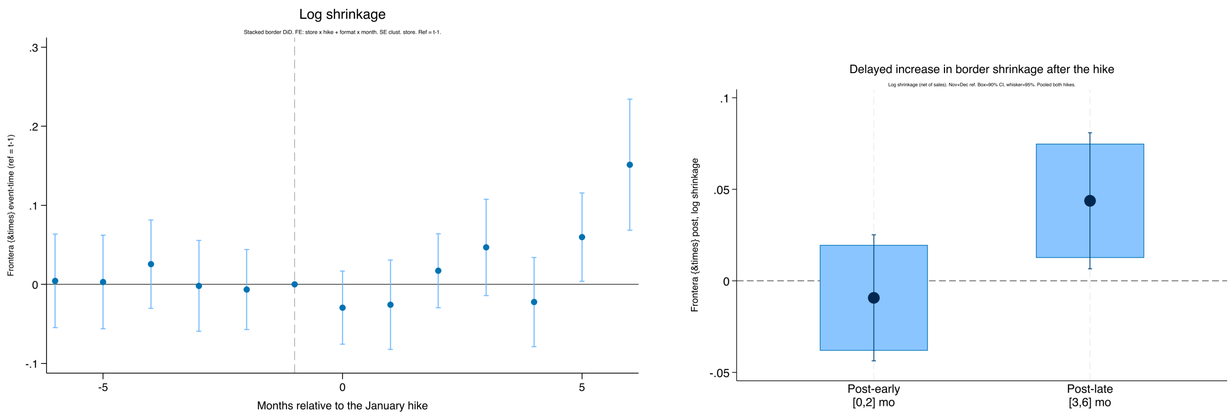
The border floor is recent. The government created the *Zona Libre de la Frontera Norte* in January 2019, and since 2020 the two floors have risen by the same percentage every January, which holds their ratio near 1.51. The hikes of January 2023 and January 2024 raised both floors by twenty percent, from permanently different bases (CONASAMI). About a quarter of the firm's stores lie inside the border zone.

I estimate, for store k of format f in hike h at event month τ,

$$Y_{kh\tau} \;=\; \alpha_{k,h} \;+\; \gamma_{f(k),m} \;+\; \sum_{s\neq -1} \beta_s \left(\text{Border}_k \times \mathbf{1}[\tau = s]\right) \;+\; \varepsilon_{kh\tau}, \tag{23}$$

where $Y_{kh\tau}$ is log shrinkage net of sales and the two hikes are stacked as separate experiments (Cengiz et al., 2019). The store-by-hike fixed effect $\alpha_{k,h}$ compares each store with itself within each hike, so no border–interior level difference enters the estimate. The format-by-calendar-month fixed effect $\gamma_{f(k),m}$ absorbs the seasonal pattern of each store format, so a border store is set against interior stores of its own format in the same month of the year. Border_k marks stores inside the border zone, $\mathbf{1}[\tau = s]$ marks event month s, and the month before the hike is the omitted category, so each β_s is the border–interior gap in month s relative to December. Standard errors are clustered by store. The two-phase summary replaces the event-month indicators with two, $\text{Impact}_{[0,2]}$ and $\text{Compression}_{[3,6]}$, measured against the two months before the hike. The comparison follows the border-comparison tradition of the minimum-wage literature (Dube et al., 2010).

Figure 7 reports the estimates. The pre-hike coefficients are flat, border shrinkage is unchanged on impact (-0.009, $p = 0.60$), and it rises by 4.4 percent in the compression phase ($p = 0.022$). The increase is delayed: it arrives as the compression persists, not on the date the floor moves.



(a) Event study.

(b) Two-phase summary.

Figure 7: The delayed shrinkage response to the border minimum-wage hikes.

Notes: Border $\times$ event-month coefficients of Equation 23 on log shrinkage net of sales, both hikes stacked, with the month before the hike as reference (panel a), and the two-phase summary against the two months before the hike (panel b). Store-by-hike and format-by-calendar-month fixed effects; whiskers are 95 percent confidence intervals, standard errors clustered by store.

Figure 8 shows the first stage for the January 2023 hike: the same event study at the worker level, with the worker's premium over her statutory floor, in pesos per day, as the outcome, a worker fixed effect in place of the store effect, and a position-by-format-by-month fixed effect. The dashed line marks the full statutory squeeze, the difference between the two floors' steps.

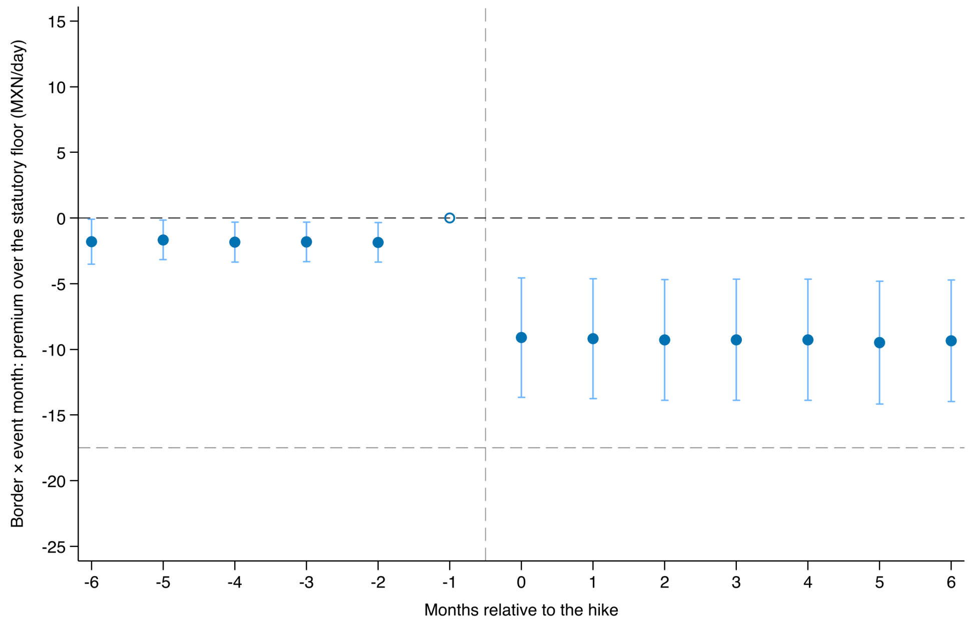


Figure 8: First stage: the worker's premium over the statutory floor, January 2023 hike.

Notes: Border × event-month coefficients on the contractual daily wage minus the statutory floor of the worker's zone, in pesos per day, for non-managerial full-time workers present throughout the window; the month before the hike is the omitted category and is drawn as a hollow marker. Worker and position-by-format-by-calendar-month fixed effects; whiskers are 95 percent confidence intervals, standard errors clustered by store. The dashed horizontal line marks the full statutory squeeze, the difference between the two floors' steps.

Whether the damage depends on the manager on the floor when the shock lands I take up in Section 9.

8 Managers and the Wage Premium: Complements in Discipline

Are the wage premium and the manager's monitoring *complements* or *substitutes* in producing discipline? The question decides how a firm should deploy the two most expensive instruments it holds against its losses. If pay and supervision are substitutes, they are interchangeable: a firm can pay less where its managers are strong, lean on wages where managerial talent is scarce, and budget the two policies apart. If they are complements, neither instrument works alone: the best managers belong in the stores whose jobs are worth most, cutting the premium quietly disables the manager, and any shock that compresses the rent damages a store through its monitor. The two inputs are almost always studied one at a time, so which regime holds is an open empirical question—and this setting can answer it, because Section 5 delivers a validated manager margin and the firm's position-based pay delivers wide variation in the premium a store carries, fixed before any manager acts on it. This section crosses the two and tests whether they multiply.

8.1 Design

To answer it, I take the high-type arrival design of Section 5 and ask one question of it: is the shrinkage reduction a high-type manager delivers larger where the store's wage premium is high?

I classify each store's municipality, in the month its manager arrives, as a high-wage-premium (HWP) or low-wage-premium (LWP) environment—HWP if the premium the firm pays over the local ENOE market wage lies above that municipality's own median over the sample, LWP otherwise. Fixing the environment at arrival makes it pre-determined with respect to everything that follows and constant within a store's event window. I then re-estimate the arrival difference-in-differences of Equation (21), interacting the $\mathrm{Post}_{[0,8]}$ indicator with this environment, separately for high- and low-type arrivals.

The complementarity hypothesis makes one prediction: the high-type shrinkage reduction should be larger—more negative—in high-premium stores than in low-premium stores. If the rent and the monitor were independent deterrents, that difference would be zero; if they are the two factors of one product, it is negative.

8.2 Good Managers Accomplish More Where the Premium Is High

Table 11 reports the estimates for log shrinkage value, high-type arrivals in the upper panel. A high-type manager arriving in a high-premium environment cuts shrinkage by 12.6 log points—about 12 percent—roughly three times the 4.3 the same arrival produces in a low-premium environment. The difference between the two environments, the complementarity itself, is 8.4 percentage points, significant at five percent. Where the premium gives the manager's dismissal threat something to confiscate, the same manager accomplishes more. The ratio measures behave the same way, and two further margins in the full battery point in the same direction: non-justified absence does not respond to the manager in either environment—consistent with Section 5, the manager margin operates on the inventory discipline she polices, not on the attendance a worker internalizes directly through her own rent—and turnover rises after a high-type arrival only where the premium is high, a first sign of the enforcement margin the mechanism runs through (Appendix H, Tables 39–40).

Table 11: Manager arrivals by wage environment at arrival

	Log shrinkage value
HT arrival, high-premium (HWP)	-0.126^{***}
	(0.032)
HT arrival, low-premium (LWP)	-0.043^{**}
	(0.018)
Difference (HWP − LWP)	-0.084^{**}
	(0.037)
LT arrival, high-premium (HWP)	0.037
	(0.033)
LT arrival, low-premium (LWP)	0.059^{***}
	(0.021)
Difference (HWP − LWP)	-0.022
	(0.039)
Mean	12.43
Median	12.35

$\text{Post}_{[0,8]} = 1$ for event times 0 through 8, interacted with the wage environment of the store's municipality in the arrival month (HWP = premium above the municipality's median; LWP otherwise). Each panel reports the arrival effect in each environment and then their difference (the complementarity); high-type arrivals above the rule, low-type below. Store, calendar-month, and year fixed effects; standard errors clustered at the store; ever-treated stores only. Classification, event definition, and outcome as in Table 6; the ratio measures and the remaining outcomes are in Appendix H. Mean and Median are of the log outcome. $^{***}p<0.01$, $^{**}p<0.05$, $^{*}p<0.1$.

The differential is not circular. Because managers are classified from their fixed effects on shrinkage, the arrival effect on shrinkage itself is read as a validation of the classification (Section 5), and I lean no further on it. The differential across wage environments is a different object. The environment is classified from wage data alone—the premium against the municipality's own median, fixed in the month the manager arrives—so it is predetermined and owes nothing to shrinkage or to the manager's label. Whether the label's imperfections net out of the difference between environments is something I measure rather than assume, three ways. A label rebuilt from an outcome that never touches shrinkage—the manager's fixed effect on non-justified absence, recorded by the time clock—returns a differential of −6.0 percentage points, significant at five percent. Permuting residuals within store and re-running the entire pipeline—estimation, labeling, and event definition—one thousand times shows that the mechanics alone generate −2.4 points; the observed differential lies outside the full placebo distribution, and netting out the mechanical component leaves -6.2 points. And a label estimated only on odd months and tested only on even months shares no residual between the labeling and the test, and returns -6.8 to -8.4 points. Appendix I reports the full battery, including a reclassification stress test that preserves 88 to 93 percent of the magnitude at the same significance, and Appendix J the arithmetic behind the mechanical component.

Three consequences organize how the rest of the section reads. First, in Table 11 the row to read causally is the *difference*: the per-environment rows are levels, and they share the validation status of Section 5; the difference row is the object the tests above protect. Netting the measured mechanical component from it still leaves roughly six points—the same magnitude the shrinkage-free key returns—so the complementarity does not rest on the mechanics. Second, the low-type panel carries the same protection: its null differential—the premium does not rescue a weak monitor—is a statement about the difference, not about the levels. Third, the mechanism margins of Section 8.3 are not built from the shrinkage fixed effect at all, so their environment differentials—the suspensions that rise only where the rent is high—face no circularity to difference out; they corroborate the complementarity from outside the classification.

Figure 9 shows the same split as event studies.

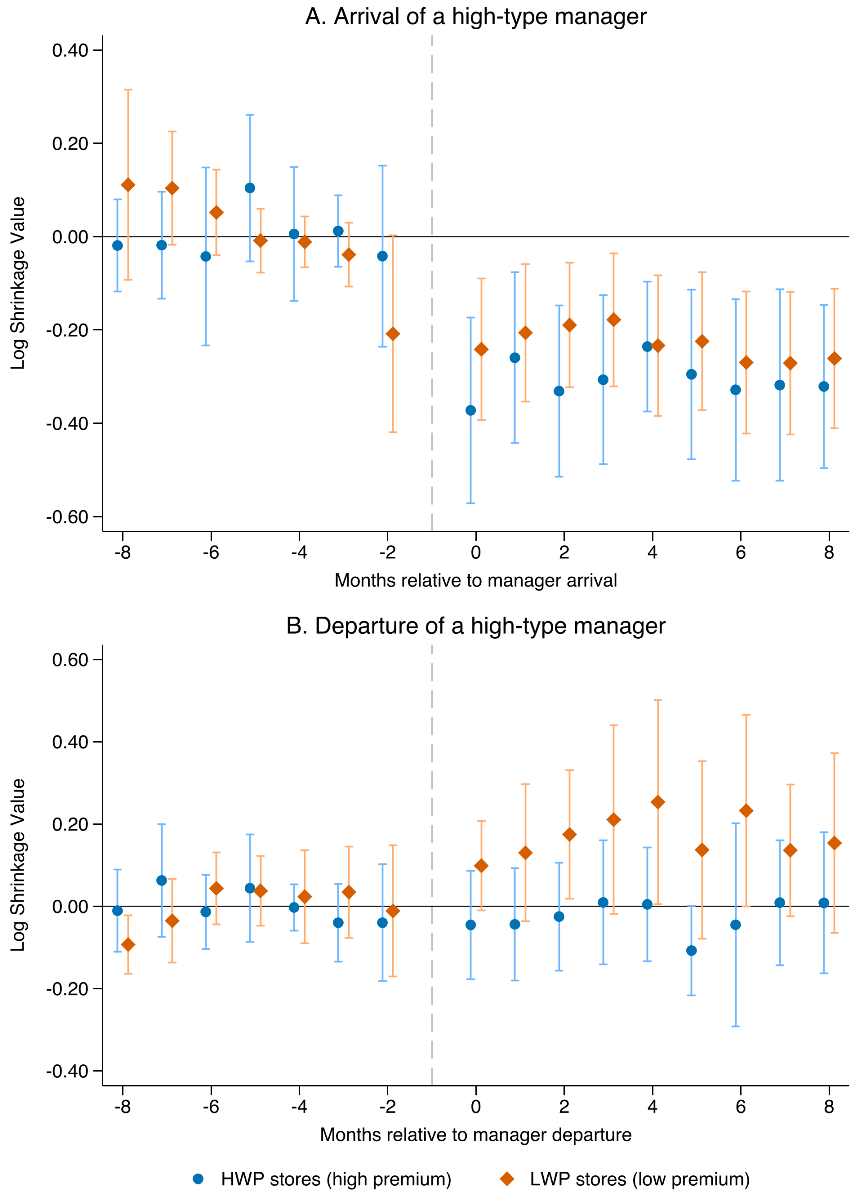
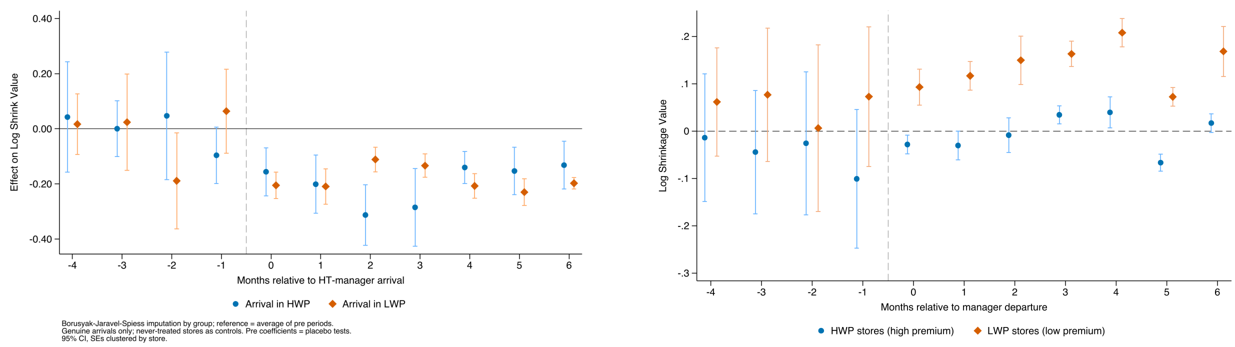


Figure 9: Arrival and departure of a high-type manager, by wage environment

Notes: Event studies of log shrinkage value around a high-type manager's arrival (panel A) and departure (panel B), estimated separately for high-premium (HWP) and low-premium (LWP) store environments, with the environment fixed at the event month as in Table 11. Coefficients are relative to the average of event months -8 to -2; the handover month $t = -1$ is omitted from the baseline and from the plot. Bands are 95 percent confidence intervals, clustered at the store. The departure split rests on few events per environment and is descriptive only.

The interaction is a two-way fixed-effects design, and staggered arrival timing raises the usual concern that already-treated stores serve as controls. I therefore re-estimate the arrival dynamics separately for high-type arrivals in high- and low-premium environments with the imputation estimator of Borusyak et al. (2024), on the genuine (non-left-censored) arrivals with never-treated stores as controls. Figure 10 plots the two paths, and two features matter. The pre-arrival placebos are flat and centered on zero in both environments, so the corrected design is valid; and on impact and through the post-window the high-premium path lies below the low-premium path—the same ordering the binary table reports, now under a heterogeneity-robust estimator. I do not read the gap as separately identified: with few genuine high-premium arrivals, the two series' confidence intervals overlap, so the powered test of the differential remains the full-sample per-type interaction of the main text (Table 11). What the figure adds is that the direction of the complementarity survives the correction for staggered timing.



(a) Arrival of a high-type manager

(b) Departure of a high-type manager

Figure 10: High-type arrival and departure dynamics by wage environment (Borusyak–Jaravel–Spiess imputation)

Notes: Imputation event study (Borusyak et al., 2024) estimated separately by wage environment fixed at the event month; reference = average of pre-treatment periods; never-treated stores serve as controls; genuine (non-left-censored) events only. Pre-event coefficients are placebo tests. Panel (b) rests on far fewer events per environment and is descriptive. 95 percent confidence intervals; standard errors clustered at the store.

The binary split has a continuous companion (Appendix Figure 20). It replaces the high/low dichotomy with the standardized premium itself, plotting the marginal effect of a high-type arrival on log shrinkage as a smooth function of the premium the store's municipality pays at the moment the manager arrives. The line slopes steadily downward: the arrival effect is a modest −4 percent where the premium is low and reaches −11 percent where it is high, tracing out as a gradient the same contrast the table reports as two cells. The complementarity is an artifact neither of where the median falls nor of how the environment is defined: re-estimating the high-type interaction under three additional definitions, all fixed at arrival, preserves its sign and the ordering of the two environments in every case (Appendix H.1).

The high-type result shows that the premium amplifies a good manager. The model speaks to bad managers as well. If the rent disciplines the shrinkage margin only through a monitor who makes detection credible, then under a low-type manager, who supplies little supervision, no premium the firm pays has anything to threaten on that margin. A low-type arrival should therefore raise shrinkage in *both* environments. Whether the damage scales with the premium the model does not settle: on the interior the same cross-partial that amplifies a good manager would make a bad one costlier where the premium is high (Section 3.5), so the comparison across environments below is a description, not a test.

The lower panel of Table 11 confirms it. A low-type arrival is followed by a shrinkage *increase* of 5.9 percent in low-premium stores, and the estimate in high-premium stores carries the same sign and a similar magnitude, with no significant difference between environments; the ratio measures repeat the pattern (Appendix H, Table 40). A bad manager is bad in every wage environment; the premium does not rescue the store, and the difference between environments is not the object the model signs. This result does not make the premium useless under a bad manager—it continues to discipline the margins the worker controls directly, attendance and retention—but on the margin the manager must police, the rent works only through the monitor. In levels, then, the two inputs are complements: the premium amplifies what a good manager accomplishes and cannot replace one.

