Managing Organizational Complexity: Evidence from Retail in Latin America*

Achyuta Adhvaryu†, Parker Howell‡, Andrea Neyra-Nazarrett§, Anant Nyshadham¶, and Jorge Tamayo‖

March 21, 2025

Abstract

How does a manager's willingness and ability to engage with complex problems affect productivity? We study this question using administrative data from a large retail firm operating over 200 stores and employing more than 20,000 workers across Colombia. Leveraging both the inherent complexity of store management and a policy of rotating middle managers across stores, we examine the role managers play in driving productivity. Specifically, we ask: Who are the managers that increase sales, what do they do differently, and how do they manage complexity?

On average, stores carry 55,000 products and work with 800 suppliers—making it impossible for managers to oversee everything at once. We find that managers who successfully increase sales navigate this complexity through four distinct strategies that less effective managers fail to implement. First, they reduce stockouts. Second, they decrease the size of the inventory—both in terms of overall value and the inventory-to-sales ratio. Third, they simplify the store's portfolio by reducing the number of products and suppliers. Fourth, they execute more effective pricing decisions. We also observe that the arrival of a high-performing manager leads to changes in store organization, reallocating personnel toward sales, pricing, and inventory-related roles. Finally, we draw on a survey of managers at our partner firm to explore potential mechanisms underlying these findings.

Keywords: managerial quality, span of attention, retail, Colombia
JEL: J24, M10, M53

*We thank Juan Diego Forero, Joaquin Rueda, Daniel Franco Laverde, Andres Rengifo, and David Rodríguez-González for excellent research assistance. We thank Harvard Business School for supporting travel and other research-related expenses.

†University of California, San Diego, NBER, J-PAL, Good Business Lab; aadhvaryu@ucsd.edu
‡University of Michigan; parhow@umich.edu
§Harvard University; aneyra@g.harvard.edu
¶University of Michigan, NBER, J-PAL, Good Business Lab; nyshadha@umich.edu
‖Harvard University, Harvard Business School; Digital Reskilling Lab; jtamayo@hbs.edu

1 Introduction

A large body of research has established that managers account for a substantial share of variation in firm performance (Bloom and Van Reenen, 2007). In particular, managerial fixed effects have been shown to explain a significant portion of productivity differences across firms and establishments (Lazear et al., 2015; Metcalfe et al., 2023). Dessein and Santos (2021) further develop a theoretical model in which, under conditions of high complexity, differences in how managers allocate their attention amplify these fixed effects—suggesting that complexity can intensify the impact of individual managerial styles on performance.

Management involves a broad array of tasks, and managers must constantly make difficult decisions about how to allocate their scarce attention across competing priorities such as employee supervision, hiring, firing, promotions, pricing, product administration and general operations. However, managers face significant informational constraints and often do not know where their own marginal effort—or that of their subordinates—has the highest impact on outcomes. These span-of-control and cognitive hierarchy problems can lead to the misallocation of attention and effort, particularly in complex environments (Adhvaryu et al., 2022; Bandiera et al., 2014).

Importantly, these decisions unfold in environments where complexity is high and attentional bandwidth is limited. It is well-established that human decision-makers are bounded in their cognitive capacity and often rely on heuristics to manage complexity (Simon, 1973). In such environments, having more choices or more information does not necessarily lead to better decisions—in fact, large assortments can induce cognitive strain and reduce engagement (Iyengar and Lepper, 2000). This suggests that what sets effective managers apart may not only be what they know or what they do, but also how they allocate attention and simplify decision-making under constraints.

A growing literature has begun to unpack the specific types of skills that matter for effective management, moving beyond broad measures such as education or tenure. For instance, Weidmann et al. (2024) and Adhvaryu et al. (2023) highlight how interpersonal and cognitive skills are closely tied to managerial effectiveness and productivity outcomes. Complementing this, another set of recent papers opens the black box of *what* managers actually do to influence performance. Rather than focusing solely on traits, these studies document specific actions and behaviors—such as performance monitoring (Adhvaryu et al., 2024) and rapport-building with subordinates (Adhvaryu et al., 2025)—that distinguish high-performing managers from their peers.

Building on this literature, we examine how differences in managers' ability—and willingness—to navigate complexity translate into differences in firm performance. We

focus on a setting where complexity is not just present, but defining: large retail stores managing tens of thousands of products, hundreds of suppliers, and dozens of employees.

This complexity is especially pronounced in core operational areas such as supply chain coordination, inventory management, and pricing. In our setting, each store works with an average of 826 suppliers and carries approximately 54,875 products—making it impossible for managers to monitor every item individually. Yet they remain responsible for ensuring that products are consistently stocked, competitively priced, and aligned with firmwide goals. These conditions create a natural testbed for studying how managers cope with complexity: who succeeds, what decisions they prioritize, and how attention is allocated when it cannot be spread evenly.

To answer this question, we use rich administrative data from a large Colombian retailer. Our dataset includes scanner-level sales data, detailed records on inventory, supplier characteristics, and supplier-store transactions. Crucially, we exploit the firm's policy of randomly rotating managers across stores between January 2017 and June 2020. Leveraging this quasi-experimental variation, we first estimate an Abowd-Kramarz-Margolis (AKM) model using store sales to separately identify manager and store fixed effects. We then classify managers as high type or low type, with high type managers defined as those whose estimated fixed effects fall above the median of the distribution.

Next, we combine these estimates with personnel records to characterize what high type managers (good managers)look like in terms of observable behaviors and team composition. We complement this with a managerial survey, which captures information on managerial style and traits, allowing us to better understand which types of managers drive superior performance. Finally, we return to the internal data to examine how these high type managers operate—focusing in particular on inventory management, pricing, and other internal metrics that reflect how they handle complexity on the ground.

We find that effective management of complexity is a key differentiator in managerial performance. Good managers distinguish themselves not by increasing attention to all aspects of operations, but by strategically focusing on critical areas. Specifically, we identify four mechanisms through which high-performing managers navigate this complexity:

First, good managers reduce stockouts by approximately 5 percent, directly enhancing sales by ensuring product availability. Second, they simultaneously decrease inventory size, with reductions in both overall dollar value of inventory and the ratio of inventory to sales. The combination of fewer stockouts alongside leaner inventory suggests that good managers allocate their attention more efficiently to inventory management. Third, good managers simplify their store's product portfolio by reducing both the number of products and suppliers. Finally, they make more effective pricing decisions, as evidenced by higher sales despite fewer price changes.

Upon a good manager's arrival, we also observe a significant restructuring, with more personnel being allocated to sales, pricing, and inventory-related departments. Specifically, good managers reduce turnover among SPI (sales, pricing, and inventory) managers by nearly 40 percent while increasing hiring in these departments by more than 50 percent. This organizational adaptation suggests that good managers recognize the outsized impact of effective SPI management on overall store performance.

Our survey evidence, while highly suggestive, complements these empirical findings, revealing that good managers report spending more time on accounting and inventory-related procedures and are more likely to identify pricing as a critical determinant of store performance. For example, while not statistically significant, we observe that good managers are nearly three times as likely to list basic math skills as critical to their job compared to managers of lower-performing stores—suggesting that quantitative skills may facilitate better decision-making in complex environments.

The rest of the paper is organized as follows. Section 2 provides background on the institutional context, store operations, and the organizational structure of our partner firm. Section 3 describes the administrative and survey data used in our analysis. Section 4 presents our empirical strategy, beginning with the estimation of manager fixed effects and continuing with our main event-study design and mechanism analysis. In Section 5, we complement our administrative results with suggestive evidence from a managerial survey. Section 6 offers robustness checks and addresses potential concerns around identification. Section 7 concludes with a discussion of the broader implications of our findings for managerial effectiveness and organizational design in complex environments. An online appendix provides additional results and robustness exercises.

2 Context

Our partner firm is the Colombian branch of one of the largest supermarket chains in Latin America. We leverage administrative data from firm's 220 locations in Colombia from January 2017 and June 2020. The company has different stores format including a mini market, a hyper market and a super market format.

2.1 Store operations

Stores setup, layout and operational structure is highly standardized and centrally determined by Central Headquarters. Including the organizational structure of the company and its management practices. The organizational structure has three layers comprising of central headquarters (where the main strategic direction, financial decisions,

production processes, and HR policies are determined), middle management working in the stores (these managers supervise production activities and manage the day-to-day operations of each store), and front-line workers (who oversee production activities).

Importantly, each store's data systems are integrated and standardized by our partner firm, ensuring that the metrics we rely on are standard across stores.

2.2 Employees and Sections of the Store

The typical store operated by our partner firm employs 103 workers, works with an average of 826 suppliers, and sells 54,875 distinct products. As in many large supermarkets (e.g., Walmart or Costco), the type of products sold varies, with any given store selling everything from non-perishable goods like flat-screen TVs and car batteries to perishable goods like sushi and donuts. While many of these (on average) 54,875 products are standard across stores, demand varies greatly store-to-store.

Employees are spread across several departments (e.g., Bakery, Butchery, Customer Service, Product Replenishment and Display, Cashier and Payment, Logistics and Storage, Security, Administrative Support, etc.). We refer to workers and managers whose main responsibilities are *sales*-, *pricing*-, or *inventory*-related as "SPI" employees. Any worker or manager whose job title is not primarily associated with a sales, pricing, or inventory task is a "non-SPI" employee (for example, stock control or warehouse workers, supply and sales coordinators, and price analysts are all SPI roles, while the store manager, cashiers, janitorial staff, call center consultants, customer management, or butchery/bakery roles would all be non-SPI roles).

This distinction is an important one: SPI workers are focused on the logistics of stocking a large store (e.g., receiving shipments in the warehouse, physically stocking shelves, negotiating with suppliers, setting prices, etc.), while non-SPI managers focus more on customer-facing tasks (e.g., ensuring speedy checkouts, managing customer service and returns, overseeing the deli, meats, and/or bakery, etc.) and personnel-specific tasks (e.g., scheduling, general training, and hiring). While the complexities of personnel management and customer-facing work are not negligible (Adhvaryu et al., 2025), we find it useful to differentiate between the types of problems inherent to these roles and the types of problems that come from managing the logistics of keeping 55,000 products well-priced and appropriately stocked.

Of the 103 workers in an average store, 25 are SPI workers and 77 are regular, non-SPI workers (see Table 1). Stores hire an average of 4.73 workers per month, with 4.34 hires going to non-SPI jobs, and only 0.40 hires working within SPI.[1]

[1]The low share of SPI hires appears to be offset by the fact that non-SPI workers are more likely to be pulled into an SPI role than vice versa.

Table 1: Pre-treatment summary statistics for all stores

	All stores	
	Mean	SD
Total workers (all)	102.94	101.37
Total workers (SPI)	25.33	27.26
Total workers (non-SPI)	77.62	75.61
Total managers (all)	8.01	7.13
Total managers (SPI)	6.09	5.48
Total managers (non-SPI)	1.92	2.67
Ratio managers to non-managers (all)	0.11	0.10
Ratio managers to non-managers (SPI)	0.50	0.71
Ratio managers to non-managers (non-SPI)	0.06	0.09
Turnover managers (all)	0.20	1.05
Turnover managers (SPI)	0.15	0.83
Turnover managers (non-SPI)	0.05	0.30
Total hiring (all)	4.73	9.88
Total hiring (SPI)	0.40	1.07
Total hiring (non-SPI)	3.87	6.56
Total hirings (managers)	0.01	0.12
Hiring managers (SPI)	0.01	0.09
Hiring managers (non-SPI)	0.00	0.08
Moves managers (all)	0.051	0.351
Moves managers to SPI	0.045	0.333
Moves managers to non-SPI	0.006	0.078
Arrival of any manager	0.14	0.35

Table 1 reports means and standard deviations for organizational variables for all stores prior to the arrival of a good manager. All statistics are reported at the store-monthly level for positions in all, SPI and non-SPI departments. *Workers* refer to any person employed in the retail store during the period analyzed. *Managers* are defined as those workers who are coordinators, leaders, bosses and directors. Conversely, *Non-managers* are those workers who are not in those positions. *Turnover* indicates the number of workers in the store that left the company entirely (i.e., it excludes moves between stores), while *Hiring* represents the number of workers in the store that were hired. Finally, *Moves* indicates the number of moves from one position to another within the company (including vertical or horizontal moves).

2.3 Managers

The lower panel of Table 1 presents statistics specific to the management team within stores. On average, stores employ 8 managers, 6.1 of whom are SPI managers. SPI departments are far more rectangular than non-SPI departments: the ratio managers to

non managers in SPI departments is 0.50, compared with 0.06 in non-SPI departments. Movement between SPI and non-SPI managerial roles is rare, with far more managers shifting from non-SPI roles to SPI roles (0.05 versus 0.01 moves per month). Movement across stores is a bit more common than within: on average, 0.16 managers arrive at a store per month.

2.4 What managers do

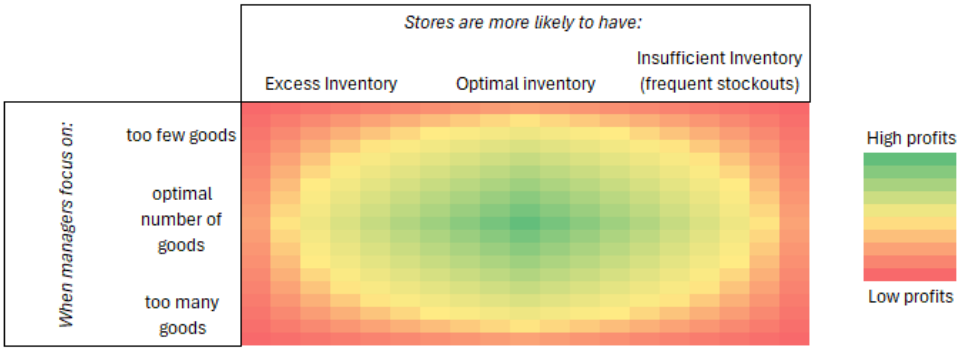
Managers also oversee personnel management, training, workflow coordination, and are responsible for maintaining quality standards and addressing operational issues. They typically lead teams, support front-line workers, and assign workers to tasks to ensure tasks are completed following centralized guidelines.

In addition to personnel-related tasks, for any given product a manager must decide *when*, *how much*, and *from which suppliers* to order. Demand for consumer goods varies by region (e.g., a winter coat will sell better in Medellín than Cali) meaning that each store needs to stock different products. However, even within a store, demand is sensitive to everything from weather, time of year, and day of the month.

To appreciate the complexity of this task, imagine a toy version of this problem in which a manager's only duty is to worry about maximizing profits for a single product (i.e., this manager doesn't worry about personnel issues like training or hiring). In this example, the manager's only job is to ensure that a single good is appropriately stocked (i.e., she wants to avoid excess inventory while avoiding stockouts) and priced (i.e., regular sales should generate enough margin to make the product worth selling, but not so much so that consumers seek out the good via other sales channels) so as to maximize store profits.

To fix ideas, suppose that product is a soccer jersey for Colombia's national soccer team. After looking through historical data to determine the right proportion of shirts to order in each size, our manager might predict that her store will experience increased demand before each game. If so, she may consider it a good use of time to look up the national team's schedule and place orders ahead of time to ensure her store has enough jerseys for each game, perhaps ordering a few more jerseys for games she considers more important than others. Once the jerseys arrive (not too far in advance so the store doesn't carry excess inventory, but also far enough in advance so that consumers can purchase the jerseys in time for the big game) she must consider how best to price the jerseys. The profit-maximizing price may be a function of both time and the team's success; It likely makes sense to charge a high price the day of the game, but it's unclear how to price the jerseys the day after the game: a loss may portend a price cut, while a win may require an increase.

Figure 1: Trade-offs in expected profits between attention and inventory management



In this example our manager is likely to do a good job maximizing profits for her one product (she may also end up spending a lot of time watching soccer). Of course, as the manager's attention becomes more divided, she will need to spend less time maximizing profits on soccer jerseys to ensure she maximizes profits on all the products she manages. Intuitively, there is a trade-off between extracting a higher margin from a lower number of products and extracting a lower margin from a higher number of products.

Thus, managers who are most likely to increase store profits are those who optimize along both dimensions illustrated in Figure 1. Specifically, allocating attention to any one product increases the chances that profits are maximized on that product, but simultaneously decreases the likelihood that the manager will maximize store-level profits. Conversely, spending too little attention on all products increases the chances that a store carries excess inventory or experiences excessive stockouts. As the number of products (and/or suppliers) increases, so too does the likelihood that the attention constraint binds.

3 Data and Descriptive Statistics

We use two data sources. The first and primary source is administrative data from the supermarket chain's 220 stores that operated across all of Colombia between January 2017 and June 2020. This data includes start and end dates, wages, promotions, and attendance logs for every employee. We also have sales, pricing, and inventory for every product as well as a record of all suppliers for any product sold by the firm. Importantly, each store's data systems are integrated and standardized by our partner firm, ensuring that the metrics we rely on are standard across stores.

The secondary source is survey data from a survey fielded after the period of analysis on an overlapping sample of managers[2].

[2] Weidmann et al. (2024) partnered with Good Business Lab to conduct this survey in 2024. While we did not design the survey instrument explicitly for our study, this survey captures a broad array of

3.1 Administrative Data

The most important indicator of store performance is sales. We also define productivity as a store's sales divided by its number of employees.[3] Two variables that are particularly important in the retail context are "stockouts" and value of inventory.

A *stockout* refers to the situation in which a store runs out of a given product due to excess demand.[4] Minimizing the frequency of stockouts is a key task for SPI managers, both because a stockout means that the store is actually losing out on sales every time a customer would have purchased that product[5], and because of potentially negative impacts on customer satisfaction. Prior to receiving a good manager, stores experience an average of 60 stockouts per month costing approximately 20,000 USD each month (see Table 2). On average, stores have

Stockouts can be avoided given a large enough inventory, but, of course, large inventories can have significant monetary-, time- and space-related costs. While it is unclear whether store managers fully internalize the cost of tying up money in inventory costs (i.e., it is unlikely that holding too many products in inventory means the store managers cannot afford to buy more products), it is very likely that managers experience time- and space-related costs of holding too much inventory.

For example, carrying excess inventory of perishable goods–or even non-perishable, seasonal products–may require stores to sell these products at lower prices to make room for newer products. Moreover, space constraints bind not only on customer-facing shelves, but also in the store's warehouse: indeed, as seasons change a manager may prefer to sell larger non-perishable products (e.g., synthetic Christmas trees) for a loss to avoid storing them for several months.

Finally, a key input in determining both the frequency of stockouts and the value of inventory is how managers price their products. Table 2 reports summary statistics for several variables relating to stockouts, inventory size/value, and product pricing. On average, stores sell \$1.72 million USD per month, while holding \$900 million USD in inventory. Price increases (5,622 per month) are far less common than price cuts (9,015 per month).

---

managerial practices, making it an excellent descriptive supplement to our empirical analysis.

[3] Our data does not include a reliable measure of the costs incurred when a store acquires a product from a supplier. Thus, we focus on sales instead of profits.

[4] For simplicity we categorize cases when a product sells out but is not re-purchased (i.e., the product is discontinued) as stockouts, even though this means we may slightly overcount unintentional stockouts. We do so due to data limitations–because we do not observe sales or inventory after the end of our sample we cannot say whether a manager has not yet had a chance to reorder the product or whether it was truly discontinued.

[5] Of course, some within-store substitution of products is likely to occur, but on average, stockouts certainly reduce a store's sales.

Table 2: Pre-treatment summary statistics for sales-, pricing- and inventory-related outcomes

	All stores	
	Mean	SD
Sales USD	1,723,302.38	1,529,728.75
Productivity USD	19,217.22	9,067.28
Num. out of stock products	886.13	674.99
Stock out value in millions	1,284.48	1,805.38
Num. of days of item out of stock	1,156.05	821.77
Stock value in millions	1,511.11	1,174.55
Stock value over sales	0.25	0.10
Number of products with discounts	8,874.56	9,559.26
Value of products with discount	102,866,384	136,174,592
Total price changes	14,637.87	13,149.14
Positive price changes	5,622.45	4,909.51
Negative price changes	9,015.41	8,918.52
Number of suppliers	826.16	335.13
Number of products	54,875.33	49,284.88
Num of bad sold products	8,453.18	9,575.90
Number of any price changes	101,820.60	112,375.80
Number of positive price changes	36,672.39	36,159.95
Number of negative price changes	101,818.55	112,377.60

Notes: Table 2 reports means and standard deviations for sales-, pricing-, and inventory-related variables for all stores prior to the arrival of a good manager. All statistics are reported at the store-month level. *Productivity* refers to the sales by worker (sales divided by number of workers). *Out of stock products* indicates the products with negative inventory levels in a given month, while *Stock value* shows the product value with positive inventory levels. *Number of days out of stock* refers to the total number of days a store was out of stock. *Price changes* indicates counts the number of times a product have a positive or negative price change between two months. *Bad sold products* are products with sales 0.5 standard deviations below the mean.

3.2 Survey data

In 2023, Good Business Lab, in conjunction with Weidmann et al. (2024) surveyed the universe of managers at our partner firm (achieving a 77.6% response rate), capturing information on managers' demographics, job histories, psychometric profiles, personality traits, managerial practices, leadership styles, job descriptions and job satisfaction. The survey also included measures for mathematical ability, recall, fluid intelligence, emotional intelligence, and performance on a task simulation. Since some managers from our 2017-

2020 sample were not the same as the managers in 2024, we report responses for the 171 individuals who were managers in both time periods.[6]

[6]We note that the survey was administered to only a subset of managers. Specifically, only the 500 highest-level managers out of approximately 3000 within-store managers were targeted by this survey. This difference in sampling criteria may have implications for the generalizability of the findings, which we address below when discussing survey results.

4 Empirical Analysis

We first establish the existence of manager fixed effects. We then briefly review various theoretical models on choice overload, including the recent model developed by Dessein and Santos (2021) and prior findings by Bandiera et al. (2014) at the executive level to help us develop predictions about what good management in complex settings should ideally look like. We observe that the arrival of a good manager does, in fact, increase store productivity and sales. Finally, we uncover the specific mechanisms managers used to achieve such increases and then consider how the organizational structure responds to the arrival of a good manager.

4.1 Measuring Value-Added of a Manager

We start with the, by now well-established, observation that individual managers may exert large effects on productivity (Lazear et al., 2015; Metcalfe et al., 2023) and classify managers according to their fixed effects on store performance. These fixed effects will be biased unless moves are uncorrelated with time-varying components of store productivity. Consequently, we are able to leverage the rotation of managers across stores only if these rotations are for reasons unrelated to performance. In our context, the rotation of managers across store branches is unrelated to store or manager performance (see Section 2), permitting us to estimate the differences in sales attributable to individual managers. We use the AKM framework (Abowd et al., 1999) to estimate the fixed effect on log productivity attributable to a given manager. Specifically, we estimate the following two-way fixed effects (TWFE) model:

$$\log(sales)_{ijt} = \theta_i + \psi_{j(i,t)} + \delta_t + \epsilon_{ijt} \tag{1}$$

where $log(sales)_{ijt}$, is total log sales in store j reporting to manager i, in month t; $\psi_{j(i,t)}$ is a store-fixed effect, δ_t is month fixed effect, and θ_i is a manager fixed effect.

Following Card et al. (2013), we further decompose the error term in (1) into a match-specific component $\eta_{j(i,t)}$, a store root component ξ_{it} and a transitory error ν_{ijt}:

$$\epsilon_{ijt} = \eta_{j(i,t)} + \xi_{it} + \nu_{ijt} \tag{2}$$

Formally, our identification of manager and store fixed effects relies on the assumption that the assignment of managers to stores is conditionally mean-independent of past, present, and future values of ϵ_{ijt}. This assumption excludes the possibility of managers being assigned to stores based on their match-specific component ($\eta_j(i,t)$) or transitory shocks to store performance (ν_{ijt}). Managerial sorting based on either of $\eta_{j(i,t)}$ or ν_{ijt}

would mean that this sorting is endogenous, resulting in biased and inconsistent estimates.