Pooling the first arrival of either type into a single regression and estimating the full triple interaction reproduces the per-type message cell by cell—high-type in a high-premium environment is the only cell in which arrivals reduce shrinkage—with the triple-difference coefficient carrying the predicted negative sign but imprecise, because high-premium arrivals are scarce; Appendix H.3 reports the design, the table, and the cell breakdown; Figure 11 plots the four cells.

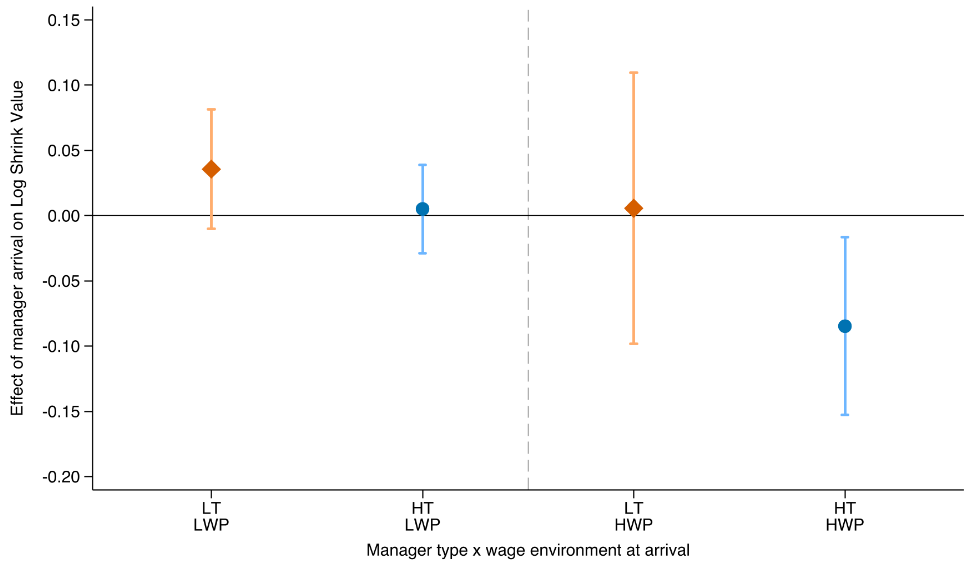


Figure 11: Effect of the first manager arrival on log shrinkage value, by manager type and wage environment at arrival

Notes: Cell effects from the pooled triple-interaction design via linear combinations; 95 percent confidence intervals, standard errors clustered at the store. Cell effects come from the pooled triple-interaction design on each store's first manager arrival (sample and definitions as in Appendix Table 42), so they need not match the per-type estimates of Table 11. In this design, only the high-type-in-high-premium cell delivers a significant reduction.

8.3 Does the Manager Do More Where the Premium Is High?

Does the manager *do* something different where the premium is high? The complementarity has two readings, and they part on her own conduct: a good manager could supply more discipline where the rent is large, or supply the same actions everywhere and have each action return more where the rent is large. To separate them I re-estimate the arrival design of Equation (21) with each action of Section 5 as the outcome—the presence and team margins exactly as that section reports them—interacting the post-arrival indicator with the wage environment at arrival exactly as in Table 11. Figure 12 puts the answer and the question side by side—the return by environment on the left, the differential of every action margin on the right—and Table 12 reports the same estimates margin by margin; the wider battery, including the separations decomposition, is in Appendix G.

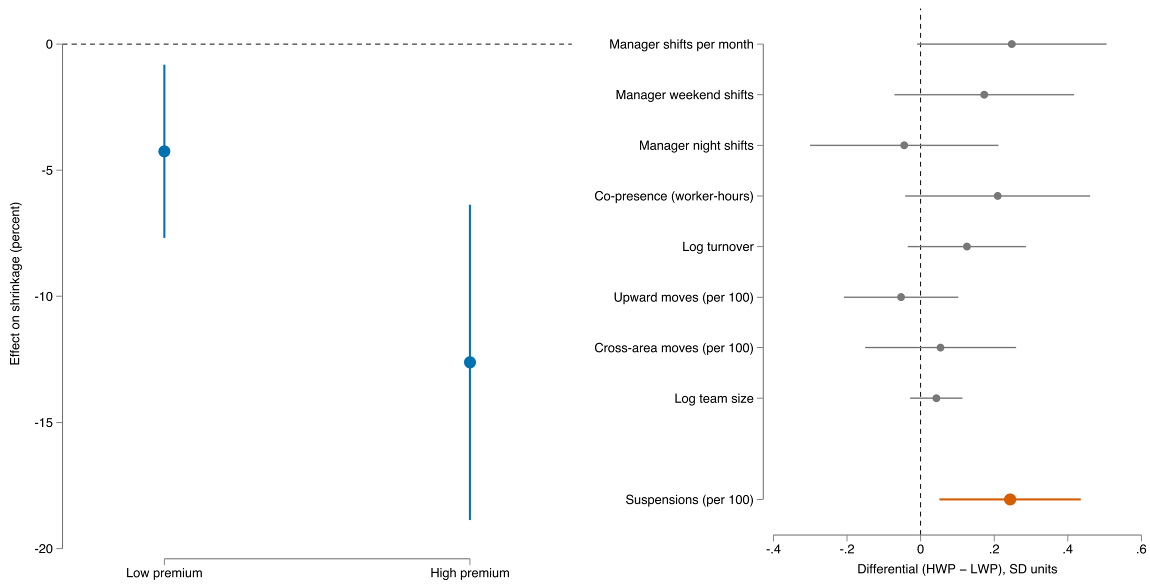


Figure 12: The gap is enforcement, not reorganization

Notes: Both panels come from the arrival design of Equation (21) with the post-arrival indicator interacted with the wage environment fixed at the month of the high-type arrival (store, calendar-month, and year fixed effects; standard errors clustered at the store; ever-treated stores). Panel (a): the effect of a high-type arrival on log shrinkage value in low- and high-premium stores, in percent—-4.3 against -12.6, a difference of 8.4 percentage points ($p = 0.02$). Panel (b): for each margin, the difference between the two environments, scaled by the standard deviation of that outcome so the margins are comparable; suspensions in orange. Bars are 95 percent confidence intervals.

The actions do not adjust in any clear way. In the right panel of Figure 12, most differentials are indistinguishable from zero; the few that approach significance—the manager's shift count—do so with intervals wide enough to admit effects of any economic size. This is a demanding question to put to the data: each differential splits a modest number of arrivals across two environments, so I read the action margins as undetermined rather than as nulls. What the data do show, sharply, is the enforcement margin and its yield—the one margin whose interval excludes zero. Suspensions rise by 0.31 per hundred workers after a high-type arrival in a high-premium store and *fall* by 0.14 in a low-premium one—a differential of 0.45, significant at five percent: the same arrival sanctions where the job is worth keeping and forgives where it is not. The return moves with it, in the left panel: the same arrival cuts shrinkage three times as much in the high-premium environment. On this evidence the mechanism is not a manager who reorganizes more where the rent is high; it is the same manager whose enforcement bites harder where dismissal costs the worker more—the pattern Proposition 4 predicts when the sanction is the manager's act and follows the rent: the same presence, sanctions where the job is worth keeping, and a larger fall in shrinkage there. Appendix G reports the wider battery of team-composition margins.

Table 12: Does the manager do more where the premium is high? Every margin crossed with the wage environment at arrival

	Effect in LWP stores	Effect in HWP stores	Difference (HWP − LWP)
Panel A. The manager's actions			
Manager shifts per month	0.916	3.042***	2.126*
Manager weekend shifts	0.252	0.745***	0.493
Manager night shifts	0.715**	0.497	−0.218
Co-presence (worker-hrs)	0.020	0.063***	0.043
Log turnover (p.e.)	0.115	0.309***	0.194
Upward moves (per 100)	0.208**	0.131	−0.077
Cross-area moves (per 100)	0.173	0.272*	0.098
Log team size	0.001	0.023	0.023
Panel B. Enforcement and the return			
Suspensions (per 100)	−0.140	0.311**	0.451**
Log shrinkage value	−0.043**	−0.126***	−0.084**

Each row re-estimates the arrival design of Equation (21) for the indicated store-month outcome, with the post-arrival indicator interacted with the wage environment fixed at the month of the high-type arrival (the municipality's premium above its own median): store, calendar-month, and year fixed effects; standard errors clustered at the store; ever-treated stores. Columns: the arrival effect in low- and high-premium stores and their difference. Rates are per hundred workers. $^{***}p<0.01$, $^{**}p<0.05$, $^{*}p<0.1$.

Taken together, the designs support the reading that managers and efficiency wages are complements *in levels*: the premium roughly triples what a high-type manager accomplishes and cannot substitute for one. It matches Coviello et al. (2022) from the opposite direction—there a rent matters only where monitoring is intense; here a monitor matters most where the rent is large—completing the levels half of the paper's central argument: the wage and the manager discipline together, each worth most where the other is present.

9 The Monitor at the Moment of the Shock

This closing section brings the manager to the border design, and asks whether the squeeze of Section 7 damages every border store, or only the stores where no one is watching the floor. Section 8 established the level form of that dependence. The marginal question is a different one, and Proposition 5 answers it: a small narrowing of the rent should move behavior only where the care margin is interior—where little monitoring is supplied—while the well-watched store, whose workers a present manager already holds at full compliance, is insured.

The design is the one already estimated. I take the two-phase border specification of Section 7 and interact the border-by-phase terms with the manager's floor presence at the moment the shock lands: border stores are split above or below the median presence within format-by-hike cells, measured in the months before each hike so that the hike itself cannot move the split. Each presence group is then compared with the same interior control stores, and the object of interest is the difference between the two compression responses.

Figure 13 reports the result, and the two responses trace opposite paths. Where the manager is on the floor, the raise pays off on impact—border shrinkage falls 4.9 percent relative to sales ($p = 0.10$)—and the later compression does no measurable damage (+1.9 percent, indistinguishable from zero): the erosion of the premium is, in effect, insured. Where she is not, the raise buys nothing on impact and the compression damage lands in full, a late rise of 4.9 percent ($p = 0.02$); at the monthly frequency the difference between the two late responses is not sharply estimated (-3.1 points, $p = 0.42$). The moderation runs through what the manager does, not what she is: splitting the same border stores by high- versus low-type instead, both groups suffer the compression and their difference is insignificant (Appendix K).

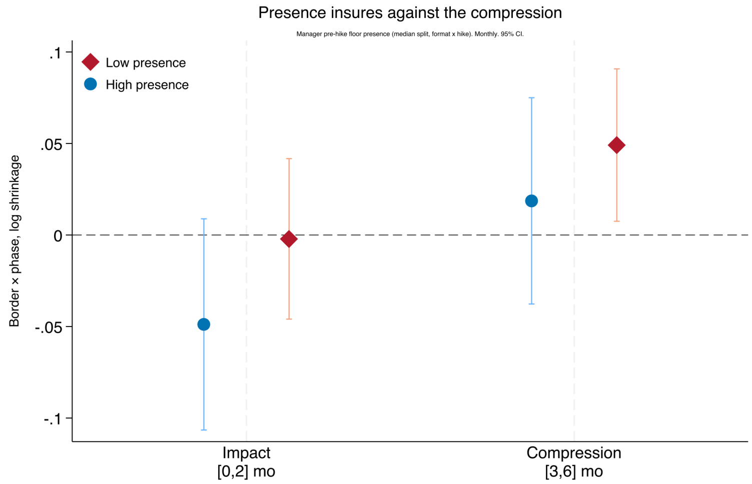


Figure 13: The compression response by the manager's floor presence at the hike. Border stores split at the median presence within format-by-hike cells; 95 percent confidence intervals.

This is not the marginal form of the level complementarity of Section 8. On the interior, that complementarity would put the marginal damage where monitoring is high; what insures the watched store is a corner: under a present manager, workers sit at full compliance and a small loss of rent leaves the Kuhn–Tucker slack positive (Proposition 5). The two facts are two comparative statics of one first-order condition, the interior cross-partial and the corner. Withdrawing the premium does its damage where no one is watching.

10 Conclusion

Every large employer faces the problem this paper studies: work that cannot be perfectly watched, performed by workers who earn more or less than their outside option, under managers who watch more or less closely. Economics has long had a theory for that problem—pay a rent and threaten to take it away—but testing it requires seeing things firms rarely let researchers see: what each worker earns against her own outside market, what she does when no one checks, and who is watching. This paper assembled all three inside one large Mexican retailer, using inventory shrinkage—a hundred-billion-dollar blind spot of the retail industry—as the audited residue of shirking. Both halves of the theory turned out to matter. Stores run by better managers lose less; workers with larger premia over their outside option shirk less; and each instrument raises the return to the other.

The complementarity has two faces, and the paper shows both. Adding an instrument pays most where the other is strong: a good manager accomplishes roughly three times more where the premium is high, and a premium that no one watches deters little. Losing a little of an instrument hurts where the other is weak: when the border hikes compressed the premium, shrinkage rose with a delay—but only in stores whose managers were seldom on the floor, while stores with a present manager rode out the squeeze insured. The two are not one finding read twice. The first is the cross-partial of the deterrent, the product of the rent and the monitor; the second is a corner, a store whose workers a present manager already holds at full compliance, where a small loss of rent has nothing left to move. The statute revealed, store by store, who was doing the disciplining—the premium, the manager, or no one.

For firms, pay policy and supervision are not separable levers. A wage ladder set without asking who monitors misprices both instruments: the same rent buys discipline in a watched store and nothing in an unwatched one, so compression shocks carry hidden costs concentrated precisely where management is thin, and the allocation of managers across stores functions as insurance against them. The business case for paying above the market runs through the monitoring capacity that converts rents into discipline.

For policy, minimum-wage effects operate inside the wage ladder, not only at the floor. Part of the cost of a hike arrives months later, on hidden margins—inventory losses—that standard datasets do not record, and its incidence depends on management quality. Small average effects in the minimum-wage literature are consistent with substantial heterogeneity across establishments that differ in how they are run.

The evidence comes from one firm in formal retail, and presence is a proxy for monitoring, not a direct measure of it. These boundaries point to the open questions: whether firms can create presence through the selection, training, and assignment of managers; how workers select into rents; how turnover evolves after a squeeze; and whether the complementarity between rents and monitoring generalizes to other hidden-action settings.

A Appendix: Proofs and Remarks for Section 3

A.1 Proofs

Throughout, derivatives of V are envelope derivatives of (6): a perturbation of W^p or μ moves the maximand directly, the induced movements of (e^*, a^*) contributing nothing at first order, so

$$\frac{\partial V}{\partial W^p} = \frac{1+r}{\rho} \;>\; 0, \qquad \frac{\partial V}{\partial \mu} = \frac{\pi_\mu(e^*;\mu)}{\rho} V = -\,\frac{q'(\mu)\,h(e^*)}{\rho} V \;\leq\; 0. \tag{24}$$

By (5), $\rho = r + b + q(\mu)h(e^*) + \big(1 - \pi^A(a^*)\big)$, so

$$\rho - q(\mu)h(e^*) \;=\; r + b + 1 - \pi^A(a^*) \;>\; 0. \tag{25}$$

Proof of Lemma 1. (i) The right-hand side of (6) is a contraction in V with modulus at most $1/(1+r) < 1$ because Definition 1(a) bounds Π in $[0,1]$ at every (e,a), so V exists and is unique. Under (3), $\pi_e(e;\mu) = q(\mu)(-h'(e))$, so (7) is $M(e^*) = qV/(1+r)$ with $M \equiv c_e/(-h')$, and

$$M'(e) \;=\; \frac{c_{ee}}{-h'} + \frac{c_e\,h''}{(h')^2} \;>\; 0,$$

so the condition has at most one solution; the derivative of the maximand in e equals $(-h')\big[qV/(1+r) - M(e)\big]$, positive to the left of the crossing and negative to the right, so the care problem is single-peaked and e^* is its unique maximum. At the optimum, substituting $c_e = q(-h')V/(1+r)$ into M' gives $M'(e^*) = \big[c_{ee} + qh''V/(1+r)\big]/(-h') = D/(-h')$, and $\pi_{ee} = -qh'' \leq 0$ gives $D > 0$. Attendance is unique given V by $D_A > 0$. (ii) is the display above. (iii) is (24); at an interior optimum, Definition 1(d) with (9) requires $W_i^p > c(x_i, e^*) + \kappa(a^*)$, and then $V > 0$. □

Proof of Proposition 1. (i) is (14), Step 2 of the proof of Proposition 3, which uses only (3): the direct effect of μ_m on the slope $\pi_e = q(\mu)(-h')$ raises the right-hand side of (7), the indirect effect through the surplus, $V_\mu = -q'h(e^*)V/\rho \leq 0$, lowers it, and (25) shows the first dominates. (ii) By (27) (Step 3 of the same proof), the ratio of the two absence responses is the ratio of the two surplus responses, $\big(\partial a^*/\partial\mu\big)/\big(\partial a^*/\partial W^p\big) = V_\mu/V_{W^p} = \pi_\mu V/(1+r) = -q'h(e^*)V/(1+r)$, using (24); dividing (14) by (26) gives the care identity. Both are (11); the sign is Step 3, and the absence response is bounded by $q'(\mu)h(e^*)$ times a positive factor, so it vanishes with $h(e^*)$ while (14) involves $h(e^*)$ only through ρ, which stays bounded. (iii) Let an unanticipated shock $\varepsilon \sim G$ to the value of the outside option be realized after actions, so that V is the stationary surplus of the text; the worker quits if and only if $\varepsilon > V$, so the quit probability is $1-G(V)$ and $\partial(1-G(V))/\partial\mu = -g(V)V_\mu = g(V)\,q'h(e^*)V/\rho \geq 0$. (If the shock recurs and is anticipated, V solves a Bellman with the option value $\Omega(V) = \int_V^\infty (\varepsilon - V)\,dG$ added to the continuation, and the same expression holds with $V+\Omega$ in place of V and $1+r-\Pi G(V)$ in place of ρ; the sign is unchanged.) The dismissal rate of a compliant worker is $q(\mu)h(e^*)$, which is of the order of $h(e^*)$. □

Proof of Proposition 2. Differentiating (7) implicitly and using (24),

$$\frac{\partial e^*}{\partial W^p} \;=\; \frac{\dfrac{\partial V}{\partial W^p}\cdot\dfrac{\pi_e(e^*;\mu_m)}{1+r}}{c_{ee}(x_i,e^*) - \dfrac{\pi_{ee}(e^*;\mu_m)}{1+r}\,V(x_i)} \;=\; \frac{\pi_e}{\rho\, D} \;>\; 0, \tag{26}$$

where the numerator is strictly positive—at an interior optimum, (7) with $c_e > 0$ and $V > 0$ forces $\pi_e(e^*;\mu_m) > 0$—and $D > 0$ by Lemma 1. The identical argument on (8) gives $\partial a^*/\partial W^p = \pi^{A\prime}/(\rho D_A) > 0$, with $\pi^{A\prime}(a^*) > 0$ forced by $\kappa' > 0$. Part (iii): $d\Pi(e^*,a^*;\mu_m)/dW^p = \pi_e\,\partial e^*/\partial W^p + \pi^{A\prime}\,\partial a^*/\partial W^p > 0$, and with the shock of Proposition 1(iii) the quit probability $1-G(V)$ falls with V and hence with W^p. Without such a shock the model has no voluntary quits. □

Proof of Proposition 3. Step 1: the product form. Lemma 1(ii), with $M'(e^*) = D/(-h')$.

Step 2: the two slopes. Differentiate (13) in W^p: $M'(e^*)\,\partial e^*/\partial W^p = qV_{W^p}/(1+r) = q/\rho$ by (24), so $\partial e^*/\partial W^p = q(-h')/(\rho D) = \pi_e/(\rho D)$, which is (26). Differentiate in μ:

$$M'(e^*)\,\frac{\partial e^*}{\partial\mu} \;=\; \frac{q'V + q\,V_\mu}{1+r} \;=\; \frac{q'V}{1+r}\Big(1 - \frac{q\,h(e^*)}{\rho}\Big) \;=\; \frac{q'V}{1+r}\cdot\frac{r+b+1-\pi^A(a^*)}{\rho} \;>\; 0,$$

by (24) and (25); dividing by $M'(e^*) = D/(-h')$ gives (14). The computation makes explicit what Proposition 1(i) holds fixed: a stricter monitor also trims the surplus of a worker who still carries exposure ($V_\mu \leq 0$), and (25) shows the deterrence channel dominates unconditionally.

Step 3: the attendance responses. Implicit differentiation of (8) gives, for $\theta \in \{W^p,\mu\}$,

$$\frac{\partial a^*}{\partial\theta} \;=\; \frac{\pi^{A\prime}(a^*)}{(1+r)\,D_A}\,\frac{\partial V}{\partial\theta}, \tag{27}$$

so $\partial a^*/\partial W^p = \pi^{A\prime}/(\rho D_A) > 0$ and $\partial a^*/\partial\mu = -\pi^{A\prime}q'h(e^*)\,V/\big(\rho(1+r)D_A\big) \leq 0$.

Step 4: the cross-partial. Both sides of (13) are positive; take logarithms and write $\ell \equiv \log M$, $v \equiv \log V$:

$$\ell(e^*) \;=\; \log q(\mu) + v(W^p,\mu) - \log(1+r). \tag{28}$$

Differentiating in W^p gives $\ell'(e^*)\, e^*_{W^p} = v_{W^p}$, with $\ell'(e^*) = M'(e^*)/M(e^*) = D/c_e > 0$ and $v_{W^p} = V_{W^p}/V = (1+r)/(\rho V) > 0$. Differentiating that identity in μ and solving,

$$\frac{\partial^2 e^*}{\partial W^p\, \partial\mu} \;=\; \frac{v_{W^p\mu} \;-\; \ell''(e^*)\, e^*_\mu\, e^*_{W^p}}{\ell'(e^*)}. \tag{29}$$

The second numerator term is nonnegative: $\ell'' \leq 0$ by Assumption 2 and $e^*_\mu, e^*_{W^p} > 0$ by Step 2. For the first, differentiate $v_{W^p} = (1+r)/(\rho V)$ totally in μ, using $d\rho/d\mu = -\big(\pi_e e^*_\mu + \pi^{A\prime} a^*_\mu + \pi_\mu\big)$ and $\rho\, V_\mu = \pi_\mu V$ from (24): the π_μ terms cancel, leaving

$$v_{W^p\mu} \;=\; \frac{(1+r)\,\beta}{\rho^2\, V}, \qquad \beta \;\equiv\; \pi_e(e^*;\mu)\, e^*_\mu + \pi^{A\prime}(a^*)\, a^*_\mu. \tag{30}$$

β is the induced-retention response: the movement in Π produced by the worker's behavioral answer to a stricter monitor, over and above the mechanical effect π_μ. Substituting the closed forms of Steps 2–3 and reducing with (7) ($q(-h')V/(1+r) = c_e$) and (8) ($\pi^{A\prime}V/(1+r) = \kappa'$),

$$\beta \;=\; \frac{q'(\mu)}{\rho\, D\, D_A}\Big[\big(-h'(e^*)\big)\, c_e\, \big(r + b + 1 - \pi^A(a^*)\big)\, D_A \;-\; h(e^*)\, \pi^{A\prime}\kappa'\, D\Big],$$

which is strictly positive if and only if (12) holds. Hence $v_{W^p\mu} > 0$, and (29) delivers (15). By Young's theorem the same inequality reads in the other direction: the premium's grip on care, $\partial e^*/\partial W^p$, is increasing in μ.

Step 5: integration. $\Delta_\mu e^*(W^p) = \int_{\mu_{LT}}^{\mu_{HT}} \partial e^*/\partial\mu\, d\mu > 0$ by (14), and (16) follows from the fundamental theorem of calculus applied twice, the integrand in (15) being continuous and strictly positive on the compact rectangle. □

Proof of Proposition 4. (i) Under enforcement ($s = 1$) the model is the baseline, and Lemma 1(iii) gives V strictly increasing and continuous in W^p, so $V(W^p,\mu) \geq \bar{v}$ if and only if $W^p \geq \overline{W}(\mu)$ for the unique root $\overline{W}(\mu)$ of $V(\cdot,\mu) = \bar{v}$ (taking $\overline{W} = \infty$ if $V < \bar{v}$ on the whole range). With $s = 0$, $\pi = 1 - b$ does not depend on e, the maximand of (6) is strictly decreasing in e because $c_e > 0$, and $e^* = 0$ by Definition 1(b); neither W^p nor μ enters the care condition. Consistency: the rule is stated on the surplus under enforcement, so if that surplus is at least $\bar{v}$ the manager sanctions and the baseline solution obtains, and if it is below $\bar{v}$ she does not and the corner obtains; the two cases are exclusive and exhaustive. (ii) The sanction rate of a worker with care e under monitor μ is $s\, q(\mu) h(e)$, with s evaluated at the monitor in place. Because $V_\mu \leq 0$, $\overline{W}(\mu)$ is nondecreasing in μ, so credibility under μ_H implies credibility under μ_L. At the care level e^*_L chosen under μ_L, replacing the monitor changes the rate by $[q(\mu_H) - q(\mu_L)]\, h(e^*_L) > 0$ where $W^p \geq \overline{W}(\mu_H)$, because $q' > 0$ and $h \geq 0$, and by zero where $W^p < \overline{W}(\mu_L)$; in the band between, the arrival ends enforcement and the rate falls to zero.

(iii) Above the threshold the hazard is (3) and Proposition 3 applies; below it $e^* = 0$ at both monitors, so $\Delta_\mu e^* = 0$. For the store, the arrival effect on average care is $\frac{1}{N_j}\sum_{i:\, W^p_{ij}\geq \overline{W}} \Delta_\mu e^*_{ij}$. As $\overline{W^p}_j$ rises with the distribution of η_i fixed, every existing summand is nondecreasing by Proposition 3(iii) and every worker who crosses the threshold adds a nonnegative summand, so the sum is nondecreasing in $\overline{W^p}_j$. □

Proof of Corollary 1. Write $\ell(u) = \ell_0 + \ell_1 u$ with $\ell_1 > 0$. The first two signs differentiate the sum term by term at fixed (x_i, η_i) and apply (26) and (14); the cross-partial is $\partial^2 S_j/\partial\overline{W^p}_j\partial\mu_j = -(\ell_1/N_j)\sum_i \partial^2 e^*_{ij}/\partial W^p \partial\mu < 0$ by (15), and the arrival statement integrates it over $[\mu_{LT}, \mu_{HT}]$ as in Step 5 of the proof of Proposition 3. For logs, the quotient rule applied to $\partial \log S_j/\partial\overline{W^p}_j = S_{W^p}/S_j$ gives $\partial^2 \log S_j/\partial\overline{W^p}_j\partial\mu_j = (S_{W^p\mu}S_j - S_{W^p}S_\mu)/S_j^2 < 0$, because $S_{W^p\mu} < 0$ and $S_{W^p}S_\mu > 0$. If instead $S_j = f(1-\bar{e}_j)$ with $f'' > 0$, $S_{W^p\mu} = f''\,\bar{e}_{W^p}\,\bar{e}_\mu - f'\,\bar{e}_{W^p\mu}$, which is the caveat in the text. □

Proof of Proposition 5. (i) By the envelope theorem on (6) with e fixed at 1, $V^{(1)}_{W^p} = (1+r)/\rho^{(1)}$ and $V^{(1)}_\mu = -q'h(1)V^{(1)}/\rho^{(1)}$, where $\rho^{(1)} \equiv r + b + q(\mu)h(1) + 1 - \pi^A(a^*)$. Then $\Gamma_{W^p} = q/\rho^{(1)} > 0$ and

$$\Gamma_\mu \;=\; \frac{q'V^{(1)}}{1+r}\Big(1 - \frac{q\,h(1)}{\rho^{(1)}}\Big) \;=\; \frac{q'V^{(1)}}{1+r}\cdot\frac{r+b+1-\pi^A(a^*)}{\rho^{(1)}} \;>\; 0,$$

the computation of (25) at $e = 1$. Strict monotonicity in both arguments gives the threshold, with slope $\bar{\mu}'(W^p) = -\Gamma_{W^p}/\Gamma_\mu < 0$. The policy is unique given V by single-peakedness (Lemma 1) and V is unique by the contraction property, so no interior solution coexists with the corner. In the notation of the interior solution, $\Gamma \geq 0$ reads $\tilde{e}(W^p, \mu) \geq 1$, with $\tilde{e}$ the root of (13) when M is extended monotonically beyond 1; the lower corner $e^* = 0$ is $q(\mu)V^{(0)}/(1+r) \leq M(0)$ and plays no role in Proposition 5. (ii) Γ is continuous in W^p, so $\Gamma(W^p + dW^p, \mu_H) > 0$ for $|dW^p|$ small and $e^* = 1$ at both premiums; at μ_L the optimum is interior and (26) applies. (iii) is (15) integrated over $[\mu_L, \mu_H]$. □

A.2 Remarks: Where the Data Refine the Model

The simplest version of the model makes sharp predictions, and several do not survive contact with the data unchanged.

Monitoring is embodied, not purchased. The model writes monitoring as an intensity the firm can dial up. High-type managers, however, do not buy more monitoring apparatus or reorganize the store—they add no loss-prevention, supervisory, or security staff, and the team's composition is unchanged on arrival. They are physically present more, and their co-presence disciplines worker punctuality shift by shift. μ_m is therefore an embodied attribute of the manager that is portable across stores and reverts when she departs, not a store-level technology the firm can set independently of who runs the store. As a moderator of the compression response, the manager's measured presence is predetermined, passes the pre-trend checks, and carries the interaction (Section 9); her estimated fixed effect fails those checks at the border and cannot serve as the moderator (Appendix K).

The complementarity is intrinsic, not behavioral. A natural reading of the interaction evidence is that good managers optimize against the rent—supplying more supervision where the premium gives supervision more to protect. The data reject it: action by action, the manager's conduct barely responds to the premium, while her enforcement and her results respond sharply. The model accommodates this by keeping μ_m out of the manager's choice set altogether: the complementarity lives in the worker's first-order condition, where the detection rate multiplies the rent, not in the manager's behavior. The premium scales what supervision is *worth*, not how much of it the manager supplies.

Shirking is not one margin. A single-action model predicts that whatever disciplines shrinkage should discipline absence too. The data split them: absence responds to the premium but not to the manager, in either environment, while shrinkage responds to both and to their interaction. The refinement is the two-margin structure—care must be detected by someone, absence records itself—and it converts an apparent set of nulls into the model's cleanest signature. The squeeze adds an asymmetry the same structure absorbs: the statutory raise moves attendance everywhere, yet the squeeze's erosion, which read symmetrically should reduce it, leaves attendance at its status quo and surfaces instead as shrinkage in unwatched stores (Sections F and 9). The verifiable margin is disciplined tightly—most workers sit at or near full compliance, where a marginal loss of surplus has no room to move behavior—while the hidden margin is interior, and interior margins are where marginal changes in the rent do their work. Proposition 5 formalizes the corner: the Kuhn–Tucker slack at full compliance decides whether a small movement in the rent moves a margin at all, and the same slack, evaluated on the care margin under a present manager, is what insures the watched store in Section 9.

Complements in levels, saturation at the margin. Written as a no-shirking condition, the premium and monitoring look like substitute instruments—the firm could sustain the same care with more of one and less of the other. The evidence is two-signed: a good manager accomplishes more where the standing premium is high (Section 8), yet the statutory erosion of the rent damages only the stores where little monitoring is supplied, and the statutory raise converts into shrinkage discipline only where it is (Section 9). The refinement is to distinguish the level complementarity of the detection-times-rent product, Proposition 3(ii)–(iii), from the marginal redundancy of an extra unit of rent under a manager whose workers already comply, Proposition 5(ii). The two are not one comparative static read at two horizons: on the interior the complementarity puts the marginal damage of a compression where monitoring is *high* (Proposition 5(iii)), so the insured store is evidence that a present manager holds care at the corner, not evidence of the interior cross-partial.

B Appendix: Identification and Limited-Mobility Bias in the AKM Decomposition

B.1 Identification

OLS estimation of Equation (19) imposes the moment conditions

$$\mathbb{E}\left[\theta_{m(j,t)}\,\varepsilon_{jt}\right] = 0 \qquad \text{and} \qquad \mathbb{E}[\psi_j\,\varepsilon_{jt}] = 0. \tag{31}$$

Identification therefore relies on the assignment of managers to stores being conditionally mean-independent of past, present, and future values of ε_{jt}. This assumption is weaker than random assignment: it permits sorting on the permanent components, so that better managers may systematically run better or worse stores, since both fixed effects are estimated. What it excludes is assignment based on the match component η or on transitory shocks—for instance, a manager dispatched to a store because shrinkage there is spiking, or a manager who moves only to stores where her particular style works. Such endogenous mobility would bias the estimated effects, and Section 4 tests its observable implications directly.

The *timing* of a manager change is not fully orthogonal to performance—a store is somewhat more likely to change manager after a high-shrinkage month—but managers are not assigned to systematically better or worse stores (origin and destination shrinkage are statistically indistinguishable), and who moves is selected on permanent traits (seniority and pay) rather than transitory shocks. What the design requires is mean-independence of mobility from the transitory error given the fixed effects, not literal randomness.

Table 13: Manager mobility and connected sets

	General-manager panel
Panel A: Coverage	
Store-months with a matched general manager	$> 90\%$
Median months observed per manager	13
Panel B: Manager mobility	
Share of movers (managers in ≥ 2 stores)	55.0%
Share of managers in ≥ 3 stores	25.4%
Share of stores with at least one mover	91.7%
Panel C: Connected sets	
Number of connected sets	10
Share of stores without movers (FE not identified)	8.3%
Largest connected set: share of the panel	$\approx 73\%$
Largest connected set: mover share	60.7%
Share of store effects identified	88%

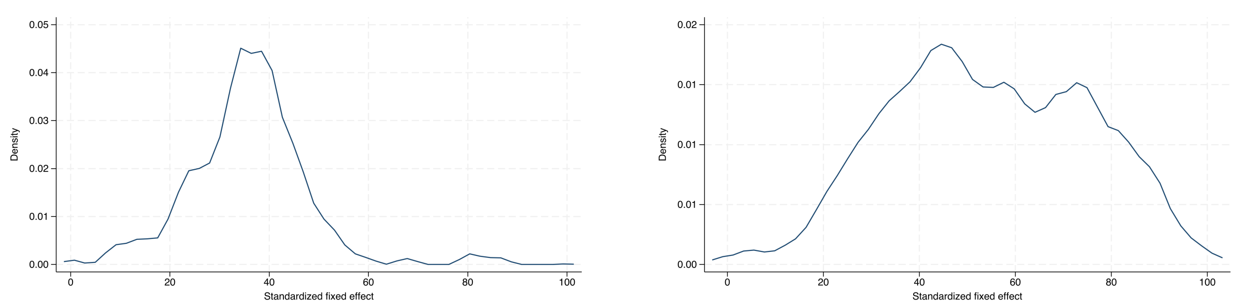
General managers matched to monthly store outcomes, 2022–2024. A mover is a manager observed in at least two stores. “Share of the panel” is the average of the shares of stores, managers, and manager-months in the largest connected set.

Table 14: The anatomy of general-manager moves

Store-to-store GM moves, 2022–2024	≈520
Distinct GMs who move	≈280
Distinct stores involved	≈280
Temporal concentration (arrival months)	
Wave months (monthly moves $\geq$ 75th pctile)	8 of 29
Share of moves in wave months	38.7%
Share of moves in top-3 months	15.9%
Mean moves per month: wave vs. other months	25.2 vs. 15.2
Chains and swaps (same arrival month)	
Part of a chain (origin store also receives a mover)	47.7%
Direct swap ($A \to B$ and $B \to A$)	19.9%
Destination status at arrival	
Vacant ($\geq$1 month without GM)	35.1%
Immediate handover (GM present in $t-1$)	64.9%
Origin–destination shrinkage distance (6-month pre-move avg.)	
Mean Δ log(Shrink Value), destination − origin	0.026
	(0.023)
Moves with both pre-move averages	≈480

A move is a general manager observed in one store and, in the next observed month (gaps of at most two months), in another. A wave month has monthly moves at or above the 75th percentile of the monthly counts. The shrinkage distance compares six-month pre-move averages of log shrinkage value at the destination and the origin (standard error in parentheses; moves with at least three pre-move months on each side).

B.2 Limited-Mobility Bias



(a) Manager fixed effects $\hat{\theta}_m$

(b) Store fixed effects $\hat{\psi}_j$

Figure 14: Plots of the estimated fixed effects, log shrinkage value

Notes: These are plots of *estimated* effects, which convolve the true effects with sampling noise; the bias-corrected variance components are in Table 5. Kernel densities of the fixed effects from Equation (19), standardized within connected set and rescaled to $[0, 100]$. Outcome: log monthly shrinkage value USD.

The variance decomposition of Table 5 starts from a plug-in estimate: it computes the variance of the *estimated* fixed effects. An estimated effect is the true effect plus estimation noise—a manager observed in a handful of stores gets a fixed effect estimated from few observations—and the variance of a noisy estimate mechanically exceeds the variance of the object it estimates. Plug-in variances are therefore biased upward, and the plug-in correlation between manager and store effects is biased downward (Andrews et al., 2008). This is a property of every AKM decomposition, not a flaw of this one, and it has a standard remedy: estimate the noise and subtract it. I do so twice, on the largest connected set, where the effects are best identified: with the correction of Andrews et al. (2008), which treats the noise as homoskedastic, and with the leave-out estimator of Kline et al. (2020) (KSS), which allows heteroskedasticity and is the demanding benchmark.

The corrections confirm the result and size it honestly. The manager component is the noisiest object—each manager is observed in few stores—so it is where the correction bites: $\widehat{\mathrm{Var}}(\theta)$ falls from 0.049 to 0.016 under KSS, so about a third of the plug-in manager variance is signal. The store component barely moves (0.287 to 0.251): stores are observed for many months, their effects carry little noise, and the correction leaves them nearly intact—exactly the pattern noise removal should produce. The corrected sorting correlation is −0.04, indistinguishable from zero. What survives is smaller than the plug-in but economically large: the 13-percent-per-standard-deviation figure of the main text is computed from the corrected 0.016, not from the plug-in. The correlation above one in the sales-ratio column of Table 15 (and near one for the per-worker ratio) is an artifact of a near-zero corrected variance in the denominator, not evidence about sorting.

Table 15: AKM variance decomposition with limited-mobility-bias corrections, all shrinkage measures

	log(Shrink Value)			log(Shrink/Sales)			log(Shrink/Worker)		
	Baseline	*Andrews*	*KSS*	*Baseline*	*Andrews*	*KSS*	*Baseline*	*Andrews*	*KSS*
$\widehat{\mathrm{Var}}(\theta)$	0.049	0.009	0.016	0.038	0.003	0.006	0.044	0.002	0.006
$\widehat{\mathrm{Var}}(\psi)$	0.287	0.243	0.251	0.075	0.037	0.040	0.073	0.027	0.032
$\widehat{\mathrm{Cov}}(\theta, \psi)$	−0.014	0.002	−0.002	−0.029	−0.015	−0.018	−0.024	−0.007	−0.010
$\widehat{\mathrm{Corr}}(\theta, \psi)$	−0.120	0.044	−0.038	−0.553	−1.399	−1.145	−0.424	−0.913	−0.739
N		≈5,900			≈5,900			≈5,900	

Notes: Largest connected set of the general-manager panel. θ = manager FE, ψ = store FE; observation-weighted variances and covariances. *Baseline*: OLS variance of the estimated fixed effects. *Andrews*: homoskedastic correction of Andrews et al. (2008). *KSS*: leave-out estimator of Kline et al. (2020). Correlations exceeding one in absolute value reflect near-zero corrected manager variance and are a known feature of these estimators.

C The Arrival-and-Departure Design: Shrinkage-to-Sales

The four designs on the ratio measure. Table 16 repeats the four designs of Table 6 with log shrinkage-to-sales as the outcome and reproduces the pattern of the main text: shrinkage falls on a high-type arrival, rises on a high-type departure and on a low-type arrival.

Table 16: Shrinkage-to-sales on manager arrivals and departures

	Log shrinkage-to-sales
Arrival of a high-type manager	-0.0341^{**}
	(0.0164)
Departure of a high-type manager	0.0600^{*}
	(0.0350)
Arrival of a low-type manager	0.0640^{***}
	(0.0193)
Departure of a low-type manager	-0.0096
	(0.0281)
Mean	0.84
Median	0.87

Design, samples, and notes as in Table 6, with the winsorized log ratio of shrinkage to sales as the outcome. $^{***}p<0.01$, $^{**}p<0.05$, $^{*}p<0.1$.

D What Good Managers Do: Additional Evidence

This appendix collects exercises referenced in Section 5: the composition of the manager's schedule, the monitoring apparatus, and the detail of the personnel reallocation and team diversity; the dose response and novelty placebo of the arrival effect; the cross-area component of the personnel reallocation of Section 5.3; the targeted-presence and within-shift discipline results; and supporting material on ordinary managers and the comovement of the estimated fixed effects.