This assumption does not restrict the possibility that managers are assigned to stores based on permanent components of managerial ability θ_i or store components α_j. While our partner firm has told us that the assignment of mangers to stores is quasirandom (specifically, managers are shuffled across stores periodically as part of their career development), we note that even if sorting on these (permanent) fixed effects does occur (e.g., managers prefer to work in stores closest to where they live) does occur, our identification is still valid.

Finally, these fixed effects are separately identified only within stores which are linked by managers moving across stores (i.e., "connected sets"). We identify 6 connected sets and estimate Equation (1) within each connected set. Appendix Table A.17 shows the results of the AKM decomposition, suggesting that manager fixed effects explain more than 12 percent of log productivity.

In the rest of the paper we refer to a "good" manager as any manager whose fixed effect on store sales is above-median.

4.1.1 Tests for Endogenous Mobility

In addition to assurances from our partner firm, we perform a series of formal tests for endogenous mobility.

As a first test, we follow Card et al. (2013) to conduct an event study around manager moves between stores to assess the extent to which these moves might be systematically driven by productivity shocks or by sorting on the match-specific component of log-productivity. We do this by isolating movers and ranking them first by average productivity of the store they moved *from*, and second, average productivity of the store they moved *to*. We split the average productivity of stores into quartiles. In Figure A.15 we plot the average monthly (residual) efficiency of the moving manager on the y-axis. We display only moves away from either the top (quartile 1) or bottom (quartile 4) quartiles. The key insight to this test is that if match-specific components are, indeed, important determinants of when/where managers move, we should see managers systematically moving when their productivity is either particularly high, or particularly low (depending on whether managers sort positively or negatively on match-specific components). If this is the case, then we would expect managers to either gain or lose (on average) in terms of productivity from moving to a new store, regardless of the productivity of their store of origin. On the other hand, if moves are not driven by match-specific components, managers who move to higher productivity stores will become *more* productive, while managers moving to less productive stores will also become less productive.

Figure A.15 confirms this intuition. In general, managers moving to better stores gain

in terms of productivity, while managers moving to less productive stores become less productive. Moreover, top (bottom) quartile managers moving to top (bottom) quartiles stores experience minimal changes in productivity.

We further follow Card et al. (2013) by plotting all potential combinations of origin and destination stores (ranked by quartiles of average manager productivity). Figure A.16 plots the average change in residual log productivity for managers who move from stores in quartile i to quartile j against the change in residual log productivity for managers who move from stores in quartile j to i. Thus, if manager moves are indeed conditionally mean independent of the match-specific component, then any productivity changes from moving from store i to store j should be equal and opposite any productivity changes incurred when moving from store j to store i. Conditional mean dependence due to the match-specific component would show up as deviations from the 45 degree line. Reassuringly, the deviations in Figure A.16 appear to be small and non-systematic.

4.2 Definitions

The above decomposition enables us to classify managers and stores according to their fixed effects. Specifically, we refer to any manager who exerts an above-median fixed effect (within the connected set) on log productivity as a "good" manager. To classify a store as good or bad, we first take the mean of all the fixed effects of managers who work in that store. We then compare this mean to the median fixed effect of managers across the 220 stores. Thus, "Bad" stores are those in which the mean fixed effect of all managers in that store is below median across the 220 stores.

In the rest of the paper we define "treatment" as the arrival of a "good" manager to a "bad" store.

4.3 Baseline Specification

We estimate an event study model (Freyaldenhoven et al., 2021; Roth, 2022), testing for differential pre-trends across stores before the arrival of a manager, and measuring post-arrival impacts. We estimate the following baseline specification:

$$Y_{s,t} = \alpha_0 + \sum_{\underline{C} \leq k \leq \overline{C}, k \neq -1} D_{st}^{k} \delta_k + \Phi_s + \theta_t + \epsilon_{s,t} \tag{3}$$

where $Y_{s,t}$ is the performance measure of store s in month t, D_{st}^{k} is a relative time to treatment indicator for whether a good manager arrived in period $t-k$, defined as $D_{st}^{k} = 1[t = \tau_s + k]$ for $k \in (\underline{C}, \overline{C})$, $D_{st}^{\underline{C}} = 1[t \leq \tau_s + \underline{C}]$, and $D_{st}^{\overline{C}} = 1[t \geq \tau_s + \overline{C}]$, where $1[\cdot]$ is the indicator function, k indexes the set of time indicator variables, and τ_s is the first

monthly period in which the good manager works at store s. The parameters of interest are δ_k for $k \in [\underline{C}, \overline{C}]$, which measure the impact of a good manager before, during and after arrival. We normalize $\delta_{-1} = 0$ and set $\underline{C} = -11$ and $\overline{C} = 11$.

We also control for store fixed effects Φ_s and time fixed effects θ_t, and cluster standard errors at the store level for inference. Our main specification excludes the never-treated stores (133 out of 220).

4.4 Identification

Our identifying assumption is that the match of managers to stores is exogenous (see Section 4.1). Our partner firm has assured us that the rotation of managers across stores is mechanical and assignments are made without considering manager traits or store needs. As an additional check, we compare observables for stores that received a good manager earlier versus later. Table 3 shows no significant differences between early- and late-treated stores. This provides confirmation of the anecdotal evidence that the rotation of managers across stores is indeed quasirandom.

Table 3: Balance table for early- and late-treated stores

Variable	Early arrivals		Late arrivals		Difference
	Mean	S.D.	Mean	S.D.	P-value
Sales USD	1,549.53	1,487.32	1,413.58	1,694.90	0.74
Productivity USD	17,630.70	4,989.90	15,768.92	6,104.19	0.21
Num. out of stock products	974.07	916.42	624.88	536.36	0.03
Stock out value in millions	1,288.15	1,861.08	833.06	1,528.09	0.26
Num. of days of item out of stock	1,153.13	819.23	918.95	722.39	0.21
Stock value in millions	1,406.76	1,257.31	1,152.04	1,052.44	0.36
Stock value over sales	0.25	0.08	0.27	0.12	0.50
Num products with discounts	8,217.56	8,898.10	3,994.57	5,649.46	0.01
Value of products with discount (in thous.)	92,507.45	115,201.98	37,063.48	65,024.64	0.01
Total price changes	14,202.19	13,494.43	10,778.34	12,166.25	0.28
Positive price changes	5,517.16	4,783.12	3,216.01	3,799.88	0.03
Negative price changes	8,685.02	8,737.73	7,562.33	8,839.34	0.61
Number of suppliers	801.61	365.20	710.33	331.77	0.29
Number of products	51,283.07	52,749.88	45,197.66	48,346.64	0.62
Num of bad sold products	7,440.93	8,098.49	6,939.14	7,172.45	0.79
Total workers (SPI)	24.64	28.71	25.32	32.80	0.93
Total workers (non-SPI)	74.82	75.72	81.13	103.48	0.80
Total managers (SPI)	5.52	5.38	5.47	5.65	0.97
Total managers (non-SPI)	1.93	1.53	2.78	5.81	0.51
Ratio managers to non-managers (SPI)	0.47	0.57	0.36	0.26	0.31
Ratio managers to non-managers (non-SPI)	0.06	0.06	0.09	0.11	0.24
Turnover managers (SPI)	0.07	0.08	0.10	0.16	0.40
Turnover managers (non-SPI)	0.03	0.03	0.06	0.15	0.27
Hiring managers (SPI)	0.00	0.01	0.01	0.06	0.58
Hiring managers (non-SPI)	0.00	0.01	0.00	0.02	0.91
Moves managers to SPI	0.03	0.06	0.06	0.14	0.47
Moves managers to non-SPI	0.00	0.01	0.01	0.04	0.28
Stores	21		66		

Notes: shows the mean and standard deviation of the period before the early (before September 2019) and late (September 2019 onward) arrival of a good manager for the first time to a bad store. In the last column of the table, we show the p-value associated with the difference in the mean between both types of store.

4.5 Managerial Quality and Performance

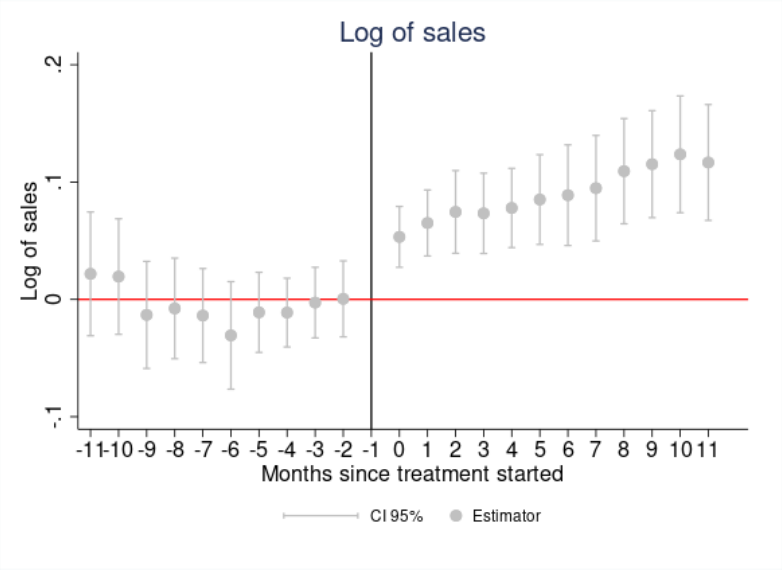
Given the large literature documenting that managerial practices drive performance (e.g., Syverson (2011)), we first confirm that this is true in our context. Table 4 shows that log sales and log productivity both increase by more than 10 percent after the arrival of a good manager. Figures 2 and A.18 show that these increases occur shortly after the good manager's arrival and increase over time.

Table 4: Log productivity AKM on log sale and log prod

VARIABLES	(1) Log of sales	(2) Log of productivity
Treatment effect	0.117***	0.112***
	(0.0218)	(0.0217)
Constant	22.06***	18.08***
	(0.0117)	(0.0177)
Observations	9,007	9,007
R-squared	0.984	0.818
Obs in post	1420	1420
Control group	nevert	nevert
Balanced Panel	nobal	nobal
stores	220	220
stores treated	87	87
Arrival controls	Yes	Yes
Sales control	No	No
Mean of Dv	21.73	17.91
Effect on mean	0.00540	0.00630

Notes: Table 4 shows the point estimates of log sales and log of productivity (measured as sales over number of workers per store). The sample consists of store-monthly panel data for the 220 stores during our analysis period of January 2017 to June 2020. The treatment effect is defined as the arrival of a good manager to a store with low performance (average manager fixed effects of the store above the median). Robust standard errors are in parentheses. * significant 10%, ** significant 5%, *** significant 1%.

Figure 2: Effect of Treatment on Log Sales



Notes: Figure 2 shows the event-study coefficients and 95 percent confidence intervals from estimating good managers' arrival on log sales. The sample consists of store-monthly panel data for the 220 stores during our analysis period of January 2017 to June 2020. The vertical line represents the arrival of the good manager.

4.6 What do good managers do?

In this section we ask what specific mechanisms are associated with the large increase in sales and productivity. We find evidence that good managers implement the following four changes shortly after their arrival to a store:

First, they reduce stockouts. Second, they decrease the size of the inventory. Specifically, both the overall dollar value of inventory and the ratio of inventory to sales decrease. Third, good managers tend to simplify their store's portfolio by reducing both the number of products and suppliers. Fourth, good managers make better (i.e., revenue-increasing[7]) price changes.

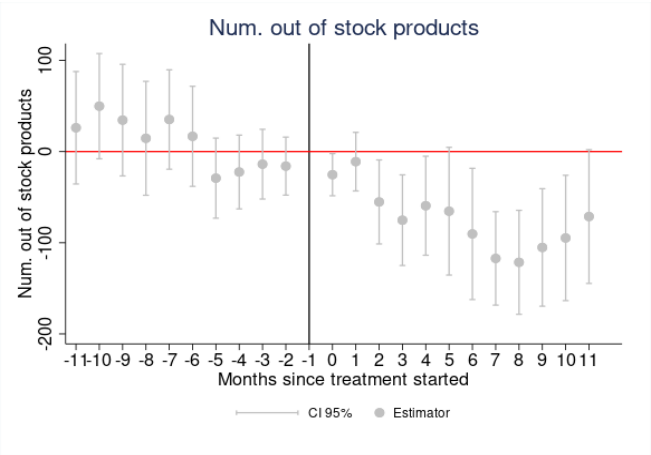
Prior to the arrival of a good manager, bad stores have on average 875 products which are out of stock (see Table A.8). Good managers reduce the number of out-of-stock products by 11.5% (5). Interestingly, they do this at the same time as making inventory actually leaner as measured by a reduction in overall stock value and stock value over sales (stock value is not going up relative to sales or even rising proportionately its actually dropping relative to sales despite sales going up and stockouts going down).