D.1 Schedule Composition, the Monitoring Apparatus, and the Team

Three tables detail margins summarized in the main text. First, the *composition* of the manager's schedule: Table 17 reports the arrival design for the manager's shift count together with her mean first-punch hour, her weekend *share*, and the store's night-shift share. A high-type arrival raises how much she works while leaving the mix essentially unchanged—her weekend share moves by less than one percentage point and the store's night share not at all—so the added weekend and night shifts of the main text scale her existing schedule rather than recompose it. A low-type arrival moves none of the presence measures and pushes the manager's average arrival time about fifteen minutes later.

Table 17: Presence and schedule composition on arrival (high- and low-type)

	(1) GM shifts /month	(2) GM arrival hour	(3) GM % weekend	(4) Store % night	(5) Co-presence (worker-hrs)	(6) Any GM overlap
HT manager arrival	1.6958***	0.0371	−0.0068*	−0.0019	0.0387***	0.0437***
	(0.4725)	(0.0909)	(0.0039)	(0.0026)	(0.0120)	(0.0135)
LT manager arrival	−0.0497	0.2479**	0.0055	0.0034	−0.0092	−0.0047
	(0.8728)	(0.1131)	(0.0037)	(0.0034)	(0.0181)	(0.0231)
Dep. mean	23.48	8.64	0.296	0.368	0.574	0.700

Each cell is the $\text{Post}_{[0,8]}$ coefficient from the arrival design of Equation (21) with store, year, and month fixed effects and standard errors clustered at the store, estimated on store-month measures built from shift-level time-clock records on the original characterization chain. "GM arrival hour" is the manager's mean first-punch time; "GM % weekend" her share of shifts on Saturday or Sunday; "Store % night" the share of all worker shifts on night hours; remaining columns as in the main text. $^{***}p<0.01$, $^{**}p<0.05$, $^{*}p<0.1$.

Second, the *apparatus*: the first three columns of Table 18 measure the machinery of monitoring a store could add—the share of loss-prevention personnel, of supervisory staff, and of security workers. None responds to a high-type arrival; if anything the supervisory share falls slightly. The monitoring behind the shrinkage reduction is the manager's own presence, not a technology she installs.

Table 18: What managers change on arrival (within-store arrival design)

	(1) Loss-prev. staff/100	(2) % super-visors	(3) % security staff	(4) Log team size	(5) Promotions /100	(6) Conflict quits/100
HT manager arrival	0.0966	−0.6352*	−0.1231	0.0210**	0.2530**	−0.0020
	(0.1652)	(0.3761)	(0.1771)	(0.0096)	(0.1259)	(0.0439)
LT manager arrival	−0.0976	0.1403	0.2151*	−0.0037	−0.0630	0.0839***
	(0.1151)	(0.4700)	(0.1296)	(0.0101)	(0.1367)	(0.0301)

Each cell is the $\text{Post}_{[0,8]}$ coefficient from a separate store-month arrival regression (Equation 21) with store, year, and month fixed effects and standard errors clustered at the store. Columns 1–3 measure monitoring organization; columns 4–6 team composition and personnel practices. Rate outcomes are per hundred workers. $^{***}p<0.01$, $^{**}p<0.05$, $^{*}p<0.1$.

The full separations-and-accessions decomposition behind the body's team table: Table 19 reports, for the four designs, log turnover and its components (quits, dismissals, hires), team size, cross-area reassignments, and promotions.

Table 19: The team on manager arrivals and departures: full decomposition

	(1) Log turnover (p.e.)	(2) Log quits (p.e.)	(3) Log dismissals (p.e.)	(4) Log hires (p.e.)	(5) Log team size	(6) Cross-area reassign. (per 100)	(7) Promotions (per 100)
HT manager arrival	0.1738**	0.1650**	0.0026	0.2084**	0.0074	0.2076**	0.1736
	(0.0724)	(0.0771)	(0.0794)	(0.0881)	(0.0108)	(0.0920)	(0.1308)
HT manager departure	0.0519	0.0150	0.0212	0.0035	−0.0081	0.2021	0.2609
	(0.1896)	(0.1801)	(0.1694)	(0.1486)	(0.0248)	(0.1993)	(0.2375)
LT manager arrival	0.0245	0.0502	0.1104	0.0233	−0.0089	0.0286	−0.0946
	(0.0610)	(0.0638)	(0.0852)	(0.0794)	(0.0091)	(0.1134)	(0.1292)
LT manager departure	0.0279	−0.0095	0.1264	0.1145	0.0065	0.5081***	0.1853
	(0.1486)	(0.1454)	(0.1173)	(0.1517)	(0.0193)	(0.1896)	(0.2024)
Mean	−3.32	−3.51	−5.72	−3.46	4.12	1.44	2.41
Median	−2.80	−2.98	−6.91	−2.91	4.03	0.79	1.93

Notes: Each cell is the $\text{Post}_{[0,8]}$ coefficient from the design of Table 6 for the row's event. Per-employee (p.e.) outcomes are logs of monthly rates; quits, dismissals, and hires decompose the separations and accessions in the personnel record; cross-area reassignments and promotions are counts per hundred workers. Mean and Median are of the dependent variable. $^{***}p<0.01$, $^{**}p<0.05$, $^{*}p<0.1$.

Third, the *team*: Table 20 decomposes the within-store personnel moves by direction on the original chain—rotation rises modestly and is concentrated in upward and lateral moves, with an independent promotions record confirming the upward margin and downward moves and demotions flat—and Table 21 shows that a high-type arrival moves none of the team-diversity measures of Harrison and Klein (2007); the lone exception, gender variety, is 0.003 on an index bounded at 0.5 and economically negligible. The shrinkage reduction is not a selection or composition story.

Table 20: Internal reallocation is upward, lateral, and across areas

	Position change (tier-based)				Move records	
	All	Upward	Lateral	Downward	Promotions	Demotions
High-type arrival	0.332**	0.165**	0.160	0.008	0.253**	−0.012
	(0.154)	(0.069)	(0.098)	(0.063)	(0.126)	(0.018)
Low-type arrival	−0.058	0.024	−0.087	−0.005	−0.063	0.030
	(0.195)	(0.079)	(0.111)	(0.071)	(0.137)	(0.025)
Mean of outcome	3.59	1.02	1.92	0.62	2.43	0.12

Each cell is the $\text{Post}_{[0,8]}$ coefficient from a separate store-month arrival regression (Equation 21), rates per hundred workers, with store, year, and month fixed effects and standard errors clustered at the store. The first columns classify a within-store position change by the hierarchical tier of its origin and destination (upward, lateral, downward); the last columns are promotion and demotion counts from the personnel-movement record. $^{***}p<0.01$, $^{**}p<0.05$, $^{*}p<0.1$.

Table 21: Team diversity does not change on a high-type arrival

	Separation (SD)			Variety (Blau)			Disparity
	Age	Tenure	Educ.	Gender	Tier	Area	Pay (CV)
High-type arrival	0.029	−0.128	−0.007	0.003**	−0.001	0.000	0.007
	(0.070)	(0.451)	(0.021)	(0.001)	(0.003)	(0.003)	(0.005)
Low-type arrival	−0.045	−0.460	−0.052***	−0.000	−0.003	−0.003	0.005
	(0.066)	(0.467)	(0.017)	(0.002)	(0.003)	(0.004)	(0.005)
Mean of outcome	12.4	31.4	2.59	0.477	0.667	0.651	0.500

Each cell is the $\text{Post}_{[0,8]}$ coefficient from a separate store-month arrival regression (Equation 21) with store, year, and month fixed effects and standard errors clustered at the store. Diversity follows Harrison and Klein (2007): separation by the within-store standard deviation of age, tenure, and education; variety by the Blau index ($1-\sum_k p_k^2$) of gender, tier, and occupational area; disparity by the coefficient of variation of monthly pay. $^{***}p<0.01$, $^{**}p<0.05$, $^{*}p<0.1$.

D.2 Dose Response and a Novelty Placebo

The arrival design classifies managers by a median split, but the shrinkage effect should be graded, not binary: a store that trades up by a wide margin should improve more than one that trades up narrowly. To test this I rank every arrival of a new general manager by the *improvement* it represents—the incumbent management's mean shrinkage fixed effect minus the arriving manager's—and estimate the arrival effect separately for the largest upgrades. Because the ranking uses the continuous fixed effect rather than the high-type indicator, it does not depend on the classification cutoff. Table 22 shows a dose response: the shrinkage reduction grows from 4.3 percent in the top half of upgrades to 7.3 percent in the top quartile and 9.3 percent in the top decile.

The same table delivers a novelty placebo. If workers simply exerted more effort for any new boss, arrivals that are *not* upgrades would also lower shrinkage. They do not. When the arriving manager is no better than the incumbent she replaces (the bottom half of the improvement distribution), shrinkage *rises* by 4.9 percent, and pooling all arrivals of new management regardless of quality yields no effect at all (+0.002, not significant). What moves shrinkage is the quality of the arriving manager, not the disruption of a new face—a distinction the high-type effect shares with the manager-quality effects of Lazear et al. (2015).

Table 22: Dose response and a novelty placebo (log shrinkage value)

Treatment definition	Effect $\text{Post}_{[0,8]}$
Top 10% upgrade (arriving $\gg$ incumbent)	−0.0928***
	(0.0220)
Top 25% upgrade	−0.0729***
	(0.0159)
Top 50% upgrade	−0.0430***
	(0.0105)
Bottom 50% (novelty placebo)	+0.0490***
	(0.0096)
Any arrival of new management	+0.0017
	(0.0077)

Each row is the $\text{Post}_{[0,8]}$ coefficient from a separate arrival difference-in-differences on log shrinkage value. "Improvement" is the store's baseline incumbent shrinkage fixed effect minus the arriving manager's; upgrade tiers are the top 10/25/50 percent of that distribution. "Novelty placebo" restricts to arrivals in the bottom half (arriving manager no better than the incumbent); the last row pools all arrivals of new management. Store, year, and month fixed effects; standard errors clustered at the store; store counts rounded. $^{***}p<0.01$, $^{**}p<0.05$, $^{*}p<0.1$.

D.3 Reallocation Across Functional Areas

Section 5.3 reports that high-type managers rotate the workers they inherit; Table 23 shows that part of this rotation moves workers across the store's functional areas, not only up and down its tiers. On a high-type arrival, the rate at which a worker's occupational area changes from one month to the next rises by about 0.21 per hundred workers, and the increase is concentrated in moves into client-facing roles (sales, service, and food; +0.09 per hundred, $p < 0.01$), with moves into back-office roles (logistics and administration) positive but imprecise (+0.05). Moves into the loss-prevention chain do not rise significantly. Low-type arrivals move none of these margins. The pattern is consistent with the redeployment of incumbents documented for good managers in the sister setting (Adhvaryu et al., 2026), and, like the upward mobility of the main text, it describes a manager who reshuffles the people she has rather than replacing them.

Table 23: High-type managers reassign workers across functional areas

	Area change (any)	Into client-facing	Into back-office	Into loss-prevention
High-type arrival	0.208**	0.088***	0.054	0.026
	(0.092)	(0.031)	(0.035)	(0.025)
Low-type arrival	0.029	0.020	−0.021	0.001
	(0.113)	(0.031)	(0.032)	(0.024)
Mean of outcome	1.44	0.32	0.25	0.14

Each cell is the $\text{Post}_{[0,8]}$ coefficient from a separate store-month arrival regression (Equation 21), rates per hundred workers, with store, year, and month fixed effects and standard errors clustered at the store. "Area change" is any month-over-month change in a worker's occupational area; "client-facing" comprises sales, service, and food roles, "back-office" logistics and administration, and "loss-prevention" the receiving–tagging–protection chain. $^{***}p$<0.01, $^{**}p$<0.05, $^{*}p$<0.1.

D.4 Presence Is Broad, Not Targeted

A further question is whether the manager directs this added presence toward the workers who matter most for shrinkage. Table 24 recomputes co-presence separately for four high-risk groups—the loss-prevention chain that receives, tags, and protects merchandise; security staff; cashiers; and merchandise organizers—and re-runs the arrival design on each. A high-type arrival raises co-presence by a nearly identical three to four percentage points for *every* group, including the average worker; the increase is broad, not concentrated on any one role. High-type managers therefore supervise everyone more rather than reallocating a fixed budget of attention toward shrinkage-prone tasks—the opposite of the targeted-monitoring pattern of Bandiera et al. (2007), and consistent with a margin of generalized presence rather than a surveillance technology.

Table 24: The rise in co-presence is broad, not targeted to risk groups

	(1) All workers	(2) Loss-prevention	(3) Security	(4) Cashiers	(5) Perishables
HT manager arrival	0.0387***	0.0370***	0.0326***	0.0385***	0.0368**
	(0.0120)	(0.0140)	(0.0107)	(0.0121)	(0.0144)
LT manager arrival	−0.0092	−0.0110	−0.0067	−0.0076	−0.0136
	(0.0181)	(0.0183)	(0.0156)	(0.0186)	(0.0194)
Dep. mean	0.574	0.611	0.485	0.568	0.609

Each cell is the $\text{Post}_{[0,8]}$ coefficient from the arrival design of Equation (21), with the outcome being the mean manager–worker co-presence share computed over the worker shifts in the indicated risk group: the loss-prevention chain that receives, tags, and protects merchandise; security staff; cashier positions; and the perishables category. Store, year, and month fixed effects; standard errors clustered at the store. $^{***}p<0.01$, $^{**}p<0.05$, $^{*}p<0.1$.

D.5 Within-Shift Discipline

The added supervision should in turn change how workers behave, and the minute-level records allow a test of this claim at the level of the individual shift. Table 25 regresses, over millions of worker shifts, three deviations from each worker's own norm—how late she clocks in, how far her break runs over, and how early she leaves—on whether the general manager is co-present that shift, absorbing worker, store, month, hour-of-day, and day-of-week fixed effects. The timing fixed effects are essential: because the manager works days, a raw co-presence comparison would confound supervision with the difference between day and night shifts; with them, identification is the *same* worker, at the *same* time of day and weekday within a store-month, on days the manager is on the clock versus days she is not. The robust margin is punctuality: when the manager is co-present, the worker arrives about one to two minutes less late than her own norm, and the sign and significance are identical whether co-presence is measured by a binary indicator or by the continuous overlap share. The other two margins—break over-run and early departure—are not robustly signed under the binary indicator, and the continuous overlap share is estimable only for the punctuality margin, so I do not interpret them. I read the tardiness result alone as the clean finding: a within-worker response to being watched, at daily frequency.

Table 25: Workers arrive on time when the manager is present

	(1) Late arrival (minutes)	(2) Break over-run (minutes)	(3) Early departure (minutes)
GM co-present (0/1)	−1.988***	1.846***	−12.414***
	(0.627)	(0.267)	(0.881)
GM overlap share	−1.007***	—	—
	(0.259)		
Worker-shifts		≈8.9 million	

Worker-shift level. Each outcome is the deviation of that shift from the worker's own median (arrival minute, break minutes, and departure minute; early departure is the median departure minus the actual, so a negative coefficient means leaving *less* early). "GM co-present" is an indicator that the manager's store-day presence interval overlaps the shift; "GM overlap share" its continuous version. Worker, store, month, hour-of-day, and day-of-week fixed effects; standard errors clustered at the store. $^{***}p<0.01$, $^{**}p<0.05$, $^{*}p<0.1$.

D.6 Managers Are Otherwise Ordinary

These actions aside, high-type managers are unremarkable on the traits a personnel file would use to explain quality. Table 26 compares the two types. They earn the same salary—a log difference of 0.015, far from significant—and are equally educated; high-type managers are about a year older, a difference of under three percent that clears significance only at the ten-percent level. Where they differ is in career path rather than credentials: high-type managers have been at the firm about three and a half months *less* (a gap of roughly nine percent that runs against a seniority reading of quality), rotate through slightly fewer stores, and hold each posting marginally longer. None of these is a dimension the firm could read as competence in advance; on pay and schooling—the observables a hiring or promotion decision would price—the two types are indistinguishable.

Table 26: Manager observables by type (spell averages)

	HTM (N≈500)	LTM (N≈500)	Diff. (HTM−LTM)	p-value
Log salary (monthly, USD)	7.285	7.270	0.015	0.408
Tenure at the firm (months)	36.51	40.08	−3.57**	0.010
Female (= 1)	0.239	0.214	0.024	0.362
Age (years)	38.31	37.33	0.98*	0.070
Education (years)	13.40	13.24	0.16	0.393
Spell length (months)	7.83	7.06	0.77*	0.056
Stores worked in sample	2.49	2.65	−0.17**	0.039

Manager characteristics averaged over the spell; each observation is a manager×store spell. HTM = shrinkage-value fixed effect below the format-by-border cell median. p-values from two-sample t-tests. ***p<0.01, **p<0.05, *p<0.1.

Two implications follow. First, the wage-setting process does not price the quality I estimate: a manager whose arrival reduces shrinkage costs the firm exactly what one whose arrival does not costs. This mirrors the finding of Hoffman and Tadelis (2021) that the people-management skills of supervisors predict performance but not pay, and suggests the firm either cannot observe the quality dimension I recover or does not act on it. Second, because the two types are observationally identical on the traits a personnel file records, the fixed-effect classification of Section 4 is not recoverable from standard human-resources observables.

One determinant of how much that presence disciplines is suggestive but worth recording: the social distance between the manager and the team she watches. Marx et al. (2021), in a Kenyan field experiment with random supervisor assignment, find that vertically homogeneous supervisor–worker pairs exert lower effort—social proximity slackens monitoring, so vertical *distance* sharpens it. My setting has no ethnic dimension, but it records the manager's and the team's gender, age, tenure, and education, so I can interact the arrival effect with the manager–team distance on each (Table 27). The evidence is at most weakly supportive, and I record it as such. The *tenure* dimension—the gap in firm experience, the dimension that most directly marks hierarchical distance—moves the enforcement margin in the predicted direction: a manager whose tenure stands farther from her team's suspends workers more (+0.19 per hundred, $p < 0.05$). But its interaction with shrinkage per worker, negative as the mechanism predicts, is not significant (−0.028, s.e. 0.020), and the gender-distance interaction is significant with the opposite sign for shrinkage value. Because the assignment is quasi-experimental and the pattern is confined to the enforcement margin, I read this as no more than a suggestive echo of the enforcement channel, not as corroboration on the shrinkage margin itself.

Table 27: Manager–team vertical distance and the disciplining effect (suggestive)

	Shrink value	Shrink/ worker	Suspensions per 100
Gender distance	0.027**	0.014	−0.037
	(0.013)	(0.015)	(0.070)
Age distance	−0.002	−0.002	−0.035
	(0.009)	(0.010)	(0.083)
Tenure distance	−0.022	−0.028	0.191**
	(0.020)	(0.020)	(0.079)
Education distance	−0.006	−0.005	−0.066
	(0.011)	(0.013)	(0.077)

Each cell is the $\text{Post}_{[0,8]} \times$ (standardized manager–team distance) interaction from a separate high-type arrival regression (arrival design of Equation 21; store, year, and month fixed effects; standard errors clustered at the store), with the distance fixed at arrival. The gender distance is the share of the team of the opposite gender to the manager; the others are the absolute manager-minus-team-mean gap, z-scored. A negative interaction on shrinkage, or a positive one on suspensions, indicates that a manager more distant from her team disciplines more (Marx et al., 2021). $^{***}p<0.01$, $^{**}p<0.05$, $^{*}p<0.1$.

D.7 Fixed-Effect Comovement: Core-Focused View

Figure 4 presents the comovement of manager fixed effects as a full scatterplot matrix. Figure 15 gives the same evidence in a core-focused layout—the shrinkage-value fixed effect against each other outcome, one panel apiece, with the pairwise correlation shown in each.

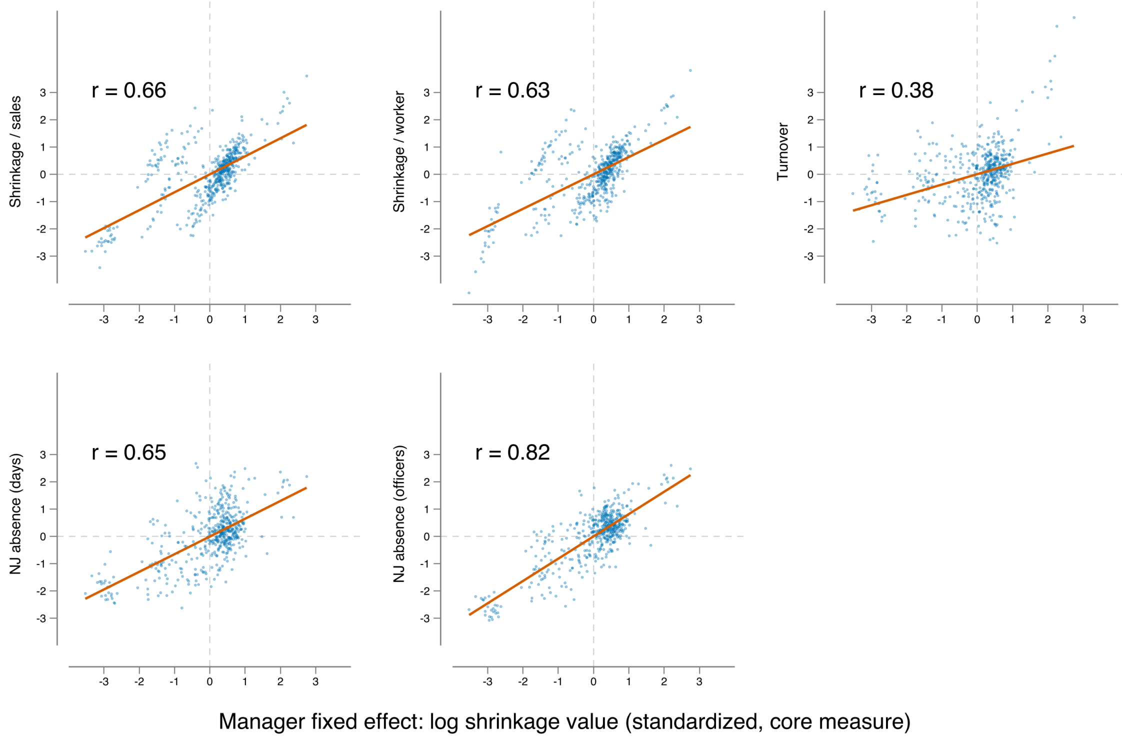


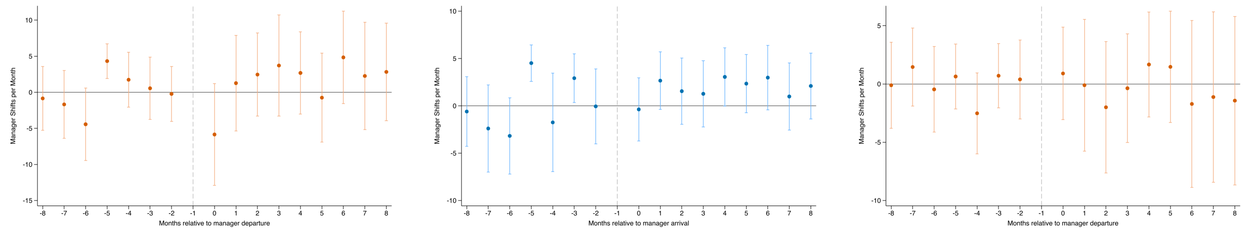
Figure 15: Manager fixed effects: core shrinkage effect against each other outcome

Notes: Each panel plots the manager's standardized AKM fixed effect for log shrinkage value (horizontal) against her fixed effect in one other outcome (vertical). The line is the OLS fit and r the Pearson correlation of the raw fixed effects (matching Figure 4).

D.8 Departures and Low-Type Designs: Event Studies

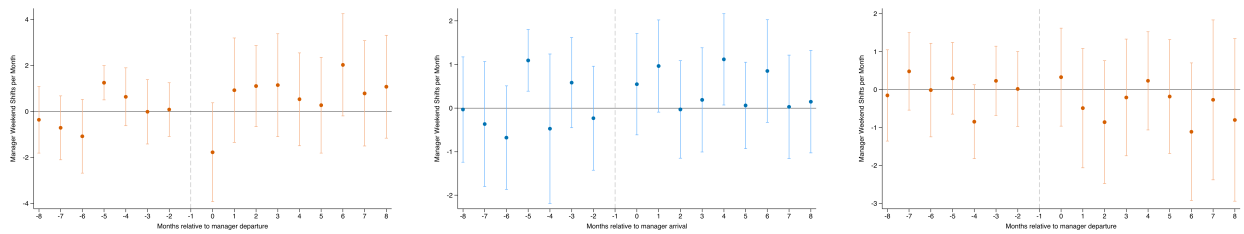
Figures 16–17 report, for every outcome of the main-text arrival figures, the three remaining designs.

Manager shifts



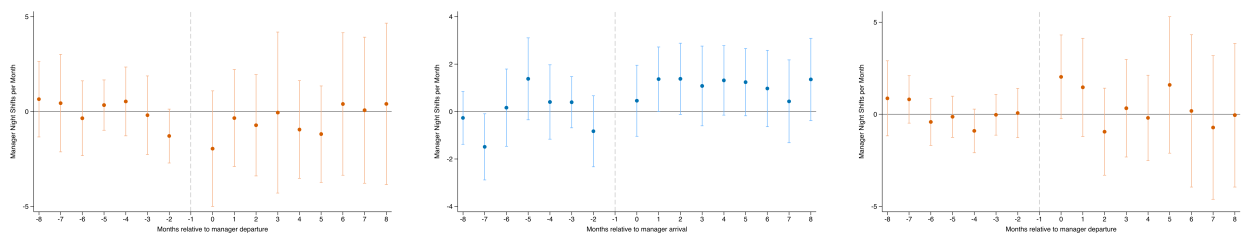
(a) HT departure **(b)** LT arrival **(c)** LT departure

Weekend shifts



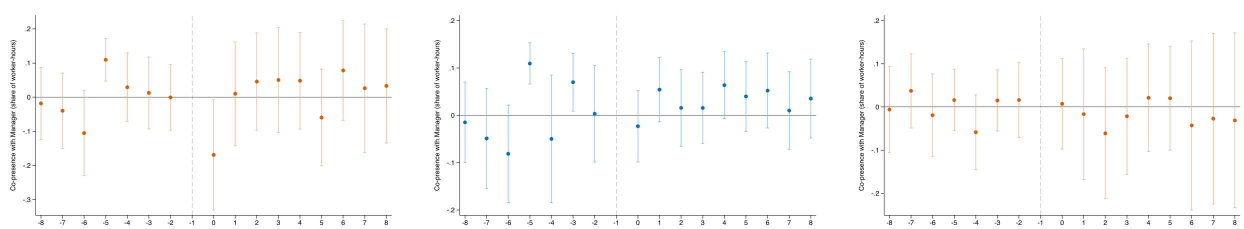
(d) HT departure **(e)** LT arrival **(f)** LT departure

Night shifts



(g) HT departure **(h)** LT arrival **(i)** LT departure

Co-presence

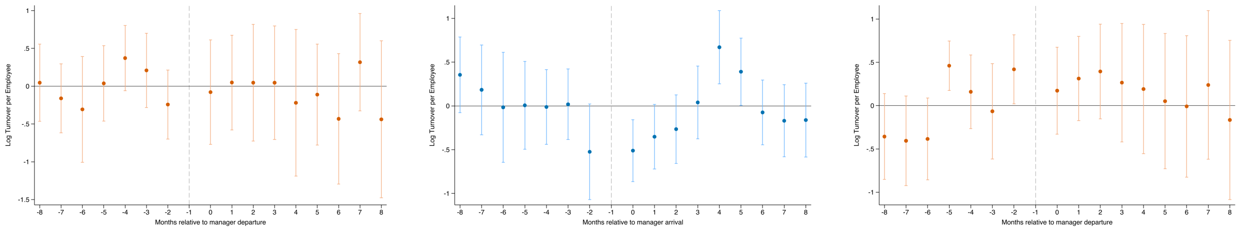


(j) HT departure **(k)** LT arrival **(l)** LT departure

Figure 16: Presence: the remaining designs

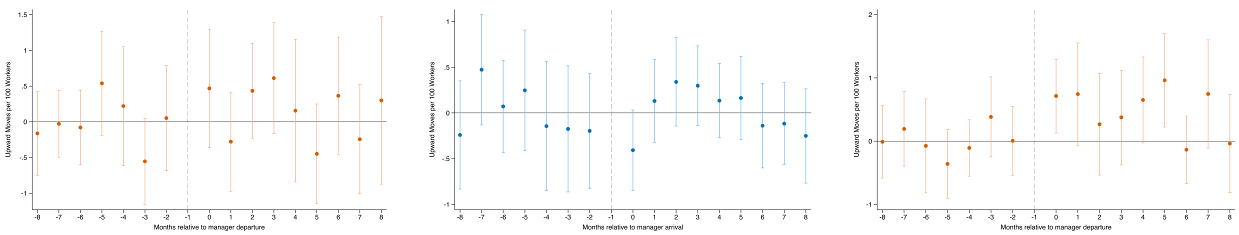
Notes: The three remaining designs for each outcome of the corresponding main-text figure (high-type departures, low-type arrivals, low-type departures), in the conventions of that figure; the departure designs rest on far fewer events and are read as suggestive.

Log turnover



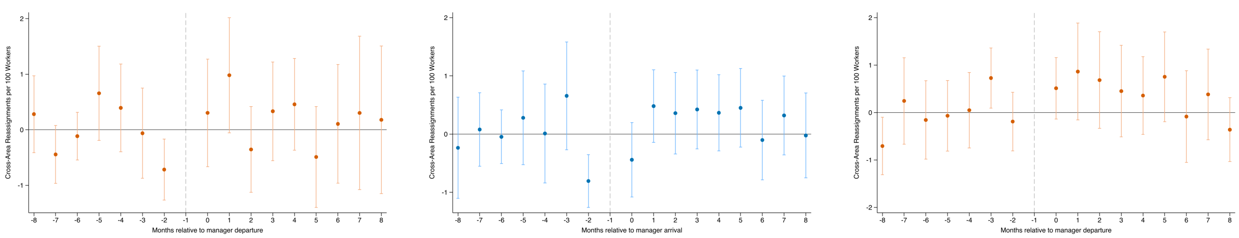
(a) HT departure **(b)** LT arrival **(c)** LT departure

Upward moves



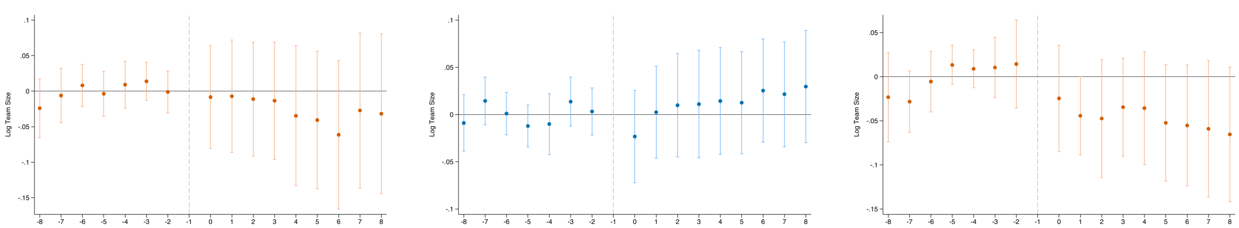
(d) HT departure **(e)** LT arrival **(f)** LT departure

Cross-area moves



(g) HT departure **(h)** LT arrival **(i)** LT departure

Log team size



(j) HT departure **(k)** LT arrival **(l)** LT departure

Figure 17: The team: the remaining designs

Notes: The three remaining designs for each outcome of the corresponding main-text figure (high-type departures, low-type arrivals, low-type departures), in the conventions of that figure; the departure designs rest on far fewer events and are read as suggestive.

E Appendix: Two-Way Fixed-Effects Evidence

The store-level estimates of Section 6 exploit the wide dispersion of local market wages across Mexico. Figure 18 maps this dispersion across municipalities; with pay fixed by position, this geographic variation in the outside option is the main source of variation in the store wage premium.

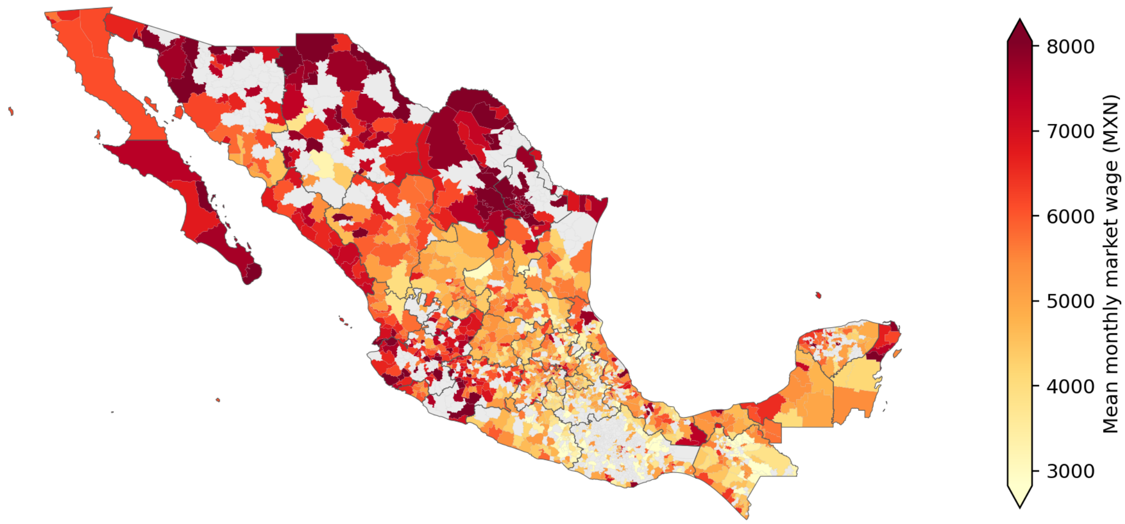


Figure 18: Dispersion of local market wages across Mexico. Each municipality is shaded by its mean monthly market wage, computed from the ENOE labor-force survey over the sample period; municipalities not sampled by ENOE are shown in grey and the color scale is clipped at the 5th and 95th percentiles. The middle 90 percent of municipalities span about USD 150 to USD 440, so the same firm position commands a substantially different wage premium depending on where the store is located.

Table 28 re-estimates the shrinkage gradient of Table 9 with shrinkage normalized by sales, in value and in physical units. The store premium covaries with store scale—stores with a larger premium are smaller—so the ratio removes scale by construction rather than through the log-sales control of column 3. The gradient is negative in every specification for both measures.

Table 28: Store-level TWFE: shrinkage normalized by sales (Log–Log)

Wage premium (Log)	(1)	(2)	(3)	(4)
Panel A. Log shrinkage-to-sales (value)				
Coefficient	-0.062^{**}	-0.044^{*}	-0.045^{**}	-0.063^{**}
SE	(0.027)	(0.022)	(0.021)	(0.026)
Mean	0.82	0.82	0.82	0.82
Median	0.90	0.90	0.90	0.90
Panel B. Log shrinkage-to-sales (physical units)				
Coefficient	-0.146^{**}	-0.100^{*}	-0.081	-0.145^{**}
SE	(0.057)	(0.052)	(0.048)	(0.056)
Mean	0.22	0.22	0.22	0.22
Median	0.48	0.48	0.48	0.48
Store format	No	Yes	Yes	No
Log sales	No	No	Yes	No
Municipality FE	Yes	Yes	Yes	Yes
Year×Qtr FE	Yes	Yes	Yes	No
Year + Quarter FE	No	No	No	Yes

Store-quarter level; the sample of Table 9. Panel A: log of shrinkage value over sales value (percent); Panel B: log of shrinkage units over units sold (percent). Mean and Median are of the log outcome. Standard errors clustered at the municipality level. $^{***}p<0.01$, $^{**}p<0.05$, $^{*}p<0.1$.

Table 29 reports the store-level gradients of Table 10 under the four specifications of Table 9.

Table 29: Store-level TWFE: attendance, turnover, and effort across specifications

	Absenteeism	Non-justified absenteeism	Turnover	Monthly shifts	Daily hours
Coefficient on the log wage premium					
(1)	−0.036***	−0.129***	−0.158***	+0.032***	+0.012***
	(0.010)	(0.023)	(0.025)	(0.009)	(0.002)
(2)	−0.038***	−0.131***	−0.158***	+0.032***	+0.011***
	(0.011)	(0.023)	(0.025)	(0.009)	(0.001)
(3)	−0.033***	−0.125***	−0.162***	+0.032***	+0.010***
	(0.010)	(0.022)	(0.025)	(0.009)	(0.001)
(4)	−0.039***	−0.127***	−0.140***	+0.030***	+0.011***
	(0.011)	(0.021)	(0.021)	(0.009)	(0.002)
Fixed effects and controls					
(1)	Municipality FE; year×quarter FE				
(2)	(1) + store format				
(3)	(1) + store format + log sales				
(4)	Municipality FE; year FE + quarter FE				
Mean	0.68	0.00	2.55	2.98	2.10
Median	0.71	0.06	2.69	2.99	2.10

Store-quarter level. Each cell is the coefficient on the log store wage premium from a separate regression of the log outcome, with the standard error below in parentheses; rows (1)–(4) are the specifications of Table 9. Turnover is a rate per hundred workers. Mean and Median are of the log outcome. Standard errors clustered at the municipality level. $^{***}p<0.01$, $^{**}p<0.05$, $^{*}p<0.1$.

E.1 Worker-Level Results

The store-level estimates of Section 6 aggregate the responses of each store's workforce. Here I follow individual workers, re-estimating Equation (22) at the worker level: the outcome is a worker's own total absence days, non-justified absence days, monthly shifts, or average daily hours; the premium is her own premium from Equation 1; a worker fixed effect replaces the municipality fixed effect and store fixed effects are also absorbed; Z_{jt} optionally includes managerial tier, tenure, and age group; and standard errors are clustered at the store level.

Table 30 reports the estimates. Within the same worker, a larger premium goes with fewer absence days, total and non-justified, and with more monthly shifts and longer daily hours; the estimates are stable across control sets and across alternative time fixed effects.

Table 30: Worker-level TWFE: continuous outcomes (log wage premium)

	Absence days	Non-justified absence days	Monthly shifts	Daily hours
Coefficient on the log wage premium				
(1)	-0.048^{***}	-0.046^{***}	$+0.046^{***}$	$+0.0093^{***}$
	(0.003)	(0.002)	(0.001)	(0.0003)
(2)	-0.050^{***}	-0.047^{***}	$+0.046^{***}$	$+0.0069^{***}$
	(0.003)	(0.002)	(0.001)	(0.0003)
(3)	-0.049^{***}	-0.047^{***}	$+0.046^{***}$	$+0.0069^{***}$
	(0.003)	(0.002)	(0.001)	(0.0003)
(4)	-0.045^{***}	-0.038^{***}	$+0.045^{***}$	$+0.0086^{***}$
	(0.003)	(0.002)	(0.001)	(0.0003)
(5)	-0.015^{***}	-0.032^{***}	$+0.039^{***}$	$+0.0081^{***}$
	(0.003)	(0.002)	(0.001)	(0.0003)
Fixed effects and controls				
(1)	Worker FE; store FE; year×quarter FE			
(2)	(1) + tier + tenure			
(3)	(1) + tier + tenure + age group			
(4)	Worker FE; store FE; year FE			
(5)	Worker FE; store FE; quarter FE			
Mean	0.61	0.22	2.77	2.09
Median	0.69	0.00	3.11	2.07
Mean (level)	2.03	1.09	19.28	8.21
Median (level)	1.00	0.33	22.33	7.91

Worker-quarter level. Each cell is the coefficient on the log monthly wage premium (mean USD 93) from a separate regression of the log outcome, with the standard error below in parentheses. Mean and Median are of the log outcome; the level rows report the outcome in its original units. Standard errors clustered at the store level. $^{***}p<0.01$, $^{**}p<0.05$, $^{*}p<0.1$.

Table 31 reports logit estimates for the discrete worker-level outcomes. A higher wage premium is associated with lower odds of any absence, any non-justified absence, and separation.

Table 31: Worker-level logit: discrete outcomes (log wage premium)

	Any absence	Any non-justified absence	Separation
Coefficient on the log wage premium (log-odds)			
(1)	-0.201^{***}	-0.436^{***}	-0.370^{***}
	(0.009)	(0.010)	(0.010)
(2)	-0.182^{***}	-0.414^{***}	-0.369^{***}
	(0.009)	(0.010)	(0.010)
(3)	-0.182^{***}	-0.380^{***}	-0.323^{***}
	(0.009)	(0.010)	(0.010)
(4)	-0.148^{***}	-0.264^{***}	-0.314^{***}
	(0.009)	(0.010)	(0.011)
(5)	-0.149^{***}	-0.283^{***}	-0.322^{***}
	(0.009)	(0.011)	(0.011)
Controls and fixed effects			
(1)	None		
(2)	Education		
(3)	Education + tenure		
(4)	Education + tenure + tier		
(5)	Education + tenure + tier; time FE		
Mean	0.70	0.57	0.15
Median	1	1	0

Worker-quarter level; logit coefficients (log-odds) on the log monthly wage premium, with the standard error below in parentheses. Each outcome is an indicator equal to one in the quarter. Mean and Median are of the indicator. Standard errors clustered at the store level. $^{***}p<0.01$, $^{**}p<0.05$, $^{*}p<0.1$.