Table 5: Effects on on-hand inventory value and stockouts

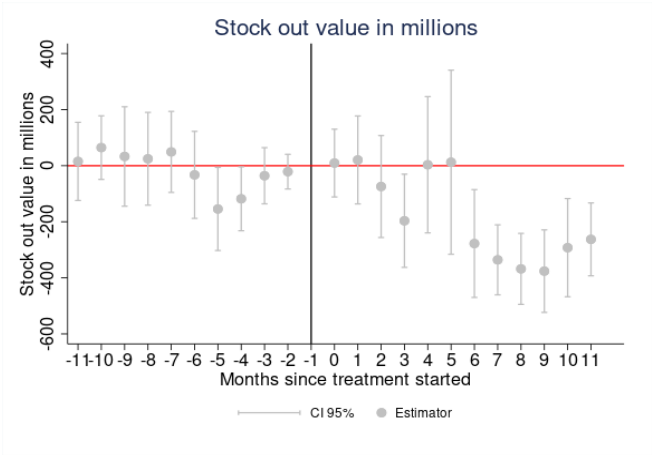
VARIABLES	(1) Num. out of stock products	(2) Stockout products proportion	(3) stockout_ part_onh	(4) Stock out value in millions	(5) Num. of days of item out of stock	(6) Stock value in millions	(7) Stock value over sales
Treatment effect	-86.97***	0.00228**	0.00343**	-165.4***	-41.23*	-93.46***	-0.0311***
	(31.55)	(0.00103)	(0.00159)	(48.25)	(21.63)	(27.38)	(0.00596)
Constant	810.0***	0.0381***	0.0460***	1,104***	1,064***	1,360***	0.238***
	(27.19)	(0.000586)	(0.000869)	(40.94)	(23.71)	(30.91)	(0.00380)
Observations	9,007	9,007	9,007	9,007	9,007	9,007	9,007
R-squared	0.889	0.964	0.939	0.885	0.945	0.985	0.805
Obs in post	1420	1420	1420	1420	1420	1420	1420
Control group	nevert	nevert	nevert	nevert	nevert	nevert	nevert
Balanced Panel	nobal	nobal	nobal	nobal	nobal	nobal	nobal
stores	220	220	220	220	220	220	220
stores treated	87	87	87	87	87	87	87
Arrival controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sales control	No	No	No	No	No	No	No
Mean of Dv	754.6	0.0376	0.0442	1024	1041	1265	0.250
Effect on mean	-0.115	0.0607	0.0776	-0.162	-0.0396	-0.0739	-0.124

Notes: Table 5 shows the point estimates of stockouts and on-hand inventory. A stockout is defined as a product whose number of items is equal to or below zero. The sample consists of store-monthly panel data for the 220 stores during our analysis period of January 2017 to June 2020. The treatment effect is defined as the arrival of a good manager to a store with low performance (average manager fixed effects of the store above the median). Robust standard errors are in parentheses. * significant 10%, ** significant 5%, *** significant 1%.

[7]We deal with revenue instead of profit due to the data limitations described in Section 3. Even so, in the retail context revenue is an excellent proxy for profits so long as most or all goods are sold for a weakly positive profit.



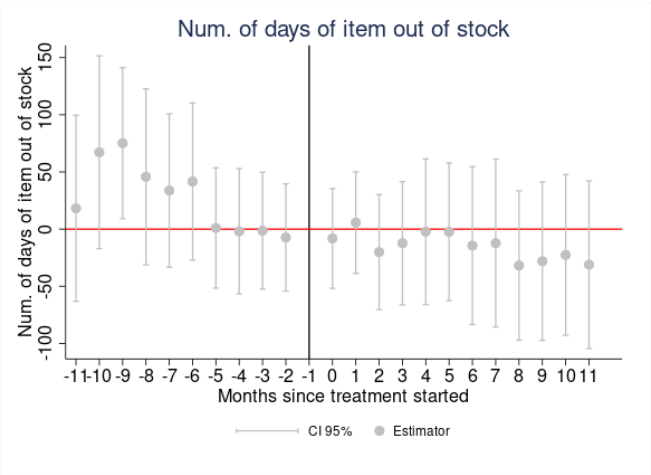
Number of products out of stock



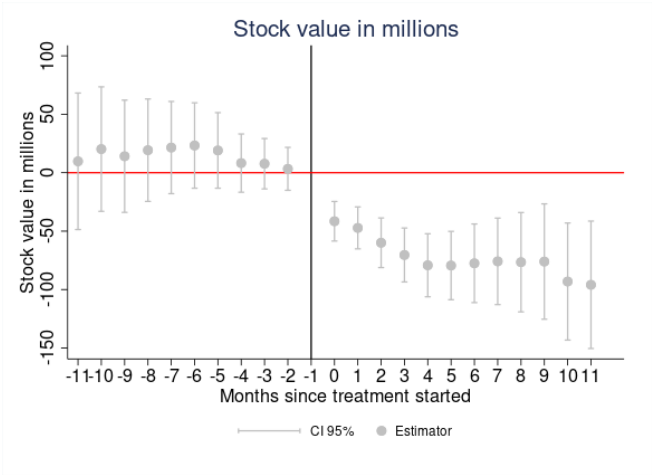
Stock out value in millions

Figure 3: Number and Value of Stockouts

Figure 4: Number of days out of stock and stock value over sales

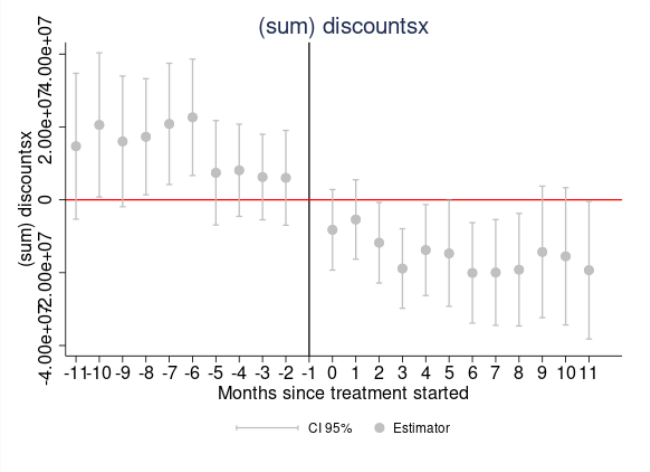


Number of days out of stock

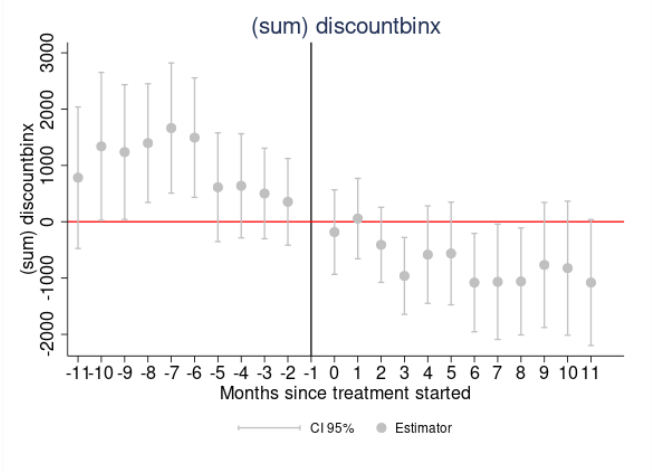


Stock value in millions

Figure 5: Total losses and number of products with losses

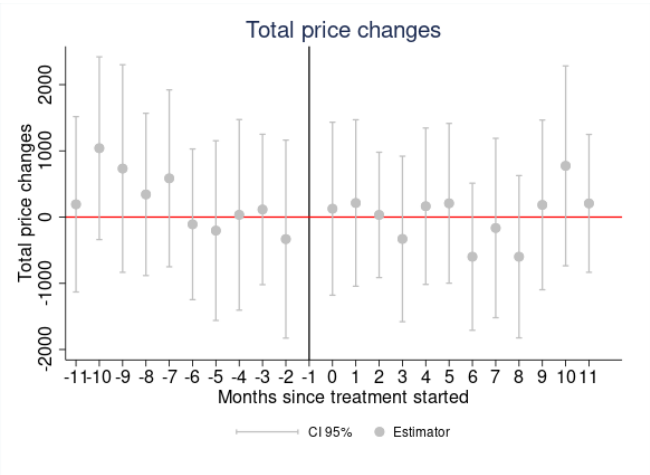


Total sum of losses

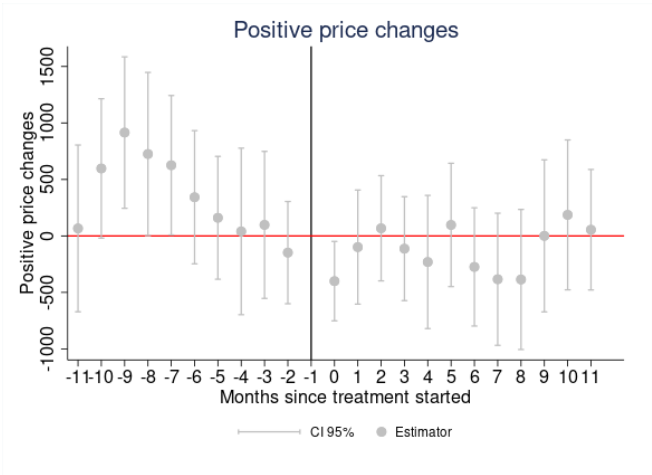


Total number of products with losses

Figure 6: Price Changes

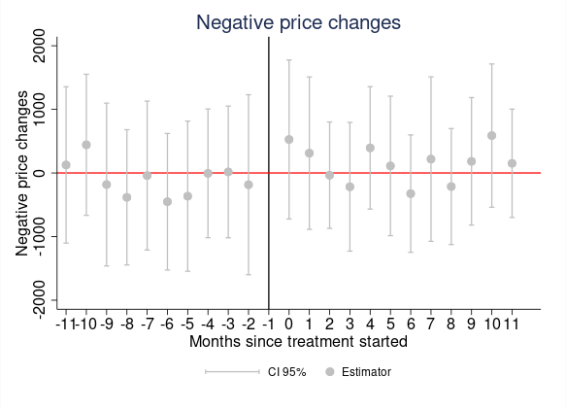


Total number of price changes



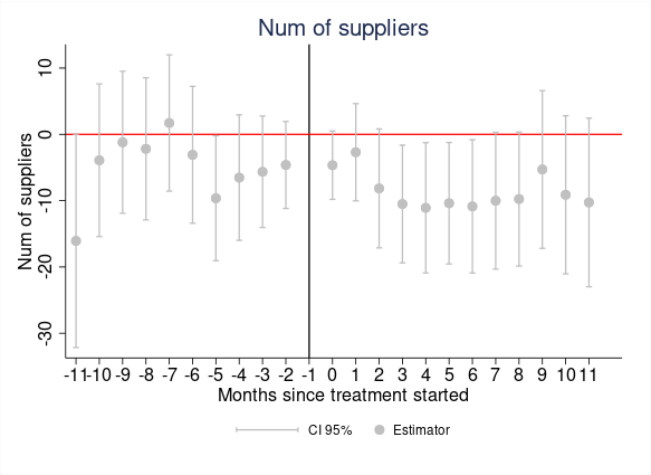
Total number of positive price changes

Figure 7: Total number of negative price changes

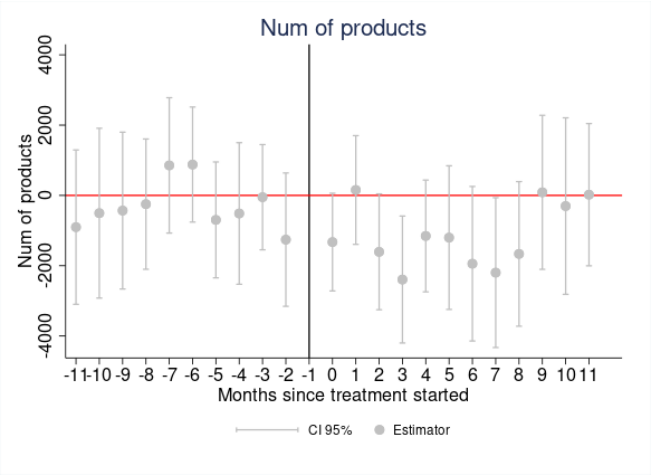


Notes: Figures 5a, 5b, 6a, 6b, 7 show the event-study coefficients and 95 percent confidence intervals from estimating good managers' arrival on losses and price changes variables. The sample consists of store-monthly panel data for the 220 stores during our analysis period of January 2017 to June 2020. The vertical line represents the arrival of a good manager.