Table 32 places the store (S) and worker (W) coefficients side by side for the four outcomes observed at both levels. The two levels tell the same story: a larger wage premium is associated with less absenteeism and more time on the job, whether I aggregate to the store or follow individual workers.

Table 32: Comparative TWFE: effect of the log wage premium (baseline)

	Absenteeism		NJ Absenteeism		Daily Hours		Monthly Shifts	
Wage premium (Log)	(S)	(W)	(S)	(W)	(S)	(W)	(S)	(W)
Coefficient	-0.036^{***}	-0.048^{***}	-0.129^{***}	-0.046^{***}	$+0.012^{***}$	$+0.009^{***}$	$+0.032^{***}$	$+0.046^{***}$
SE	(0.010)	(0.003)	(0.023)	(0.002)	(0.002)	(0.0003)	(0.009)	(0.001)
Mean	0.68	0.61	0.00	0.22	2.10	2.09	2.98	2.77
Median	0.71	0.69	0.06	0.00	2.10	2.07	2.99	3.11

(S) = store-quarter level (municipality and year×quarter FE). (W) = worker-quarter level (worker, store, and year×quarter FE). No controls. Mean and Median are of the log outcome. Clustered standard errors in parentheses (municipality for (S), store for (W)). $^{***}p<0.01$, $^{**}p<0.05$, $^{*}p<0.1$.

F The Raise and the Squeeze, Worker by Worker

Here is what the paper knows so far, and what it cannot yet see. Section 7 shows that when the statute squeezes a border store's premium, shrinkage rises months later, and its closing anatomy names the reason a store-level number cannot be the last word: the same January morning *raises* the wage of the worker below the incoming floor and *narrows* the premium of her coworker just above it, and a store-level regression averages those two experiences into one number. What is missing is the worker-level view: who inside the store responds, on which margin, and on whose clock.

The payroll supplies it, because every worker's distance to the incoming floor is visible in November, before the floor arrives. I classify each full-time worker by the ratio of her daily salary to the January floor, and the four groups are the cast of the story. The *swept* (ratio below 1) are overtaken by the floor: the statute hands them a raise they did not bargain for. The *compressed-near* (1 to 1.3) keep their wage while the floor closes in from below—the squeezed workers of the store section, seen individually. The *compressed-far* (1.3 to 1.6) lose ground too, but from a safer distance, which makes them a built-in dose gradient. The *cushion* (1.6 and above) stand beyond the reach of the floor and of the firm's spillover raises, and serve as the control group.[11] I track each group through the same two-phase clock as the store design. Identification is within store-month: a worker-by-hike fixed effect absorbs everything permanent about the worker, a store-by-month-by-hike fixed effect absorbs everything that happens to her store in a given month—no store or border shock can masquerade as a group effect—and each group is measured against the cushion coworkers of the same store in the same month where the cushion exists (see the note above on its thinness at the border), with standard errors clustered by store.

F.1 Results: The Worker-Level Clock Matches the Store's

What should each group do? The model of Section 3 makes one prediction per group; the clock is not the model's but the premium's, whose path each observation feeds through the comparative statics one period at a time. The swept receive a large unbargained raise, so on gift-exchange and no-shirking logic alike their discipline should *improve* from the

[11] The classification is predetermined and held fixed through each event window. I set the cushion cutoff where the firm's own post-hike raise schedule flattens out, so the control is not itself treated through internal spillovers (Autor et al., 2016; Dustmann et al., 2022); measurement details of the salary concept follow the payroll conventions of Section 2. At the border the cushion is thin by construction—the compression leaves 2.1 percent of border workers at or above 1.6 times the incoming floor in the 2023 hike, and none in the 2024 hike—so border-level group *levels* lean on the 2023 cohort, and the results I read at the border are the within-border contrasts and the dose–response, which difference out the common store-month. The thin border cushion is not a nuisance of the design but its point: no rung of the border ladder escapes the floor.

month the raise lands—a raise is an event. The compressed lose rent every month the new floor persists, so their discipline should *erode*, slowly—a squeeze is a state; and the compressed-far, who lose more ground, should erode more. The cushion, untouched, pin the counterfactual. The margins I measure are the ones a worker controls and the time-clock verifies: the probability of an unjustified absence and unjustified absence days (the care margin the worker internalizes), and suspensions (the firm's own disciplinary response). Shrinkage does not appear here because it is a store outcome, already measured in the store section.

Read Table 33 as follows. Each row is an outcome; the columns report each group's effect relative to its cushion coworkers, separately for the early phase (January–February, the raise's window) and the late phase (March–June, the squeeze's window); the rows d_{21} and d_{31} report the contrasts compressed-near−swept and compressed-far−swept within phase, which is where the squeeze shows once the common store month is netted out.

Three results organize the table, one per margin. On the worker's care margin, the swept repay the raise: their probability of an unjustified absence falls by two percentage points and their unjustified absence days by about four percent of a day per month, against a median own raise of roughly a quarter—the reciprocity a large, unbargained raise invites (Akerlof, 1982). On the firm's margin, the employer disciplines the swept more: suspensions rise in both phases, the firm's response to a worker whose statutory cost just rose. And on the clock—the exercise's central result—the compressed fall behind the swept exactly when the store section says the squeeze binds: d_{21} is statistically zero in January and February and opens to 1.3 percentage points of absence probability and 0.03 days in March through June, with the same late-phase pattern, larger still, for the compressed-far (d_{31})—the dose gradient inside the store. The raise is an event whose reciprocity arrives with it and builds through the spring; the erosion of position is a state whose behavioral cost accumulates (Card et al., 2012; Breza et al., 2018).

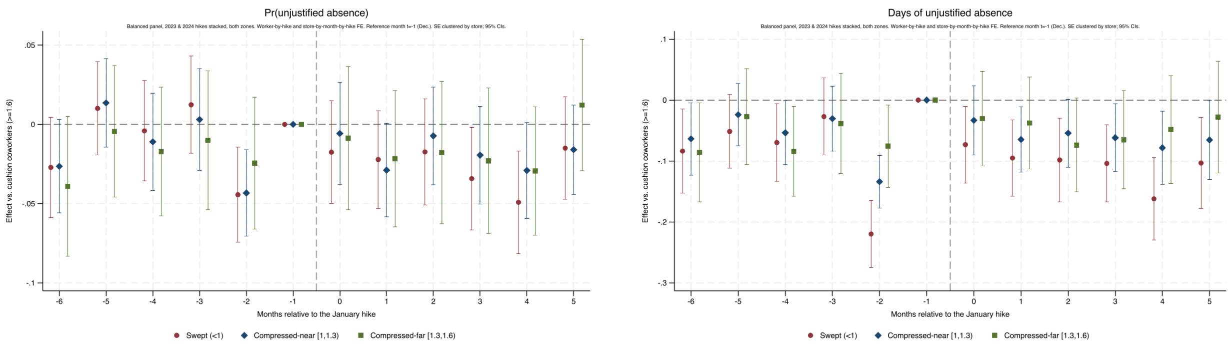
Table 33: Exposure group × phase: worker-level attendance and discipline

	Swept	Compr. near	Compr. far	d_{21}	d_{31}
Panel A. Pr(unjustified absence)					
Early$_{[0,1]}$	−0.011	−0.007	0.001	0.004	0.012
	(0.010)	(0.009)	(0.014)	(0.007)	(0.011)
Late$_{[2,5]}$	−0.020**	−0.007	0.001	0.013**	0.021**
	(0.008)	(0.007)	(0.011)	(0.006)	(0.010)
Panel B. Days of unjustified absence					
Early$_{[0,1]}$	−0.009	0.002	0.018	0.011	0.027
	(0.020)	(0.017)	(0.023)	(0.016)	(0.022)
Late$_{[2,5]}$	−0.041**	−0.014	−0.002	0.028**	0.040*
	(0.017)	(0.015)	(0.022)	(0.014)	(0.021)
Panel C. Pr(suspension)					
Early$_{[0,1]}$	0.004**	0.003*	0.003	−0.001	−0.001
	(0.002)	(0.002)	(0.003)	(0.002)	(0.003)
Late$_{[2,5]}$	0.005***	0.003*	−0.000	−0.002	−0.005**
	(0.002)	(0.002)	(0.002)	(0.001)	(0.002)

Balanced panel, 2023 and 2024 hikes stacked, both zones. Group effects are measured against the cushion control; d_{21} and d_{31} are the compressed-minus-swept contrasts within phase. Early = January–February; Late = March–June. Worker-by-hike and store-by-month-by-hike fixed effects; standard errors clustered by store. $^{***}p<0.01$, $^{**}p<0.05$, $^{*}p<0.1$.

Figure 19 reads the same regression on a monthly clock: each point is an exposure-group-by-event-month coefficient, measured against cushion coworkers in the same store and month, with December as the reference month. The swept improvement arrives on impact and builds through the spring—unjustified absence days fall by 0.07 in January and by 0.16 in month four—and the compressed fall behind late: the compressed-minus-swept gap in absence days is positive in every post-hike month but is largest in months four and five (0.084 days for the near group, $p = 0.004$; 0.114 for the far, $p = 0.010$)—the dose gradient of the phase table, month by month. The monthly read also shows why the table draws its inference from pooled phases against the full pre-hike window: the November point sits below December for the swept and compressed-near groups—the payroll's year-end seasonal anchor—while months −6 through −3 are flat, and the compressed-minus-swept contrasts, which difference out the common seasonality, show no pre-hike pattern on the absence-probability margin.

Figure 19: The minimum-wage event study, split by how much the hike cuts the worker's premium



(a) Pr(unjustified absence) **(b)** Days of unjustified absence

Notes: Monthly event-study version of Table 33, same sample and specification: balanced worker panel, 2023 and 2024 hikes stacked, both zones. Each point is the coefficient on an exposure-group × event-month interaction, measured against cushion coworkers (≥ 1.6 times the incoming floor) in the same store and month; the omitted month is $t = -1$ (December); groups are fixed in the pre-hike November payroll. Worker-by-hike and store-by-month-by-hike fixed effects; standard errors clustered by store; bars are 95 percent confidence intervals; the dashed vertical line marks the hike. Joint tests of the pre-hike months reject for the swept and compressed-near groups ($p < 0.01$), driven entirely by the November point—the seasonal November–December anchor—with months -6 to -3 flat; the compressed-minus-swept contrasts are clean in panel (a).

The gap is also dose-responsive, and the dose is the border. With a border increment added to each group effect—identified within store-month, so no border-wide shock can contaminate it—the compressed–swept gap is approximately zero in the interior, where compression is mild, and 3.0 percentage points at the border (0.090 days), where the same position loses far more ground to the floor; the formal dose–response contrast is 2.2 points (0.085 days). The swept improvement, by contrast, is the same in both zones: reciprocity to an own raise is a level effect, indifferent to the dose.

Table 34: The compressed–swept gap, by zone: the border dose–response

	d_{21} (Compr. near − Swept)		Dose–response (DDD)	
Outcome	Interior	Border	$F_2 - F_1$	$F_3 - F_1$
Pr(unjustified absence)	0.008	0.030***	0.022*	0.006
	(0.006)	(0.011)	(0.012)	(0.019)
Days of unjustified absence	0.006	0.090***	0.085***	0.080*
	(0.013)	(0.027)	(0.030)	(0.044)

Group × post effects split into an interior component and a border increment; worker-by-hike, store-by-month-by-hike, and position-by-month-by-hike fixed effects; the interior and border d_{21} and the dose–response contrasts are linear combinations from a single regression; balanced panel, pooled post window; standard errors clustered by store. $^{***}p$<0.01, $^{**}p$<0.05, $^{*}p$<0.1.

G The Mechanism of the Complementarity: The Full Battery

Section 8.3 adjudicates between two readings of the wage-premium complementarity by asking whether the manager's conduct scales with the rent, and reports a compact summary of the evidence. This appendix collects the full battery behind that summary, and adds a sharper cut at the composition channel. Both exercises use the high-type classification of Section 5 and the arrival design of Equation 21, replacing the shrinkage outcome with each mechanism in turn and interacting the $\text{Post}_{[0,8]}$ indicator with the wage environment of the store's municipality at arrival.

G.1 Every Mechanism, Crossed with the Premium

Table 35 takes each margin examined in Section 5 and estimates its high-type arrival effect separately in low- and high-premium stores, alongside the difference between the two and its p-value. The table groups the margins into six panels: the organization of monitoring (Panel A), the composition and Harrison–Klein diversity of the team (Panel B), the internal reallocation of workers (Panel C), the manager's own presence (Panel D), the enforcement she exercises (Panel E), and the shrinkage return (Panel F). The first four panels collect the manager's inputs; the last two collect the enforcement and the return.

The pattern is one-directional and confirms the reading of the main text. The manager's inputs do not scale with the premium. The organization of monitoring is flat—the loss-prevention, supervisory, and security shares of the workforce move by statistically indistinguishable amounts in the two environments, with the single exception of the supervisory share, whose differential is modest and isolated among the two dozen input margins in the table. Team composition and diversity are flat throughout Panel B: neither the size, the gender or occupational mix, nor any Harrison–Klein construct responds to the premium. The internal reallocation of Panel C—rotation, upward moves, and cross-area moves—rises on a high-type arrival, as Section 5.3 documents, but the rise is no larger where the premium is high. Most tellingly, the manager's own presence in Panel D does not scale: overall co-presence, the presence targeted at the shrinkage-prone roles—the loss-prevention chain, cashiers, perishables, and security—and presence in the low-oversight night and weekend hours all move by amounts that are statistically indistinguishable across the premium. The faint upward tilts that survive—overall co-presence and manager shifts drift a little higher where the premium is high—do not reach the sharpness of the enforcement and return margins, and they run against, not with, the reallocation of attention that the supply reading would require. The manager deploys the same supervisory technology regardless of the rent.

What scales is the enforcement and the return. The disciplinary instrument the manager controls day to day—the suspension—rises after a high-type arrival only where the premium is high (+0.31 per hundred workers in high-premium stores against −0.14 in low-premium, a differential of 0.45 significant at five percent), while the dismissal, which the manager does not control as directly, does not respond. The shrinkage reduction is two to three times larger in high-premium stores on the value and per-worker measures; the sales ratio moves the same way from a near-zero low-premium base.

Table 35: The full channel of the complementarity: the manager's inputs do not scale with the premium, the enforcement and the return do (Part 1: organization, composition, reallocation)

	Low-premium (LWP)	High-premium (HWP)	Difference (HWP−LWP)	p(diff)
Panel A. Monitoring organization				
Loss-prevention staff /100	0.340**	0.182	-0.158	0.492
	(0.161)	(0.187)	(0.230)	
Supervisory share (%)	-0.503	0.931	1.434**	0.024
	(0.407)	(0.629)	(0.636)	
Security-staff share (%)	0.117	0.038	-0.079	0.789
	(0.183)	(0.289)	(0.295)	
Panel B. Team composition and Harrison–Klein diversity				
Log team size	0.001	0.023	0.022	0.237
	(0.011)	(0.019)	(0.019)	
Female share (%)	-1.277***	-0.471	0.805	0.202
	(0.456)	(0.560)	(0.631)	
Sales-floor share (%)	-0.079	-0.189	-0.110	0.874
	(0.377)	(0.676)	(0.699)	
Gender variety (Blau)	0.003**	0.001	-0.002	0.362
	(0.001)	(0.002)	(0.003)	
Tier variety (Blau)	-0.002	0.005	0.007	0.137
	(0.003)	(0.005)	(0.005)	
Occupational-area variety (Blau)	0.000	0.006	0.006	0.325
	(0.003)	(0.005)	(0.006)	
Age dispersion (SD)	0.069	0.093	0.024	0.880
	(0.079)	(0.144)	(0.157)	
Tenure dispersion (SD)	-0.363	-0.315	0.048	0.950
	(0.504)	(0.689)	(0.761)	
Education dispersion (SD)	-0.038	-0.013	0.025	0.435
	(0.026)	(0.024)	(0.032)	
Pay dispersion (CV)	0.010	0.000	-0.010	0.368
	(0.007)	(0.009)	(0.011)	
Panel C. Internal reallocation				
Internal rotation /100	0.132	0.349*	0.217	0.453
	(0.207)	(0.204)	(0.289)	
Upward moves /100	0.208**	0.131	-0.077	0.504
	(0.084)	(0.082)	(0.115)	
Cross-area moves /100	0.173	0.270*	0.097	0.608
	(0.112)	(0.155)	(0.190)	

Part 1 of 2; see the notes to Part 2 (Table 36).

Table 36: The full channel of the complementarity: the manager's inputs do not scale with the premium, the enforcement and the return do (Part 2: presence, enforcement, and the return)

	Low-premium (LWP)	High-premium (HWP)	Difference (HWP−LWP)	p(diff)
Panel D. The manager's own presence				
Manager–worker co-presence (share)	0.020	0.063***	0.043	0.102
	(0.014)	(0.022)	(0.026)	
Any manager overlap (share)	0.023	0.077***	0.054*	0.078
	(0.016)	(0.025)	(0.031)	
Manager shifts per month	0.916	3.041***	2.126*	0.059
	(0.585)	(0.913)	(1.126)	
Co-presence: loss-prevention chain	0.022	0.045*	0.023	0.414
	(0.017)	(0.024)	(0.028)	
Co-presence: cashiers	0.021	0.060***	0.040	0.135
	(0.015)	(0.022)	(0.027)	
Co-presence: perishables	0.012	0.069**	0.057*	0.064
	(0.016)	(0.027)	(0.031)	
Co-presence: security	0.018	0.049**	0.031	0.197
	(0.012)	(0.021)	(0.024)	
Manager presence, night hours (%)	0.017	-0.003	-0.020	0.309
	(0.013)	(0.017)	(0.020)	
Manager presence, weekend hours (%)	-0.000	-0.007	-0.006	0.379
	(0.004)	(0.006)	(0.007)	
Panel E. Enforcement				
Suspensions /100	-0.140	0.311**	0.451**	0.013
	(0.139)	(0.153)	(0.181)	
Dismissals /100	-0.008	0.019	0.027	0.852
	(0.090)	(0.124)	(0.145)	
Panel F. Shrinkage (the return)				
Log shrinkage value	-0.043**	-0.126***	-0.084**	0.022
	(0.017)	(0.032)	(0.037)	
Log shrinkage / sales	-0.015	-0.076***	-0.061*	0.065
	(0.021)	(0.027)	(0.033)	
Log shrinkage / worker	-0.037**	-0.138***	-0.101***	0.008
	(0.018)	(0.035)	(0.038)	

Notes: Each row is a separate high-type arrival regression with the indicated store-month outcome, following the arrival design of Equation 21 (store, calendar-month, and year fixed effects; standard errors clustered at the store; ever-treated stores). The wage environment is fixed at arrival (high-premium, HWP, = a wage premium above the municipality's median). The first two columns report the arrival effect in low- and high-premium stores, with stars reflecting each effect's own significance; the third column is the difference between them and the last is the p-value that the two differ—the premium-scaling test. Rates are per hundred workers. $^{***}p<0.01$, $^{**}p<0.05$, $^{*}p<0.1$.

G.2 Within- and Between-Area Diversity

The evidence that the manager's inputs are flat across the premium closes one door that the reallocation results of Section 5.3 might seem to leave open. High-type managers reassign workers across occupational areas, and a reshuffle across areas can in principle change team diversity *within* each area even while leaving it unchanged at the store level: mixing areas raises within-area heterogeneity and lowers between-area heterogeneity, with the store total held fixed (Harrison and Klein, 2007). Table 37 tests this directly. I rebuild each Harrison–Klein construct *within* occupational area—the separation measures (the standard deviation of age, tenure, and education), the variety measures (the Blau index of gender and of tier), and the disparity measure (the coefficient of variation of pay)—as the headcount-weighted mean across areas, together with the corresponding between-area dispersion, and run each through the arrival design. This exercise asks only whether the reshuffling changes composition inside each area, so it is not crossed with the premium.

A high-type arrival moves none of them. Within-area age, tenure, and education dispersion, within-area gender and tier variety, within-area pay dispersion, and their between-area counterparts are all unchanged, and the low-type arrival, reported for contrast, moves them no more. The only movement to approach conventional significance is within-area education dispersion, which edges down on either arrival by a similar amount and so does not distinguish the high-type manager. The composition of the workforce is intact not only at the store level but inside each team the manager runs. I interpret this as evidence that the manager deploys the workers she inherits rather than recomposing them at any level of aggregation, consistent with the flat composition inputs of Table 35.