Figure 8: Number of Suppliers and Products

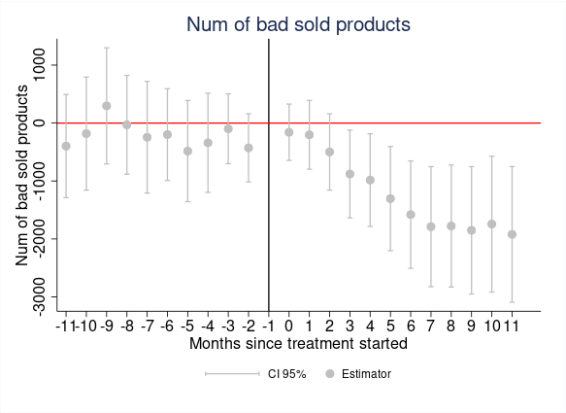


Total number of suppliers



Total number of products

Figure 9: Total number of bad sold products



4.7 Personnel Management and Changes to Store-level Organizational Structure

In this section we discuss shows how a store's organizational structure changes following the arrival of a good manager.

Table 6 shows that the arrival of a good manager causes a significant reduction in the turnover of that store's SPI managers combined with a significant increase in the hiring of new SPI managers.

We also observe a significant effect on the ratios of total number of (non-)managers to total managers in SPI and non-SPI departments.

Bad stores prior to arrival of a good manager have 22 workers total in the SPI departments with roughly 5 of them being managers, whereas non-SPI department contain more than three times that number of workers, with less than half as many managers managers (see Table A.8. Good managers add one worker to the SPI departments over their first year after arrival (see Figure 10a). Half of these new workers are managers (see Figure 10c), number of workers does not change significantly in non-SPI workers but the ratio of managers to workers does go down potentially reflecting movement of managers from non-SPI to SPI. this seems to happen by reducing turnover among SPI managers nearly 40% (similar reduction among non-SPI but not significant) and increasing hiring of SPI managers by more than 50% (no impact on hiring of non-SPI managers) we also see that the movement of managers from non-SPI to SPI departments goes up by 40% (the coefficient rises by a factor of 1.5 in the first month, and 3 in the second month after arrival but then drops to a smaller, but still positive effect after that).

Table 6: Workers and Managers (SPI/non-SPI)

VARIABLES	(1) Total workers (SPI)	(2) Total workers (non-SPI)	(3) Total managers (SPI)	(4) Total managers (non-SPI)	(5) Ratio managers to non-managers (SPI)	(6) Ratio managers to non-managers (non-SPI)
Treatment effect	1.310**	-0.423	0.358**	0.175	0.0949	-0.0184**
	(0.543)	(1.494)	(0.146)	(0.112)	(0.0707)	(0.00926)
Constant	25.91***	78.31***	5.955***	1.939***	0.417***	0.0326***
	(0.639)	(3.725)	(0.0930)	(0.0843)	(0.0360)	(0.00352)
Observations	9,007	9,007	9,007	9,007	6,166	9,007
R-squared	0.990	0.980	0.972	0.964	0.638	0.602
Obs in post	1420	1420	1420	1420	908	1420
Control group	nevert	nevert	nevert	nevert	nevert	nevert
Balanced Panel	nobal	nobal	nobal	nobal	nobal	nobal
stores	220	220	220	220	220	220
stores treated	87	87	87	87	87	87
Arrival controls	Yes	Yes	Yes	Yes	Yes	Yes
Sales control	No	No	No	No	No	No
Mean of Dv	24.81	77.84	5.496	2.387	0.423	0.0732
Effect on mean	0.0528	-0.00540	0.0652	0.0732	0.224	-0.251

Notes: Table 6 shows the point estimates of number of workers, managers, and the ratio of managers to non-managers in SPI and non-SPI departments. SPI refers to departments whose main responsibilities are related to sales, pricing or inventory activities, while non-SPI departments are those with main responsibilities other than those mentioned above. The sample consists of store-monthly panel data for the 220 stores during our analysis period of January 2017 to June 2020. The treatment effect is defined as the arrival of a good manager to a store with low performance (average manager fixed effects of the store above the median). Robust standard errors are in parentheses. * significant 10%, ** significant 5%, *** significant 1%.

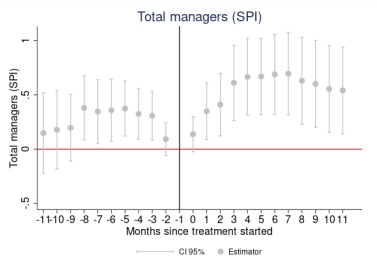
Figure 10: Workers and Managers (SPI/non-SPI)



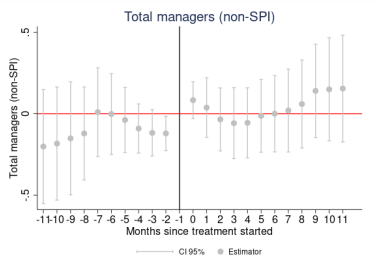
Workers in SPI



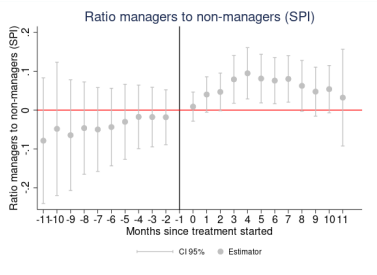
Workers in non-SPI



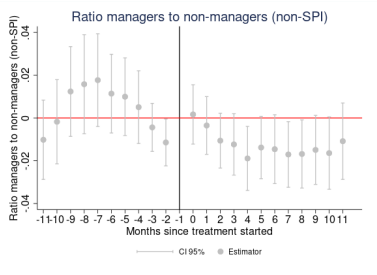
Managers in SPI



Managers in non-SPI



Mngrs/Non-Mngrs in SPI



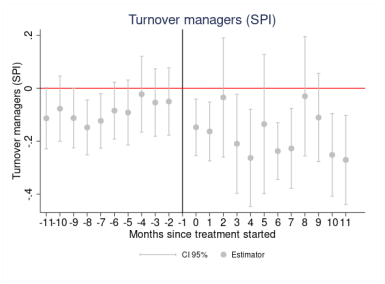
Mngrs/Non-Mngrs

Table 7: Turnover, Hiring and Moves (SPI/non-SPI)

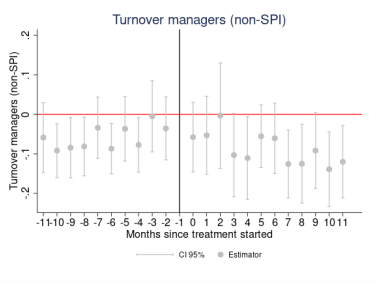
VARIABLES	(1) Turnover managers (SPI)	(2) Turnover managers (non-SPI)	(3) Total hiring (SPI)	(4) Total hiring (non-SPI)	(5) Moves managers to SPI	(6) Moves managers to non-SPI
Treatment effect	-0.117***	-0.0269	0.187***	-0.0829	0.0350	-0.00154
	(0.0335)	(0.0210)	(0.0458)	(0.134)	(0.0231)	(0.00476)
Constant	-0.0236	-0.0757	0.416***	3.623***	0.0479***	0.00568
	(0.0232)	(0.0795)	(0.0314)	(0.147)	(0.0138)	(0.00604)
Observations	9,007	9,007	9,007	8,919	9,007	9,007
R-squared	0.483	0.235	0.277	0.564	0.074	0.076
Obs in post	1420	1420	1420	1395	1420	1420
Control group	nevert	nevert	nevert	nevert	nevert	nevert
Balanced Panel	nobal	nobal	nobal	nobal	nobal	nobal
stores	220	220	220	220	220	220
stores treated	87	87	87	87	87	87
Arrival controls	Yes	Yes	Yes	Yes	Yes	Yes
Sales control	No	No	No	No	No	No
Mean of Dv	0.0754	0.0331	0.369	3.373	0.0500	0.00634
Effect on mean	-1.547	-0.813	0.507	-0.0246	0.701	-0.243

Notes: Table 7 shows the point estimates of number of workers, managers, and the ratio of managers to non-managers in SPI and non-SPI departments. SPI refers to departments whose main responsibilities are related to sales, pricing or inventory activities, while non-SPI departments are those with main responsibilities other than those mentioned above. *Turnover* is defined as the monthly number of manager departures per store. Similarly, *Hiring* is the number of workers hired in the month. As for *Moves*, this variable is constructed as all the movements that occur within the stores in a month, whether vertical or horizontal. The sample consists of store-monthly panel data for the 220 stores during our analysis period of January 2017 to June 2020. The treatment effect is defined as the arrival of a good manager to a store with low performance (average manager fixed effects of the store above the median). Robust standard errors are in parentheses. * significant 10%, ** significant 5%, *** significant 1%.

Figure 11: Turnover, Hiring and Moves (SPI/non-SPI)



Turnover in SPI



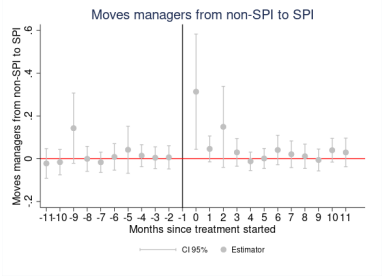
Turnover in non-SPI



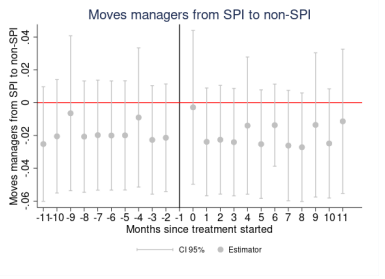
Hirings in SPI



Hirings in non-SPI



Moves to SPI



Moves to non-SPI

5 Survey Evidence

We now present the results of a survey conducted on a subset of managers. While this survey was conducted for an unrelated project (see Weidmann et al. (2024)) it provides helpful, if suggestive, evidence that is broadly consistent with our empirical findings. Because this survey targeted only higher-ranking managers while our empirical results include all managers, we emphasize that all survey results referenced are based on responses only from the managers who make up the highest levels of middle management.

Each panel of Figure 12 compares the survey responses for all managers at bad stores (left bar) to the survey responses of all good managers who moved stores (right bar). This set of the results is the best analog to our definition of treatment (arrival of a good manager to a bad store). We also present results using alternative sample splits (see Figure A.20), finding largely consistent results.

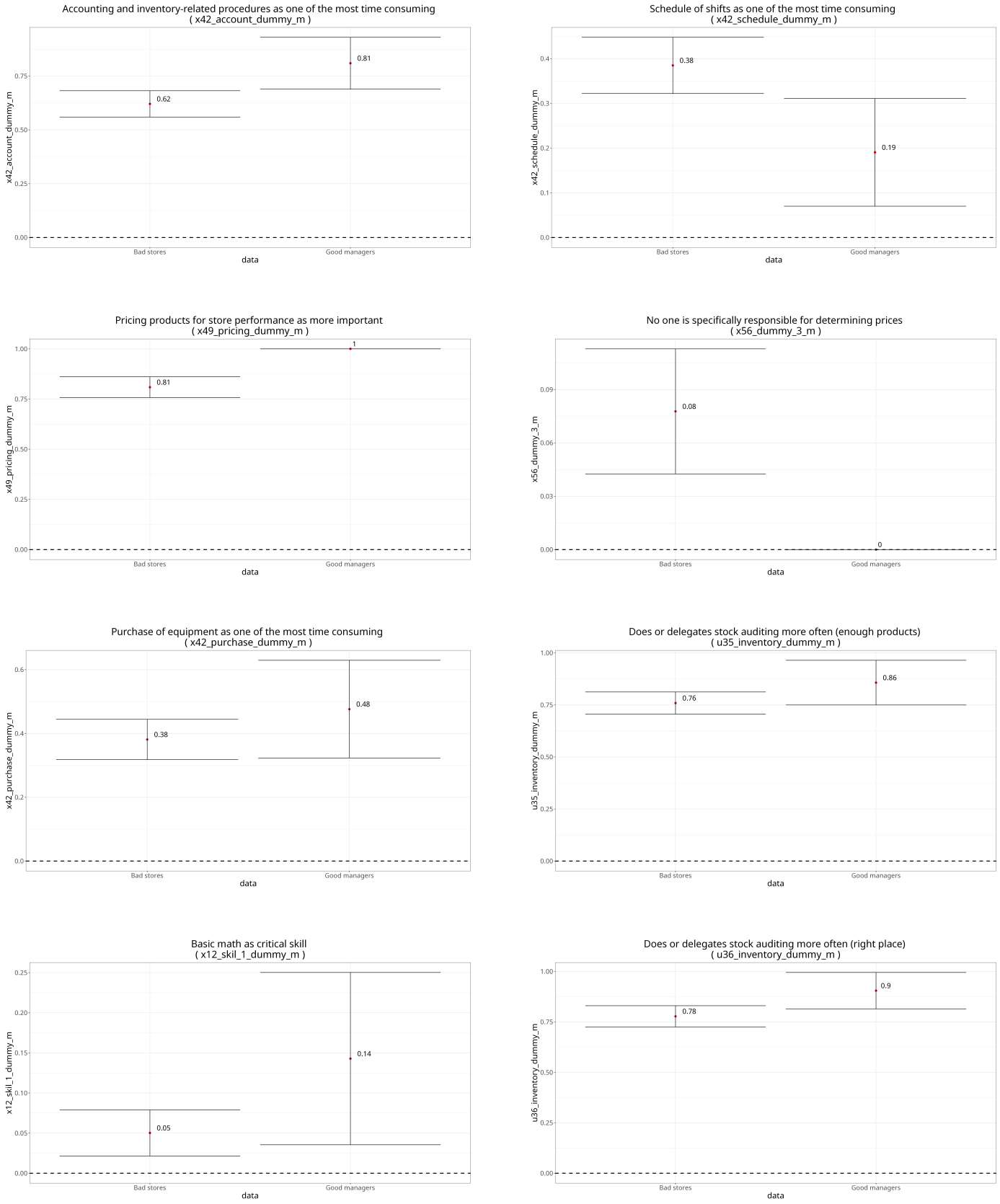
Panels 12a and 12b together support the narrative that good managers spend more

time with accounting and inventory-related procedures and less time with shift scheduling relative to managers at bad stores.

Good managers are also much more likely to report that pricing is one of the most important determinants of store performance (Panel 12c), while managers at bad stores are significantly more likely to report that no one is specifically responsible for determining prices (Panel 12d).

While not significant, the difference in reported coefficients in Panel 12g is indicative of an interesting difference in attitudes between good managers and managers of bad stores: 14% of good managers list basic math as a critical skill to their job, compared only 5% of managers at bad stores. Panel 12h shows another interesting (albeit insignificant) difference: 90% of good managers report delegating stock auditing more often, compared to 78% of managers at bad stores.

Figure 12: Survey Results



Notes: Figure 12 compares survey results for all managers at a bad store (left bar) and all good managers who moved stores (right bar).

6 Robustness

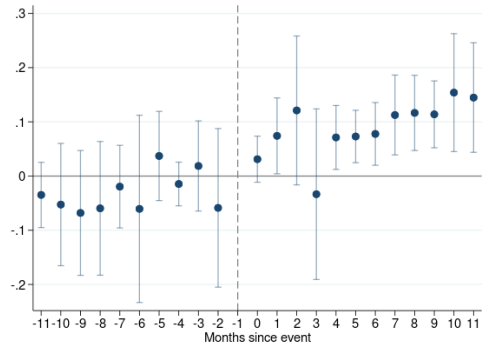
In this section we conduct a series of alternative estimation exercises to demonstrate the robustness of our baseline event-study specification and sample.

6.1 Heterogeneous Effects

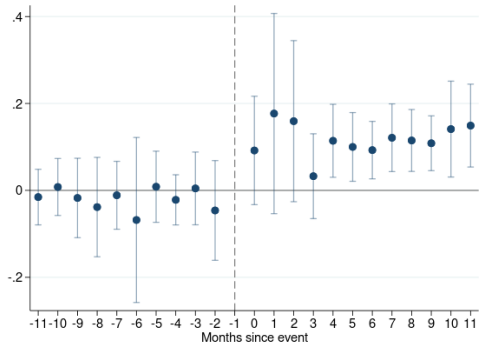
Because our treatment timing is staggered, we must account for the possibility of negative weights in multiple-period difference-in-difference estimators (Athey and Imbens, 2021; De Chaisemartin and d'Haultfoeuille, 2020). Otherwise heterogeneous treatment effects within stores over time or between groups of stores treated at different times might contaminate leads and lags in which all treated stores are pooled across groups (Sun and Abraham, 2020). Goodman-Bacon (2021) shows that difference-in-differences models of the form in (3) yield a weighted average of all possible permutations of pairwise difference-in-differences estimators, where, in our case, a pair is a cohort of stores treated at time t paired with a cohort of stores treated at time $t' > t$.

We address these issues in two ways. First, we estimate the cohort-specific average treatment effect suggested by Sun and Abraham (2020), which translates Callaway and Sant'Anna (2020) group-time average treatment effect from calendar time into relative periods, allowing us to compare cohorts while holding their exposure to the treatment constant. Second, we estimate the group-time average treatment effect, where a group is defined by the time period when stores are first treated. The key assumption in our main sample excluding "never-treated" stores is the conditional parallel trends between stores treated in period g and groups that are "not-yet-treated" by time t. The results of these robustness checks are presented in Figures 13a and 13b. In each case, the results are similar to what we observe in our baseline event-study specification and sample: The relative effect on log sales is 11.7% using our baseline specification, and 8.8% when following Callaway and Sant'Anna (2020). For log productivity, our baseline relative effect is 11.2% compared to 11.0% using the Callaway-Sant'Anna specification.

Figure 13: Effects of a good manager on log sales and productivity: Callaway-Sant'Anna specification



Callaway - Sant'Anna Log sales



Callaway - Sant'Anna Log productivity

7 Discussion

We exploit the quasi-random movement of managers across stores to document substantial sales and productivity gains attributable to manager fixed effects. Our findings reveal that effective management of day-to-day complexity is a key differentiator in managerial performance. Good managers distinguish themselves not by increasing attention to all aspects of operations, but by strategically focusing on the limited set of tasks they can reasonably address. We observe that good managers enhanced sales by ensuring product availability (i.e., reducing stockouts) and making better pricing decisions. They did this while decreasing inventory size and value and simplifying their store's product portfolio by reducing both the number of products and suppliers. This hard-to-achieve combination of fewer stockouts alongside leaner inventory suggests that good managers allocate their attention more efficiently to inventory management.

The organizational response to the arrival of a good manager is also revealing. Upon a good manager's arrival, we observe a significant restructuring, with more personnel being allocated to sales, pricing, and inventory-related departments. Specifically, good managers reduce turnover among SPI (sales, pricing, and inventory) managers while increasing hiring in these departments. This organizational adaptation suggests that good managers recognize the outsized impact of effective SPI management on overall store performance.

Our survey evidence complements these empirical findings, revealing that good managers report spending more time on accounting and inventory-related procedures and are more likely to identify pricing as a critical determinant of store performance.

Taken together, these insights have important implications for retail management practices and organizational design. Firms operating in similarly complex environments may benefit from technologies that help managers better allocate attention across products

and suppliers. Training programs that enhance managers' abilities to identify high-leverage areas for intervention could yield substantial returns. Additionally, organizational structures that facilitate focus on critical inventory, pricing, and sales functions may enhance performance.

Future research should explore whether similar patterns emerge in complex organizational settings beyond retail, and whether technological innovations that reduce the cognitive burden of managing complexity can substitute for managerial quality or instead complement it by allowing good managers to focus their attention on even higher-value activities.

References

Abowd, J. M., Kramarz, F., and Margolis, D. N. (1999). High Wage Workers and High Wage Firms. *Econometrica*, 67(2):251–333.

Adhvaryu, A., Howell, P., Nyshadham, A., and Tamayo, J. (2025). Rapport in Organizations: Evidence from Fast Food. *Working Paper*.

Adhvaryu, A., Kala, N., and Nyshadham, A. (2022). Management and Shocks to Worker Productivity. *Journal of Political Economy*, 130(1):1–47.

Adhvaryu, A., Nyshadham, A., and Tamayo, J. (2023). Managerial Quality and Productivity Dynamics. *Review of Economic Studies*, 90(4):1569–1607.

Adhvaryu, A., Nyshadham, A., and Tamayo, J. (2024). An Anatomy of Performance Monitoring.

Andrews, M., Gill, L., Schank, T., and Upward, R. (2008). High wage workers and low wage firms: negative assortative matching or limited mobility bias? *Journal of the Royal Statistical Society*, 171(3):673–697.

Athey, S. and Imbens, G. W. (2021). Design-based analysis in difference-in-differences settings with staggered adoption. *Journal of Econometrics*.

Bandiera, O., Prat, A., Sadun, R., and Wulf, J. (2014). Span of Control and Span of Attention.

Bloom, N. and Van Reenen, J. (2007). Measuring and explaining management practices across firms and countries. *The quarterly journal of Economics*, 122(4):1351–1408.

Callaway, B. and Sant'Anna, P. H. (2020). Difference-in-differences with multiple time periods. *Journal of Econometrics*.

Card, D., Heining, J., and Kline, P. (2013). Workplace Heterogeneity and the Rise of West German Wage Inequality*. *The Quarterly Journal of Economics*, 128(3):967–1015.

De Chaisemartin, C. and d'Haultfoeuille, X. (2020). Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review*, 110(9):2964–96.

Dessein, W. and Santos, T. (2021). Managerial Style and Attention. *American Economic Journal: Microeconomics*, 13(3):372–403.

Freyaldenhoven, S., Hansen, C., Pérez, J. P., and Shapiro, J. M. (2021). Visualization, identification, and estimation in the linear panel event-study design. Technical report, National Bureau of Economic Research.

Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics*.

Iyengar, S. S. and Lepper, M. R. (2000). When choice is demotivating: Can one desire too much of a good thing? *Journal of Personality and Social Psychology*, 79(6):995–1006.

Kline, P., Saggio, R., and Sølvsten, M. (2020). Leave-out estimation of variance components. *Econometrica*, 88(5):1859–1898.

Lazear, E. P., Shaw, K. L., and Stanton, C. T. (2015). The Value of Bosses. *Journal of Labor Economics*, 33(4):823–861.

Metcalfe, R. D., Sollaci, A. B., and Syverson, C. (2023). Managers and Productivity in Retail.

Roth, J. (2022). Pretest with Caution: Event-Study Estimates after Testing for Parallel Trends. *American Economic Review: Insights*, 4(3):305–322.

Simon, H. A. (1973). Applying Information Technology to Organization Design. *Public Administration Review*, 33(3):268–278.

Sun, L. and Abraham, S. (2020). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*.

Syverson, C. (2011). What Determines Productivity? *Journal of Economic Literature*, 49(2):326–365.

Weidmann, B., Vecci, J., Said, F., Deming, D. J., and Bhalotra, S. R. (2024). How Do You Find a Good Manager?