Table 37: Within- and between-area diversity do not change on a high-type arrival

Diversity measure	HT arrival	LT arrival
Within-area age SD	0.0208	-0.0550
	(0.0956)	(0.0866)
Within-area tenure SD	-0.1331	-0.4963
	(0.4156)	(0.4598)
Within-area education SD	-0.0483*	-0.0470**
	(0.0247)	(0.0218)
Within-area gender variety (Blau)	0.0032	0.0017
	(0.0035)	(0.0034)
Within-area tier variety (Blau)	0.0015	0.0018
	(0.0035)	(0.0035)
Within-area pay CV	0.0034	-0.0054
	(0.0050)	(0.0036)
Between-area age SD	0.0969	-0.0927
	(0.1606)	(0.1409)
Between-area tenure SD	-0.9678	0.1094
	(0.6275)	(0.4877)
Between-area pay SD	102.4159	218.5141
	(166.3813)	(161.2184)

Each cell is the $\text{Post}_{[0,8]}$ coefficient from a separate store-month arrival regression following Equation 21 (store, calendar-month, and year fixed effects; standard errors clustered at the store; ever-treated stores). Within-area measures are the headcount-weighted mean across occupational areas of each within-area Harrison–Klein construct (Harrison and Klein, 2007); between-area measures are the standard deviation of area means across areas. $^{***}p<0.01$, $^{**}p<0.05$, $^{*}p<0.1$.

G.3 The Complementarity at the Worker-Shift Level

The complementarity so far lives at the level of the store's aggregate losses. If it is the reinforcement that Shapiro and Stiglitz (1984)- and Coviello et al. (2022)-style models describe—the manager's supervision and the worker's rent combining in the production of *effort*—it should also be visible in how individual workers respond to being watched, at the daily frequency the store-month results aggregate over. Appendix D.5 isolated one clean worker-level response of that kind: punctuality—the same worker arrives closer to her own norm on days the general manager is on the clock. Here I ask whether that punctuality response to supervision is itself stronger where the premium is high. I return to the shift-level discipline panel and interact the manager's co-presence with the store's high-premium status, absorbing worker, store×month, arrival-hour, and day-of-week fixed effects, so that identification is the same worker, at the same time of day, on days the manager is present versus absent, compared across high- and low-premium environments.

Table 38 reports the probe, and the evidence is suggestive rather than decisive. On the clean punctuality margin the response to supervision is about twice as large in high-premium stores: co-presence pulls arrivals 1.4 minutes closer to the norm where the premium is low and 2.8 minutes where it is high, the direction the complementarity predicts. The interaction itself, however, is imprecise ($p \approx 0.11$), and the continuous-premium version is not consistently signed, so the gradient is a hint rather than a result. (The table reports the early-departure column for completeness; as in Appendix D.5 that margin is entangled with shift structure and carries little weight.) I therefore read the effort-margin evidence as consistent with, not independent confirmation of, the complementarity: the clean version of the result lives at the store level (Table 11), and the daily worker-shift data can only hint that the same reinforcement operates on the effort margin where the theory locates it. Establishing it would require richer premium variation at the worker level or a higher-frequency effort measure than punch-clock timing.

Table 38: The wage × supervision complementarity at the worker-shift level (suggestive)

	Late arrival (min)	Early departure (min)
GM co-present, low-premium store	-1.36^{*}	-13.10^{***}
	(0.74)	(1.04)
× high premium (HWP)	-1.42	1.20
	(0.88)	(0.99)
GM co-present, high-premium store (total)	-2.78^{***}	-11.90^{***}
	(0.81)	(1.02)
GM co-present × continuous premium (z)	1.78	-1.49
	(1.15)	(1.28)
Worker-shifts	≈8.7 million	

Worker-shift regressions on the discipline panel of Appendix D.5. "GM co-present" indicates the manager's presence interval overlaps the shift; outcomes are deviations (minutes) from the worker's own median arrival and departure. The × HWP row is the additional effect in high-premium stores and the total their sum; the last row uses the continuous premium z-score. Worker, store×month, arrival-hour, and day-of-week fixed effects; standard errors clustered at the store. This probe does not depend on the manager classification. $^{***}p<0.01$, $^{**}p<0.05$, $^{*}p<0.1$.

H Wage-Premium Complementarity: Robustness

Table 39 reports the full outcome battery of the arrival-by-environment design; the body reports log shrinkage value only.

Table 39: High-type manager arrivals by wage environment at arrival: full battery

	(1) Log shrink value	(2) Log shrink/ sales	(3) Log shrink/ worker	(4) Log NJ abs. p.e.	(5) Log turnover p.e.
HT arrival, high-premium (HWP)	−0.126***	−0.076***	−0.138***	−0.049	0.309***
	(0.032)	(0.027)	(0.035)	(0.053)	(0.115)
HT arrival, low-premium (LWP)	−0.043**	−0.015	−0.037**	0.018	0.115
	(0.018)	(0.021)	(0.018)	(0.030)	(0.080)
Difference (HWP − LWP)	−0.084**	−0.061*	−0.101***	−0.066	0.194
	(0.037)	(0.033)	(0.038)	(0.064)	(0.126)
Dep. mean	12.43	0.84	8.31	−0.24	−3.32

Notes: Full-outcome version of the upper panel of the body table: $\text{Post}_{[0,8]}$ interacted with the wage environment fixed at the arrival month, for the five outcomes in columns. Store, calendar-month, and year fixed effects; standard errors clustered at the store; ever-treated stores only. $^{***}p<0.01$, $^{**}p<0.05$, $^{*}p<0.1$.

Table 40: Low-type manager arrivals by wage environment at arrival: full battery

	(1) Log shrink value	(2) Log shrink/ sales	(3) Log shrink/ worker	(4) Log NJ abs. p.e.	(5) Log turnover p.e.
LT arrival, high-premium (HWP)	0.037	0.046	0.030	−0.037	−0.083
	(0.033)	(0.033)	(0.034)	(0.046)	(0.105)
LT arrival, low-premium (LWP)	0.059***	0.080***	0.075***	−0.000	0.087
	(0.021)	(0.022)	(0.023)	(0.025)	(0.071)
Difference (HWP − LWP)	−0.022	−0.035	−0.045	−0.037	−0.170
	(0.039)	(0.037)	(0.041)	(0.049)	(0.126)
Dep. mean	12.46	0.85	8.39	−0.28	−3.44

Notes: As in Table 39, for low-type arrivals. $^{***}p<0.01$, $^{**}p<0.05$, $^{*}p<0.1$.

H.1 Alternative Definitions of the Wage Environment

The complementarity should be a property of the economics, not of any one way of defining the wage environment. I therefore re-estimate the high-type interaction under three additional definitions, all fixed at the arrival month (Table 41), each using a different slice of the premium's variation. The first is a cross-sectional split—whether the store's municipality lies above the same-quarter median premium of *other* municipalities—which by construction carries no time trend. The second uses the standardized premium itself as a continuous moderator, in the spirit of the continuous relative wage of Cappelli and Chauvin (1991). The third replaces the premium with a cost-of-job-loss measure in the tradition of Schor and Bowles (1987), scaling the outside wage by the reemployment probability implied by the municipal unemployment rate. The complementarity holds under all four definitions. Under the cross-sectional split the high-type effect on shrinkage value is -8.0 percent in high-premium against -5.1 percent in low-premium stores; each standard deviation of the arrival premium amplifies the effect on shrinkage-to-sales by about 3.6 percentage points; and each standard deviation of the cost of job loss by about 5.5. The magnitudes move with how much of the premium's variation each definition uses, as they should; the sign and the ordering of the two environments—the complementarity itself—are invariant across every definition, which is the mark of a result that belongs to the setting rather than to a specification.

Table 41: High-type arrivals under four definitions of the wage environment

	(1) Log shrink value Log turnover p.e.	(2) Log shrink/ sales
A. Binary, within-municipality median (main definition)		
HT arrival in HWP	−0.126***	−0.076***
	(0.032)	(0.027)
Difference (HWP − LWP)	−0.084**	−0.061*
	(0.037)	(0.033)
B. Binary, cross-sectional (vs. same-quarter median of other municipalities)		
HT arrival in HWP	−0.080***	−0.049**
	(0.015)	(0.019)
Difference (HWP − LWP)	−0.029	−0.036
	(0.026)	(0.027)
C. Continuous premium (standardized)		
HT arrival at mean premium ($z = 0$)	−0.073***	−0.041**
	(0.015)	(0.017)
Gradient per +1 SD of premium	−0.024	−0.036**
	(0.017)	(0.018)
D. Continuous cost of job loss (standardized)		
HT arrival at mean cost ($z = 0$)	−0.081***	−0.059***
	(0.017)	(0.020)
Gradient per +1 SD of cost	−0.029	−0.055**
	(0.025)	(0.025)

Notes: All moderators fixed at the arrival month; high-type classification as in Section 5. Panel A repeats the main table. Panel B: municipality above the same-quarter median of other municipalities. Panel C: standardized premium. Panel D: outside wage scaled by the reemployment probability implied by the municipal unemployment rate (smaller sample, as the unemployment rate is missing for some municipality-quarters). Columns: log shrinkage value and log shrinkage-to-sales (remaining outcomes in the batteries above). Store, calendar-month, and year fixed effects; standard errors clustered at the store. $^{***}p<0.01$, $^{**}p<0.05$, $^{*}p<0.1$.

H.2 The Continuous Premium Gradient

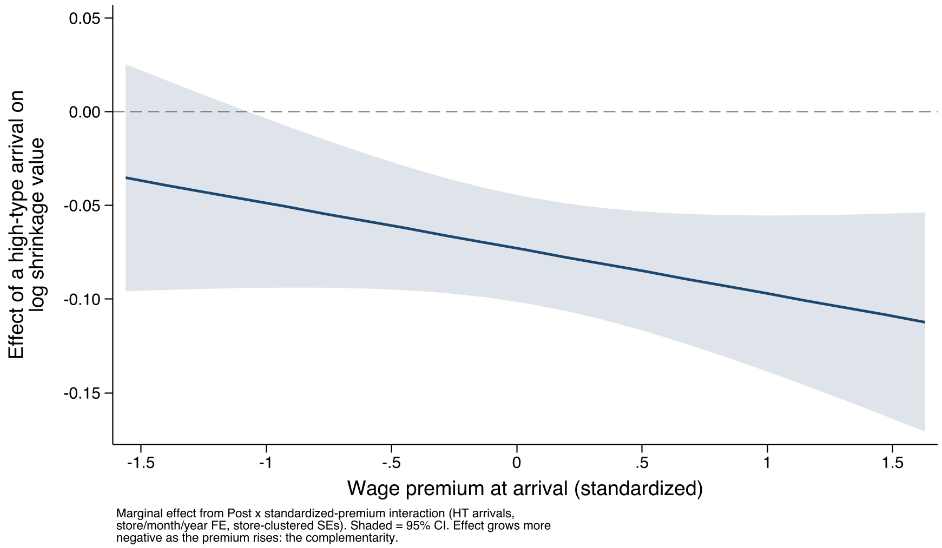


Figure 20: The high-type arrival effect on shrinkage grows more negative as the wage premium rises

Notes: Marginal effect of a high-type arrival ($\text{Post}_{[0,8]}$) on log shrinkage value as a function of the standardized wage premium of the store's municipality at arrival, from the interaction of $\text{Post}_{[0,8]}$ with the continuous premium (store, calendar-month, and year fixed effects; standard errors clustered at the store). Shaded band is the 95 percent confidence interval; the plotted range spans the 5th–95th percentiles of the arrival premium.

H.3 The Pooled Triple-Difference

Pooling the first arrival of either type into a single regression, I re-estimate the difference-in-differences companion of Equation (21), adding to the post indicator its interactions with the manager type and the wage environment and the triple interaction Post $\times$ HT $\times$ HWP, whose coefficient is the triple difference in the top row of Table 42. The design reproduces the per-type message cell by cell on log shrinkage value: the only cell in which arrivals reduce it is high-type-in-high-premium (-8.5 percent); low-type arrivals in low-premium stores raise shrinkage on the ratio measure; and the remaining value cells are indistinguishable from zero (on the ratio, the high-type-in-low-premium cell is a small positive). The triple-difference coefficient carries the predicted negative sign but is imprecise, because high-premium arrivals are scarce and the pivotal cells rest on few stores; the powered test therefore remains the per-type interactions of Table 11, which use every arrival.

Table 42: Pooled first arrivals: manager type × wage environment (DDD)

	(1) Log shrink value Log turnover p.e.	(2) Log shrink/ sales
Arrival × HT × HWP (DDD)	−0.060	0.007
	(0.065)	(0.056)
Cell effects (linear combinations):		
HT arrival in HWP	−0.085**	−0.016
	(0.035)	(0.032)
HT arrival in LWP	0.005	0.043**
	(0.017)	(0.019)
LT arrival in HWP	0.006	0.012
	(0.053)	(0.040)
LT arrival in LWP	0.036	0.077***
	(0.023)	(0.022)

Pooled first arrival of either type; HT = type of the first arriving spell; HWP = wage environment at arrival. Top row: the coefficient of the triple interaction Post × HT × HWP (see text); lower rows: the total effect in each cell (linear combinations). Store, calendar-month, and year fixed effects; standard errors clustered at the store. Columns: log shrinkage value and log shrinkage-to-sales (remaining outcomes in the batteries above). $^{***}p<0.01$, $^{**}p<0.05$, $^{*}p<0.1$.

I Classification Circularity and the Complementarity

Table 43 summarizes the full battery, including the split-sample key (labels from odd months, tests on even months: −6.8 points on shrinkage value, −8.4 per worker) and a reliability-shrunken reclassification stress that preserves 88 to 93 percent of the differential's magnitude at the same significance.

Table 43: The wage-environment differential without the shrinkage classification: summary

Test	What it removes	Differential (HWP − LWP)
Classification keyed to non-justified absence	Any use of shrinkage in the label	-0.060^{**} (0.024)
Placebo pipeline (1,000 within-store permutations)	Everything except the mechanics	Mechanical mean -0.024; observed outside the full placebo band; net -0.062
Split-sample key (labeled on odd months, tested on even)	Shared residuals between label and test	-0.068^{*} (value); -0.084^{**} (per worker)
Reliability-shrunken reclassification	Noise in the label	88–93 percent of the magnitude, same significance

Each row removes one ingredient of the classification pipeline and re-estimates the HWP − LWP differential. Differentials in log points; standard error in parentheses where applicable. Full detail in the text of this appendix and Appendix J. $^{***}p<0.01$, $^{**}p<0.05$, $^{*}p<0.1$.

Because a high-type manager is classified from her AKM fixed effect in shrinkage, estimated on the full panel, the months that follow her arrivals contribute to the very label the arrival design then tests. Section 5 draws the resulting conclusion for the *level* of the arrival effect: it is a validation of the decomposition, not an independent measurement. This appendix asks the question that matters for the paper's central result: does the *complementarity*—the difference in the arrival effect between high- and low-premium environments—inherit that circularity? The answer, established three independent ways below, is that the mechanical component is real, does *not* fully cancel in the difference, and still cannot account for the complementarity: a null simulation bounds what mechanics can produce, and the observed differential lies outside the bound; a classification free of shrinkage data reproduces it.

First, the logic—stated as the condition it actually is. The mechanical component would cancel in the HWP−LWP difference only if it were homogeneous across environments, and its size depends on three measurable ingredients: the weight of the focal store's months in the classifying fixed effect, the residual volatility of shrinkage at the arrival store, and how close the arriving manager's fixed effect sits to her classification median (marginal labels are the ones noise can flip). Comparing the three across environments among the arrival events, the first two are balanced (focal weight 0.63 in both, $p = 0.999$; pre-arrival residual dispersion $p = 0.42$), but the third is not: managers arriving in high-premium stores sit closer to their cell median (0.10 versus 0.24, $p < 0.01$). Homogeneity therefore fails, the mechanical component need *not* cancel in the difference, and I do not rest the argument on cancellation. Instead I measure the mechanical component directly.

The measurement is a null simulation that preserves the mechanics and destroys the economics. Within each store, I permute the shrinkage residuals (net of store and time effects) across months: any true manager effect is scrambled away, while the circularity channel—the same realizations entering both the classifying fixed effect and the tested post-window—remains fully intact, along with each store's residual variance, each manager's spell structure, and each store's wage environment. Re-running the entire pipeline (estimation, classification, arrival dating, and the interacted difference-in-differences) one thousand times yields the distribution of what pure mechanics can produce. For the *level*, the answer is: a great deal—the placebo arrival effect averages -3.0 percent (central 98 percent band $[-4.6, -1.3]$), roughly two-thirds of the in-pipeline estimate, which is why the paper treats the level as a validation.[12] For the *differential*, mechanics produce -2.4 points on average—the non-cancellation the imbalance above predicts—with a central 98 percent band of $[-6.6, +1.5]$ and nothing in one thousand draws beyond -7.9. The observed differential of -8.7 lies outside the entire support of the null (permutation $p < 0.001$)[13], and netting the mean mechanical component from it leaves roughly -6.2 points of genuine complementarity—almost exactly the -6.0 points the circularity-free classification below delivers independently. Two corrections that share no machinery converge on the same number.

Second, the test. I re-run the arrival design classifying managers by their fixed effect in a *different* outcome—non-justified absence—estimated on the same panel with the same format-by-border median rule. No shrinkage data enter this classification, so no mechanical channel connects the label to the tested outcome. If the complementarity were an artifact of shrinkage circularity, it would vanish under this label. Table 44 reports the result: the arrival *level* on shrinkage is small and insignificant, as expected once the mechanical component is removed from a validation object; the *differential* survives—under the absence label, a high-type arrival leaves shrinkage unchanged in low-premium stores and cuts it by 6.1 percent in high-premium stores, a difference of 6.0 percentage points, significant at the five percent level. A classification with no mechanical link to shrinkage reproduces the paper's central pattern: the manager's arrival moves the shrinkage margin only where the premium gives her threat something to confiscate. Two companion labels bound the reading honestly: classifying by the officer-recorded absence measure yields the same sign (-4.0 points, not significant), and classifying by turnover yields no differential; the absence label is, ex ante,

[12] The simulated number is not a black box: Appendix J derives it in closed form—a truncated-mean formula evaluated at the panel's residual dispersion and average managerial tenure reproduces about eighty-five percent of it—and shows that the same formula accounts for every other pattern in this appendix.

[13] The permutation battery runs under its own pre-specified replication protocol (event window and transition rule fixed in advance), whose baseline differential is -8.7 points; the canonical specification of Table 11—the one Table 44 reports—puts it at -8.4. The mechanical mean and the placebo band are properties of the battery protocol; the cross-outcome key, re-estimated in Table 44, speaks directly to the canonical specification.

the natural cross-outcome proxy—it is the fixed effect most correlated with the shrinkage effect (Figure 4)—so I read the battery as supportive with the dispersion disclosed.

Table 44: The complementarity under a classification with no mechanical link to shrinkage

	Classification of the arriving manager by:	
Log shrinkage value	Shrinkage FE (baseline)	NJ-absence FE (cross-outcome)
Arrival effect, LWP stores	-0.043^{**}	-0.001
Arrival effect, HWP stores	-0.126^{***}	-0.061
Difference (HWP − LWP)	-0.084^{**} (0.037)	-0.060^{**} (0.024)
Pooled arrival level	-0.068^{***} (0.015)	-0.020 (0.014)

Arrival difference-in-differences of Section 5 ($\text{Post}_{[0,8]}$; store, calendar-month, and year fixed effects; ever-treated stores; standard errors, in parentheses, clustered at the store), with the high-type label built from the indicated fixed effect under the format-by-border median rule, and the wage environment fixed at arrival. The baseline column reports the canonical specification of Table 11; its level is precisely estimated but partially embeds the classification (see text). Under the cross-outcome label no shrinkage data enter the classification. $^{***}p<0.01$, $^{**}p<0.05$, $^{*}p<0.1$.

Third, the flip side, reported for completeness because it disciplines how the paper words the level. Two leave-out exercises measure how much of the arrival *level* carries out-of-sample signal: re-estimating each manager's fixed effect excluding the focal store entirely—the leave-out construction that the teacher value-added and judge-leniency literatures prescribe to purge own-observation bias (Chetty et al., 2014; Goldsmith-Pinkham et al., 2025)—and re-running the design, and, separately, fixing the baseline event set and asking whether the arriving manager's *leave-out* quality predicts the size of the drop. In both, the level does not survive: the leave-out arrival effect on shrinkage is a precisely estimated zero, and the out-of-sample quality gradient within the fixed event set is zero, while the transition margins that carry no mechanical link—turnover and the disciplinary instruments—retain their baseline behavior. The pieces reconcile: the null simulation implies a residual non-mechanical level component of one to two percent, small enough to sit inside the leave-out exercises' confidence bands—present, but beyond their power to detect.[14] This is precisely why the paper presents the arrival level as a validation of the decomposition and rests its causal weight elsewhere: on the differential just shown to survive a circularity-free classification, on the mechanism outcomes of Section 5, which are not derived from the shrinkage fixed effect, and on the enforcement margins.