A Appendix

A.1 Additional Statistics

Table A.8: Additional pre-treatment summary statistics for good and bad stores

	Bad stores		Good stores	
	Mean	SD	Mean	SD
Sales USD	1,704,997.25	1,404,293.38	1,737,059.13	1,617,649.25
Productivity USD	17,745.57	6,787.31	20,323.21	10,322.26
Num. out of stock products	875.46	637.62	894.15	701.74
Stock out value in millions	1,103.41	1,475.44	1,420.57	2,007.61
Num. of days of item out of stock	1,112.96	673.49	1,188.44	916.33
Stock value in millions	1,512.39	1,019.12	1,510.15	1,279.11
Stock value over sales	0.26	0.10	0.24	0.09
Number of products with discounts	8,652.34	8,604.15	9,041.57	10,216.36
Value of products with discount	95,944,752	120,380,336	108,068,208	146,731,680
Total price changes	14,681.63	12,142.79	14,604.98	13,858.86
Positive price changes	5,598.90	4,484.18	5,640.15	5,206.84
Negative price changes	9,082.72	8,353.36	8,964.83	9,321.46
Number of suppliers	840.35	286.07	815.49	367.40
Number of products	54,081.37	43,208.06	55,472.02	53,396.27
Num of bad sold products	13,324.99	10,965.09	4,791.87	6,247.19
Number of any price changes	102,470.44	104,131.01	101,468.84	119,073.88
Number of positive price changes	36,372.20	33,092.43	36,932.28	38,586.28
Number of negative price changes	102,467.01	104,134.30	101,467.80	119,074.72

Notes: Table A.8 reports means and standard deviations for sales-, pricing-, and inventory-related variables for all stores prior to the arrival of a good manager. All statistics are reported at the store-monthly level. A *Good (Bad) store* is a store where the average fixed effect of its managers, computed at the store-month level, is above (below) the median of managers' fixed effects within the connected set. *Productivity* refers to the sales by worker (sales divided by number of workers). *Out of stock products* indicates the products with negative inventory levels in a given month, while *Stock value* shows the product value with positive inventory levels. *Number of days out of stock* refers to the total number of days a store was out of stock. *Price changes* indicates counts the number of times a product have a positive or negative price change between two months. *Bad sold products* are defined as those with sales 0.5 standard deviations below the mean.

Table A.9: Pre-treatment summary statistics for organizational variables in good and bad stores

	Bad stores		Good stores	
	Mean	SD	Mean	SD
Total workers (all)	104.12	90.98	102.06	108.52
Total workers (SPI)	25.75	25.79	25.01	28.32
Total workers (non-SPI)	78.38	67.39	77.05	81.24
Total managers (all)	8.09	7.11	7.95	7.15
Total managers (SPI)	6.07	4.88	6.11	5.89
Total managers (non-SPI)	2.02	3.60	1.84	1.64
Ratio managers to non-managers (all)	0.10	0.08	0.12	0.12
Ratio managers to non-managers (SPI)	0.52	0.75	0.49	0.67
Ratio managers to non-managers (non-SPI)	0.04	0.07	0.07	0.11
Turnover managers (all)	0.18	0.93	0.21	1.12
Turnover managers (SPI)	0.14	0.75	0.15	0.89
Turnover managers (non-SPI)	0.05	0.28	0.05	0.31
Total hiring (all)	4.49	10.97	4.92	8.97
Total hiring (SPI)	0.34	0.91	0.44	1.18
Total hiring (non-SPI)	3.54	5.59	4.11	7.20
Total hirings (managers)	0.01	0.11	0.01	0.12
Hiring managers (SPI)	0.01	0.10	0.00	0.07
Hiring managers (non-SPI)	0.00	0.05	0.01	0.09
Moves managers (all)	0.061	0.389	0.043	0.320
Moves managers to SPI	0.058	0.377	0.035	0.294
Moves managers to non-SPI	0.003	0.055	0.007	0.091
Arrival of any manager	0.13	0.33	0.15	0.36

Table A.9 reports means and standard deviations for organizational variables for all stores prior to the arrival of a good manager. All statistics are reported at the store-monthly level for positions in all, SPI and non-SPI departments. A *Good (Bad) store* is a store where the average fixed effect of its managers, computed at the store-month level, is above (below) the median of managers' fixed effects within the connected set. *Workers* refer to any person employed in the retail store during the period analyzed. *Managers* are defined as those workers who are coordinators, leaders, bosses and directors. Conversely, *Non-managers* are those workers who are not in those positions. *Turnover* indicates the number of workers in the store that left the retail store, while *Hiring* represents the number of workers in the store that were hired. Finally, *Moves* indicates the number of moves from one position to another within the company (it includes vertical or horizontal moves).

A.2 Additional Evidence

Table A.10: Effects on losses and price changes

VARIABLES	(1) Total num of products with losses	(2) Total value of losses	(3) Total price changes	(4) Positive price changes	(5) Negative price changes
Treatment effect	-1,552***	-2.802e+07***	-3.618	-321.9*	318.2**
	(481.8)	(8.131e+06)	(247.6)	(172.4)	(123.9)
Constant	7,029***	7.091e+07***	14,153***	5,186***	8,967***
	(503.9)	(8.921e+06)	(175.6)	(132.2)	(74.57)
Observations	9,007	9,007	9,007	9,007	9,007
R-squared	0.901	0.873	0.930	0.908	0.892
Obs in post	1420	1420	1420	1420	1420
Control group	nevert	nevert	nevert	nevert	nevert
Balanced Panel	nobal	nobal	nobal	nobal	nobal
stores	220	220	220	220	220
stores treated	87	87	87	87	87
Arrival controls	Yes	Yes	Yes	Yes	Yes
Sales control	No	No	No	No	No
Mean of Dv	6614	7.020e+07	13108	4801	8307
Effect on mean	-0.235	-0.399	-0.000300	-0.0670	0.0383

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

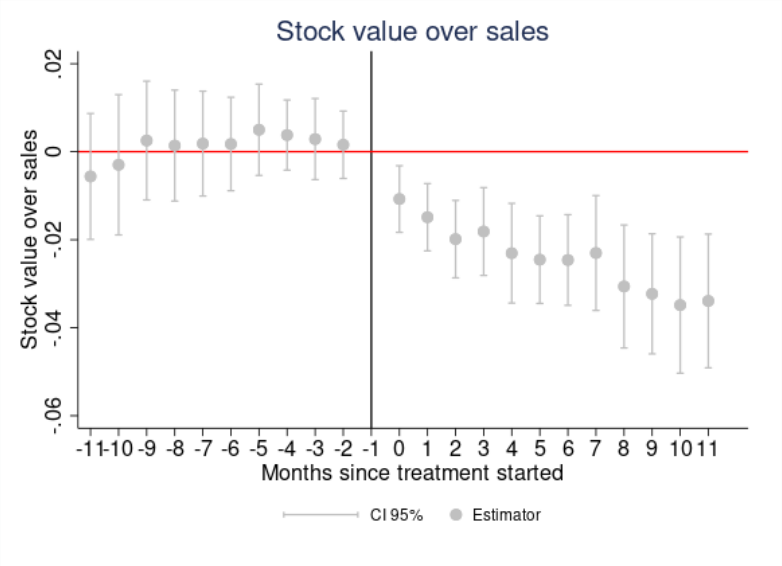
Notes: Table A.10 shows the point estimates of losses and price changes variables. Losses are define as cases when the selling price is lower than the value of inventory of a product. Price changes ar defined as those cases when there was a positive, negative or both price changes. The sample consists o store-monthly panel data for the 220 stores during our analysis period of January 2017 to June 2020. Th treatment effect is defined as the arrival of a good manager to a store with low performance (averag manager fixed effects of the store above the median). Robust standard errors are in parentheses. significant 10%, ** significant 5%, *** significant 1%.

Table A.11: Effects on suppliers, products and bad sold products

VARIABLES	(1) Num of suppliers	(2) Num of products	(3) Num of bad sold products
Treatment effect	-2.971	-146.3	-960.7*
	(4.221)	(680.9)	(490.4)
Constant	799.3***	52,951***	9,501***
	(4.370)	(652.4)	(364.9)
Observations	9,007	9,007	9,007
R-squared	0.991	0.980	0.891
Obs in post	1420	1420	1420
Control group	nevert	nevert	nevert
Balanced Panel	nobal	nobal	nobal
stores	220	220	220
stores treated	87	87	87
Arrival controls	Yes	Yes	Yes
Sales control	No	No	No
Mean of Dv	754.3	47782	7325
Effect on mean	-0.00390	-0.00310	-0.131

Notes: Table A.11 shows the point estimates of number of suppliers, products, and bad sold products. Bad sold products are those with sales 0.5 standard deviations below the mean. The sample consists of store-monthly panel data for the 220 stores during our analysis period of January 2017 to June 2020. The treatment effect is defined as the arrival of a good manager to a store with low performance (average manager fixed effects of the store above the median). Robust standard errors are in parentheses. * significant 10%, ** significant 5%, *** significant 1%.

Figure A.14: Stock value over sales - On-hand Inventory Only



Notes: This figure shows the event-study coefficients and 95 percent confidence intervals from estimating good managers' arrival on out-of-stock and on-hand variables. The sample consists of store-monthly panel data for the 220 stores during our analysis period of January 2017 to June 2020. The vertical line represents the arrival of a good manager.

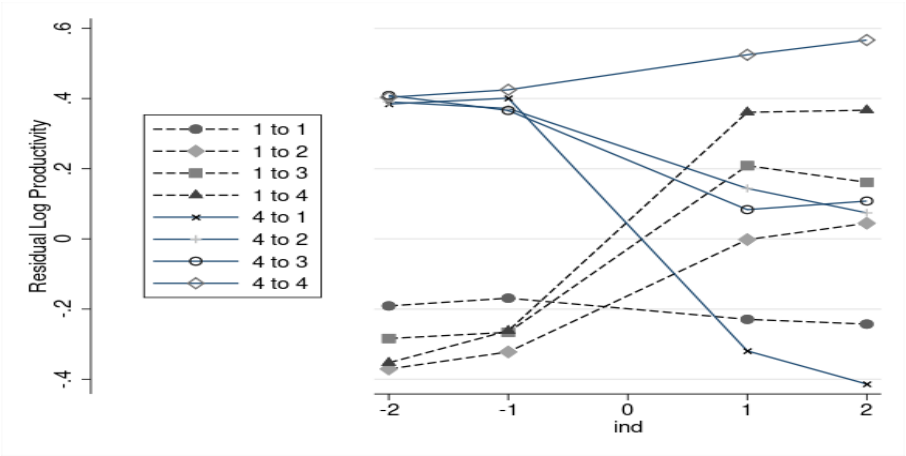


Figure A.15: Event Study Around Movers- Productivity (*logs.*)

Notes: Figure A.15 ranks movers in terms of (i) quartiles of average log productivity in their initial store and (ii) quartiles of the average log productivity in the store where they moved to. The average log productivity is computed over the entire sample period, and quartiles are calculated for each store. The graphical representation depicts the average residual of productivity (*logs.*) of movers on the y-axis; the residual is computed for specific periods: from 3 to 4 months (Period = -2) and 1 to 2 months (Period = -1) before the move from the initial store, and 1 to 2 months (Period = 1) and 3 to 4 months (Period

= 2) after the move to the new destination store, plotted on the x-axis. The analysis focuses on moves away from stores in the top quartile (lines in quartile 4) and stores in the bottom quartile (lines in quartile 1). To create the residual variable, we run a regression of the monthly log productivity of each store on year and month fixed effects. Then we predict the residuals and run the movers analysis.