[14] Both exercises use the arrival pipeline of Section 5 with format-median classification cells; results and code are available in the replication materials. The leave-out classifier is necessarily noisier—a manager's out-of-store fixed effect rests on her other spells only—so these estimates bound the portable component from below rather than measure it.

J The Arithmetic of the Mechanical Component

The simulated magnitudes of Appendix I are not artifacts of the simulation: each one follows from a single closed-form object—the mean of the lower half of a noisy average—evaluated at two numbers measured in the panel. This appendix derives them.

The label is a noisy average. Consider the null world of the simulation: no manager has any true effect. The classifying fixed effect of manager m is then pure estimation noise, and to first order it is the average of the shrinkage residuals (net of store and time effects) over her observed months,

$$\hat{\theta}_m \;=\; \nu_m \;=\; \frac{1}{T_m}\sum_{(j,t)\,\in\, m} \tilde{e}_{jt}, \qquad \sigma_\nu \;=\; \frac{\sigma}{\sqrt{T_m}}, \tag{32}$$

where σ is the monthly residual dispersion and T_m her months observed. In the panel, $\sigma = 0.172$ and the average manager is observed $T_m \approx 15$ months, so $\sigma_\nu \approx 0.044$.

Selection: keeping the lower half. The high-type label keeps the managers below their cell median—the lower half of the distribution of ν. For a normal average, the mean of the lower half is

$$E\big[\nu_m \,\big|\, \nu_m < \text{median}\big] \;=\; -\sqrt{\tfrac{2}{\pi}}\,\sigma_\nu \;=\; -0.798 \times 0.044 \;\approx\; -3.5 \text{ percent}, \tag{33}$$

the same truncated-mean object that underlies selection corrections and the shrinkage adjustments of the value-added literature (Chetty et al., 2014).

Why the tested window inherits the full draw. The post-arrival months the design tests are a *subset* of the months that formed the label. For exchangeable residuals, the mean of any subset satisfies

$$E\big[\bar{e}_{\text{tested}} \,\big|\, \nu_m\big] \;=\; \nu_m \qquad \text{(slope exactly one)}: \tag{34}$$

conditional on the manager's overall average being low by ν_m, every one of her months—including the tested ones—is expected low by the same ν_m. The tested window therefore inherits the selected noise draw one for one, and the first-order mechanical arrival effect is the full -3.5 percent of Equation (33). The simulation delivers -3.0: the gap is accounted for by three design details the formula ignores—the store fixed effect absorbs the manager's own months beyond the $[0, 8]$ window, netting out part of her draw; the fixed effect is not exactly the raw average because of the network partialling; and the medians are estimated within cells. A two-parameter formula reproduces about eighty-five percent of the simulated mechanical level.

One formula, every result. Equations (32)–(34) also explain the rest of the battery. Under leave-out or cross-outcome classification the label is built from *other* months or *another* outcome, so the tested window no longer shares the selected draw ($\mathrm{Cov}(\bar{e}_{\text{tested}}, \nu) \approx 0$) and the mechanical component vanishes—which is why the arrival level collapses there, and why the cross-outcome pattern is *diagonal*: each label inherits only its own noise.

Why the differential need not cancel. By environment, the mechanical component is $-\sqrt{2/\pi}\,\sigma_e/\sqrt{T_e}$, plus terms of the same form and opposite sign from the predecessor's label. The mechanical differential is therefore

$$\Delta_{\text{mech}} \;=\; -\sqrt{\tfrac{2}{\pi}}\left[\frac{\sigma_H}{\sqrt{T_H}} - \frac{\sigma_L}{\sqrt{T_L}}\right] \;+\; \text{predecessor and weighting terms,} \tag{35}$$

which cancels only if $\sigma/\sqrt{T}$ is identical across environments. It is not: the measured residual dispersion is about eleven percent higher at high-premium arrival stores (0.157 versus 0.141), and the remaining terms depend on career lengths and event counts—which is why Appendix I *measures* the differential null by simulation (-2.4 points on average) rather than asserting cancellation.

Why the null has a ceiling. The mechanical differential is an average of selection terms over roughly two hundred arrival events, so by the central limit theorem it concentrates: its simulated standard deviation is 1.7 points. The observed differential of -8.7 sits $(8.7 - 2.4)/1.7 \approx 3.8$ standard deviations from the center of the null—a normal probability of order 10^{-4}, consistent with zero occurrences in one thousand draws. Mechanics cannot reach the observed magnitude for the same reason a fair coin does not deliver ninety heads in a hundred tosses: not because it is impossible, but because an average of many draws does not travel that far.

Summary. Every piece of the pattern—the mechanical level of three percent, its collapse under leave-out and cross-outcome classification, the diagonal, the non-cancellation in the difference, and the ceiling of the null—is one formula, $E[\nu \mid \text{lower half}] = -\sqrt{2/\pi}\,\sigma/\sqrt{T}$, applied to a design that tests the same months it labeled.

K Manager Type Does Not Moderate the Compression

Section 9 moderates the compression by the manager's floor presence and reports that her estimated *type* does not moderate it. This appendix gives that result in full, in the two ways one can measure type. The first is the discrete high/low-type label the paper uses throughout—the exact parallel of the presence split—which delivers a clean null. The second is the AKM shrinkage fixed effect of Section 4 used continuously, where a right-signed gradient appears but fails its own pre-trend test and so cannot be read causally. Both reach the same conclusion: what moderates the squeeze is what the manager does, not the type she is estimated to be.

The discrete type split. I split the border stores exactly as the presence design does, but on manager type: high- versus low-type by the AKM shrinkage fixed effect, classified within format-by-border cells—the same high/low split the manager sections use—and fixed at the general manager in place the December before each hike. The two-phase specification is otherwise identical. Figure 21 reports it. The compression lands on both type groups—a late rise of 5.0 percent for low-type stores ($p = 0.02$) and 7.6 percent for high-type ($p < 0.01$)—and the high-minus-low difference is +2.6 points (s.e. 3.6, $p = 0.47$), small and, if anything, the wrong sign for a protective role. The split is clean: on the net-of-sales outcome the pre-hike coefficients are jointly flat in both groups (low-type $p = 0.44$, high-type $p = 0.65$), the same check the presence groups pass. Manager type, measured the way the rest of the paper measures it, simply does not blunt the compression.

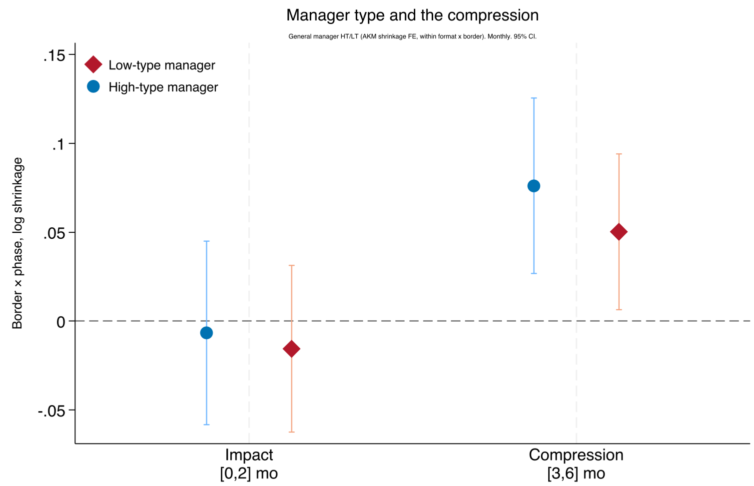


Figure 21: The compression response by the manager's estimated type at the hike. Border stores split at the median AKM shrinkage fixed effect within format-by-border cells (the paper's high/low-type label), the general manager fixed in the December before each hike. 95 percent confidence intervals.

The continuous fixed effect. I take the monthly stacked border design of Section 7 unchanged—store-by-hike and format-by-calendar-month fixed effects, the November–December reference, and the impact $[0, 2]$ and compression $[3, 6]$ phases—and add the continuous moderator. Let Q_k be the standardized quality of store k's general manager in the December before the hike, defined as the negative of her AKM shrinkage fixed effect, so that a higher Q marks a better manager, one whose stores lose less inventory.[15] The triple interaction Border $\times$ Compression $\times$ Q measures the gradient of the late-phase border effect in manager quality; the monitoring story predicts a negative sign—a better manager should blunt the compression damage.

Result. At face value the gradient obliges: it is negative and statistically strong for every shrinkage measure (-0.096, s.e. 0.025, for shrinkage net of sales; similar for value and units), the sales placebo is flat, and a tercile version shows the late border rise concentrated under the worst-ranked managers. But the design flunks the check that gives such gradients their meaning: interacting the *pre-hike* event-time indicators with Border $\times$ Q yields jointly significant coefficients (net-of-sales $p = 0.004$, value $p = 0.001$)—border stores

[15]The fixed effect is the per-manager effect from the log-shrinkage-value decomposition of Section 4; a median split on it recovers the high/low-type label used in the manager sections up to the border dimension, and managers labeled high-type carry the lower shrinkage effect, confirming the sign. The manager is the general manager in place in the December before each hike, predetermined with respect to the January shock; Q is defined for more border store-episodes than the binary split retains.

run by managers with different fixed effects were already on systematically different paths before the hikes arrived. The contamination is worst in the 2024 episode, whose distant pre-window is problematic even unconditionally (joint $p < 0.001$), and remains marginal in the clean 2023 episode ($p = 0.083$). A moderation that rides on differential pre-paths cannot be read causally: with the fixed effect estimated from the same shrinkage histories that generate those paths, the apparent buffering is indistinguishable from mean reversion in the label itself. This failure is the border contamination of the AKM person effect that motivates the paper's within-format-by-border classification (Section 4), surfacing here as a failed identification check rather than a usable result.

Reading. The contrast with the presence moderator is the point. The presence split is predetermined, measured from time-clock data rather than from outcome histories, and passes the same pre-trend test in both groups (net-of-sales joint $p = 0.41$ and $p = 0.36$ for low- and high-presence stores). The manager's estimated *quality* cannot be separated from the outcome paths it was estimated on; what she *does*—her presence on the floor, characterized in Section 5—can, and it is the moderator the paper rests on.

L What Store Shrinkage Lines Up With

Figure 22 shows how store shrinkage covaries with three families of standardized store characteristics—its workforce and organization, its operations, and its local environment—each entered one at a time, so the associations are directly comparable.[16] The personnel and organizational measures show the largest, most systematic associations (standardized associations of about 0.17 to 0.23). Customer traffic enters weakly (about 0.10), and the environmental measures—commercial robbery, homicide, and perceived insecurity—cluster near zero (all below 0.03 in absolute value).

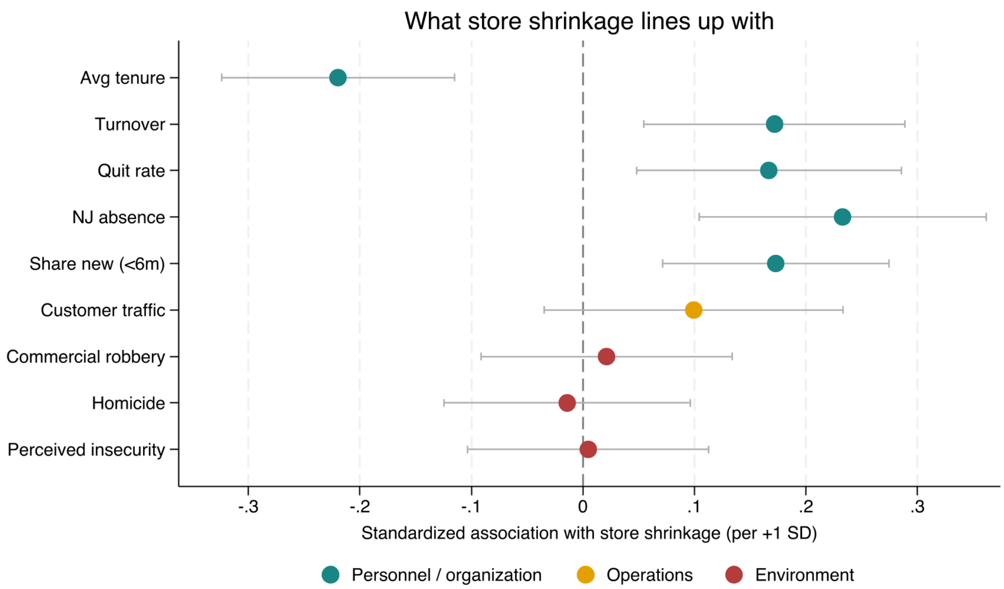


Figure 22: What store shrinkage lines up with. Standardized association of average store shrinkage with workforce, operations, and environment measures (store-level, more than 300 stores); horizontal bars are 95% confidence intervals.

[16]Each marker is the standardized coefficient (with 95% confidence interval) from a store-level regression of average shrinkage-to-sales on the indicated covariate, one at a time, across the more than three hundred stores; the personnel gradients survive conditioning on store format and municipality. Environment measures are municipal averages of SESNSP commercial robbery and homicide and of ENSU perceived insecurity.

References

Abowd, John M., Francis Kramarz, and David N. Margolis, "High Wage Workers and High Wage Firms," *Econometrica*, 1999, *67* (2), 251–333.

Adhvaryu, Achyuta, Parker Howell, Andrea Neyra-Nazarrett, Anant Nyshadham, and Jorge Tamayo, "What Do Good Managers Do and What Do They Leave Behind?," 2026. Working paper, Harvard Business School.

Akerlof, George A., "Labor Contracts as Partial Gift Exchange," *The Quarterly Journal of Economics*, 1982, *97* (4), 543–569.

Alvarez, Jorge, Felipe Benguria, Niklas Engbom, and Christian Moser, "Firms and the Decline in Earnings Inequality in Brazil," *American Economic Journal: Macroeconomics*, 2018, *10* (1), 149–189.

Andrews, Martyn J., Leonard Gill, Thorsten Schank, and Richard Upward, "High Wage Workers and Low Wage Firms: Negative Assortative Matching or Limited Mobility Bias?," *Journal of the Royal Statistical Society: Series A*, 2008, *171* (3), 673–697.

———, ———, ———, and ———, "High Wage Workers Match with High Wage Firms: Clear Evidence of the Effects of Limited Mobility Bias," *Economics Letters*, 2012, *117* (3), 824–827.

Autor, David H., Alan Manning, and Christopher L. Smith, "The Contribution of the Minimum Wage to US Wage Inequality over Three Decades: A Reassessment," *American Economic Journal: Applied Economics*, 2016, *8* (1), 58–99.

Bandiera, Oriana, Iwan Barankay, and Imran Rasul, "Incentives for Managers and Inequality Among Workers: Evidence from a Firm-Level Experiment," *Quarterly Journal of Economics*, 2007, *122* (2), 729–773.

Bertrand, Marianne and Antoinette Schoar, "Managing with Style: The Effect of Managers on Firm Policies," *Quarterly Journal of Economics*, 2003, *118* (4), 1169–1208.

Borusyak, Kirill, Xavier Jaravel, and Jann Spiess, "Revisiting Event-Study Designs: Robust and Efficient Estimation," *Review of Economic Studies*, 2024, *91* (6), 3253–3285.

Breza, Emily, Supreet Kaur, and Yogita Shamdasani, "The Morale Effects of Pay Inequality," *Quarterly Journal of Economics*, 2018, *133* (2), 611–663.

Cappelli, Peter and Keith Chauvin, "An Interplant Test of the Efficiency Wage Hypothesis," *Quarterly Journal of Economics*, 1991, *106* (3), 769–787.

Card, David, Alexandre Mas, Enrico Moretti, and Emmanuel Saez, "Inequality at Work: The Effect of Peer Salaries on Job Satisfaction," *American Economic Review*, 2012, *102* (6), 2981–3003.

———, Jörg Heining, and Patrick Kline, "Workplace Heterogeneity and the Rise of West German Wage Inequality," *Quarterly Journal of Economics*, 2013, *128* (3), 967–1015.

Cengiz, Doruk, Arindrajit Dube, Attila Lindner, and Ben Zipperer, "The Effect of Minimum Wages on Low-Wage Jobs," *Quarterly Journal of Economics*, 2019, *134* (3), 1405–1454.

Chetty, Raj, John N. Friedman, and Jonah E. Rockoff, "Measuring the Impacts of Teachers I: Evaluating Bias in Teacher Value-Added Estimates," *American Economic Review*, 2014, *104* (9), 2593–2632.

Coviello, Decio, Erika Deserranno, and Nicola Persico, "Minimum Wage and Individual Worker Productivity: Evidence from a Large US Retailer," *Journal of Political Economy*, 2022, *130* (9), 2315–2360.

Damodaran, Aswath, "Operating and Net Margins by Industry Sector (United States)," NYU Stern School of Business, data as of January 2026 2026. Net margin, Retail (Grocery and Food): 1.32 percent.

Dube, Arindrajit, T. William Lester, and Michael Reich, "Minimum Wage Effects Across State Borders: Estimates Using Contiguous Counties," *Review of Economics and Statistics*, 2010, *92* (4), 945–964.

Dustmann, Christian, Attila Lindner, Uta Schönberg, Matthias Umkehrer, and Philipp vom Berge, "Reallocation Effects of the Minimum Wage," *Quarterly Journal of Economics*, 2022, *137* (1), 267–328.

Emanuel, Natalia and Emma Harrington, "The Payoffs of Higher Pay: Labor Supply and Productivity Responses to a Voluntary Firm Minimum Wage," Staff Report, Federal Reserve Bank of New York 2024.

Goldsmith-Pinkham, Paul, Peter Hull, and Michal Kolesár, "Leniency Designs: An Operator's Manual," NBER Working Paper 34473, National Bureau of Economic Research 2025.

Goodman-Bacon, Andrew, "Difference-in-Differences with Variation in Treatment Timing," *Journal of Econometrics*, 2021, *225* (2), 254–277.

Harrison, David A. and Katherine J. Klein, "What's the Difference? Diversity Constructs as Separation, Variety, or Disparity in Organizations," *Academy of Management Review*, 2007, *32* (4), 1199–1228.

Hoffman, Mitchell and Steven Tadelis, "People Management Skills, Employee Attrition, and Manager Rewards: An Empirical Analysis," *Journal of Political Economy*, 2021, *129* (1), 243–285.

Kline, Patrick, Raffaele Saggio, and Mikkel Sølvsten, "Leave-Out Estimation of Variance Components," *Econometrica*, 2020, *88* (5), 1859–1898.

Lazear, Edward P., Kathryn L. Shaw, and Christopher T. Stanton, "The Value of Bosses," *Journal of Labor Economics*, 2015, *33* (4), 823–861.

Marx, Benjamin, Vincent Pons, and Tavneet Suri, "Diversity and Team Performance in a Kenyan Organization," *Journal of Public Economics*, 2021, *197*, 104332.

Mas, Alexandre and Enrico Moretti, "Peers at Work," *American Economic Review*, 2009, *99* (1), 112–145.

Minni, Virginia, "Making the Invisible Hand Visible: Managers and the Allocation of Workers to Jobs," *Quarterly Journal of Economics*, 2026, *141* (3), 1871–1920.

Nagin, Daniel S., James B. Rebitzer, Seth Sanders, and Lowell J. Taylor, "Monitoring, Motivation, and Management: The Determinants of Opportunistic Behavior in a Field Experiment," *American Economic Review*, 2002, *92* (4), 850–873.

National Retail Federation, "Shrink Accounted for Over $112 Billion in Industry Losses in 2022," National Retail Federation, National Retail Security Survey 2023.

Schor, Juliet B. and Samuel Bowles, "Employment Rents and the Incidence of Strikes," *The Review of Economics and Statistics*, 1987, *69* (4), 584–592.

Selyukh, Alina, "Retailers Howled About Theft Last Year. Why Not Now?," NPR, March 11 2024.

Shapiro, Carl and Joseph E. Stiglitz, "Equilibrium Unemployment as a Worker Discipline Device," *American Economic Review*, 1984, *74* (3), 433–444.

Tyco Retail Solutions and PlanetRetail RNG, "2018 Sensormatic Global Shrink Index," Technical Report, Tyco Retail Solutions 2018. Global shrink of 1.82 percent of sales, approximately $100 billion, across 14 countries.