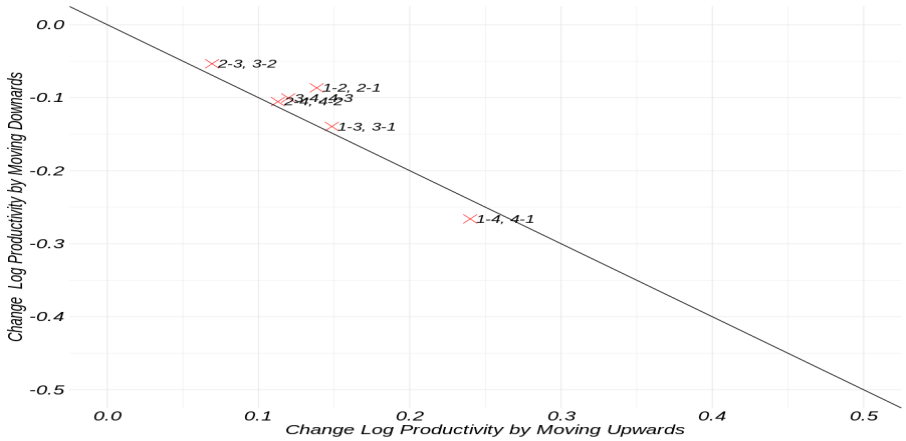


Figure A.16: Symmetry Test- Productivity (*logs.*)

Notes: Figure A.16 ranks movers in terms of (i) quartiles of average productivity (*logs.*) in their initial store and (ii) quartiles of the average productivity (*logs.*) in the store where they moved to. The average productivity (*logs.*) is computed over the entire sample period, and quartiles are calculated for each store. The Figure then plots the average change in residual of log productivity (*logs.*) of movers from lines in quartile X to quartile Y, against the change in residual of productivity (*logs.*) for movers in the opposite direction; for example, the point labeled "2 to 4, 4 to 2" corresponds to the average change for movers from lines in quartile 2 to quartile 4, plotted against the change for movers from lines in quartile 4 to quartile 2. The changes are calculated for average residual of log productivity (*logs.*) in the two months before the move and the two months after the move. The solid line corresponds to the 45-degree line. To calculate the log productivity residual, we run a regression of productivity (*logs.*) in the stores on year and monthly fixed effects. Then, we predict the residuals and run the movers analysis.

Figure A.17: Explanatory power of AKM on log productivity

```
Cov(log_prod, xb) / Var(log_prod):              .09157458
Cov(log_prod, peff) / Var(log_prod):            .12048224
Cov(log_prod, feff) / Var(log_prod):            .71834721
Cov(log_prod, res) / Var(log_prod):             .06959605
```

Table A.12: Explanatory power of AKM on log productivity

Covariance term	Fraction of $var(\omega)$
$Cov(\omega, x\beta)$	0.0916
$Cov(\omega, \theta)$	0.1205
$Cov(\omega, \psi)$	0.7183
$Cov(\omega, \varepsilon)$	0.0696

Table A.12 reports the explanatory power of each term in the AKM on the variance of productivity (ω). Each row refers to the covariance between the productivity (*logs.*) and the covariates ($x\beta$), the manager fixed effects (θ), the store fixed effects (ψ), and the residuals of the model (ε).

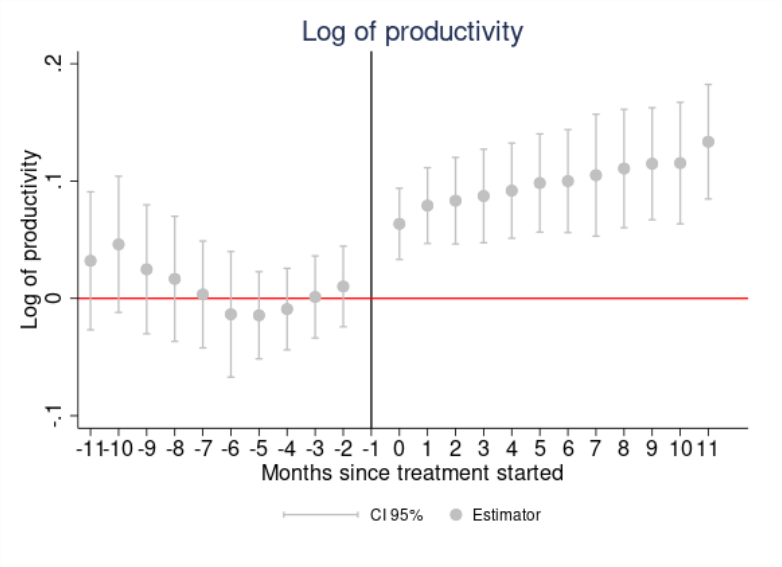
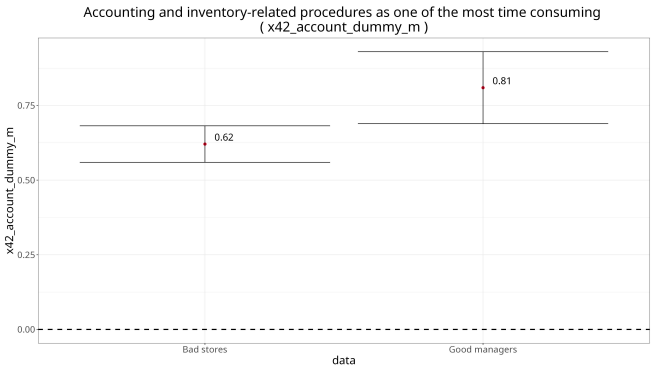
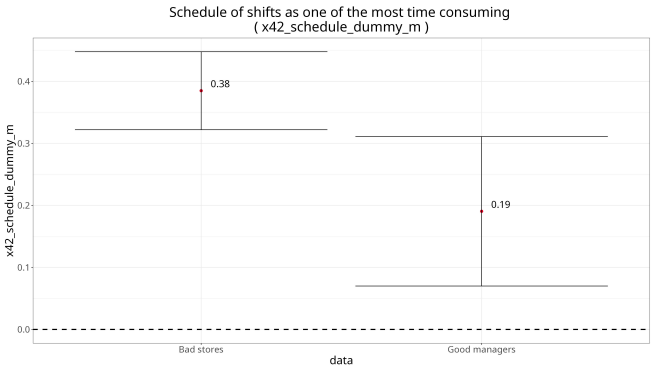


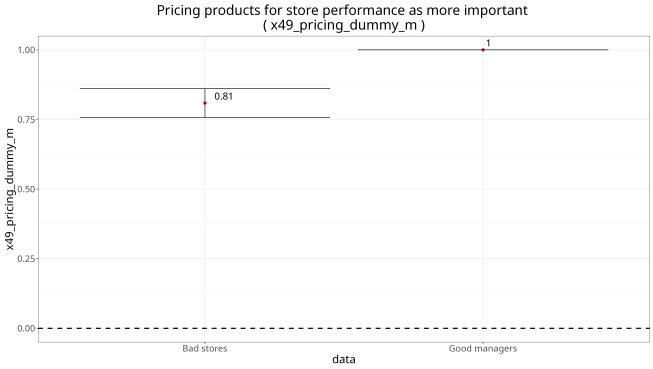
Figure A.18: Effect of Treatment on Log Sales

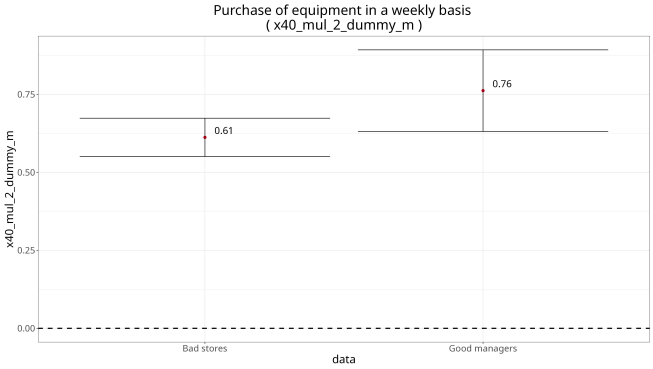
Notes: Figure A.18 shows the event-study coefficients and 95 percent confidence intervals from estimating good managers' arrival on log productivity. The sample consists of store-monthly panel data for the 220 stores during our analysis period of January 2017 to June 2020. The vertical line represents the arrival of a good manager.

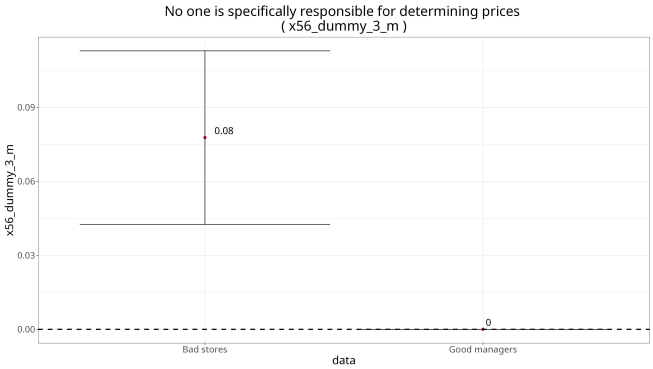
Figure A.19: Additional Survey Results

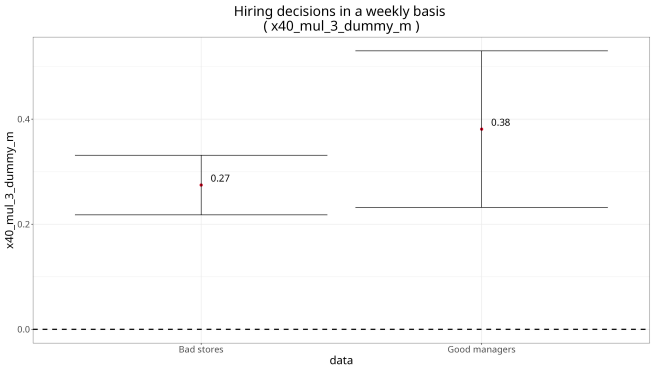


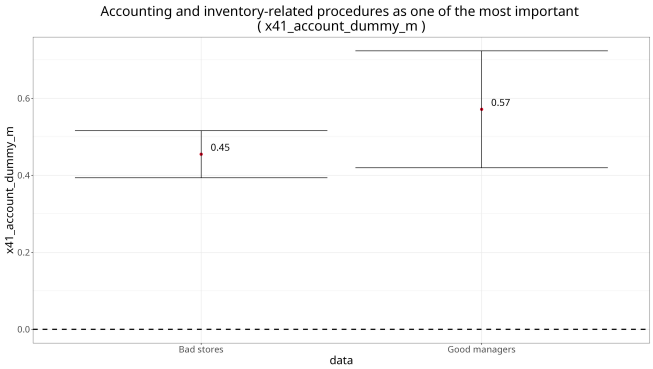


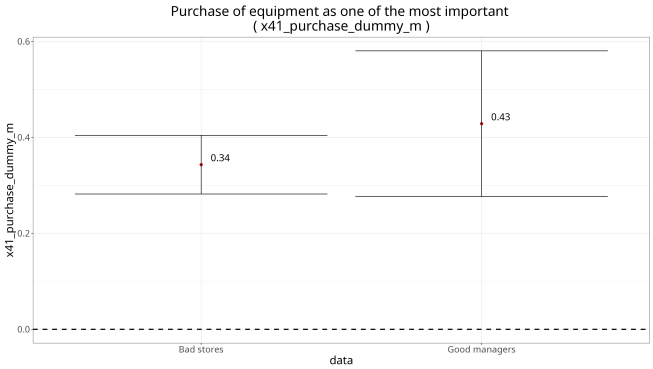


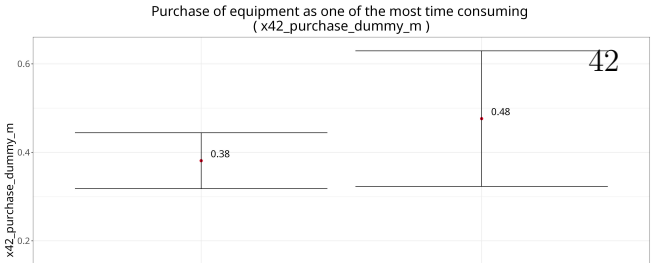












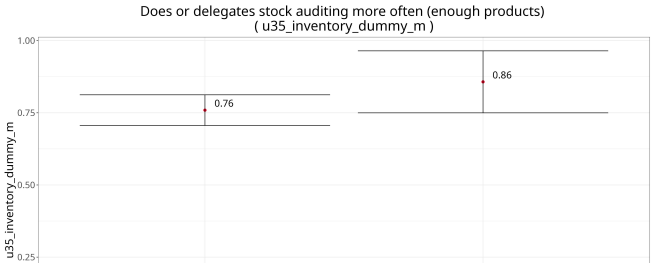
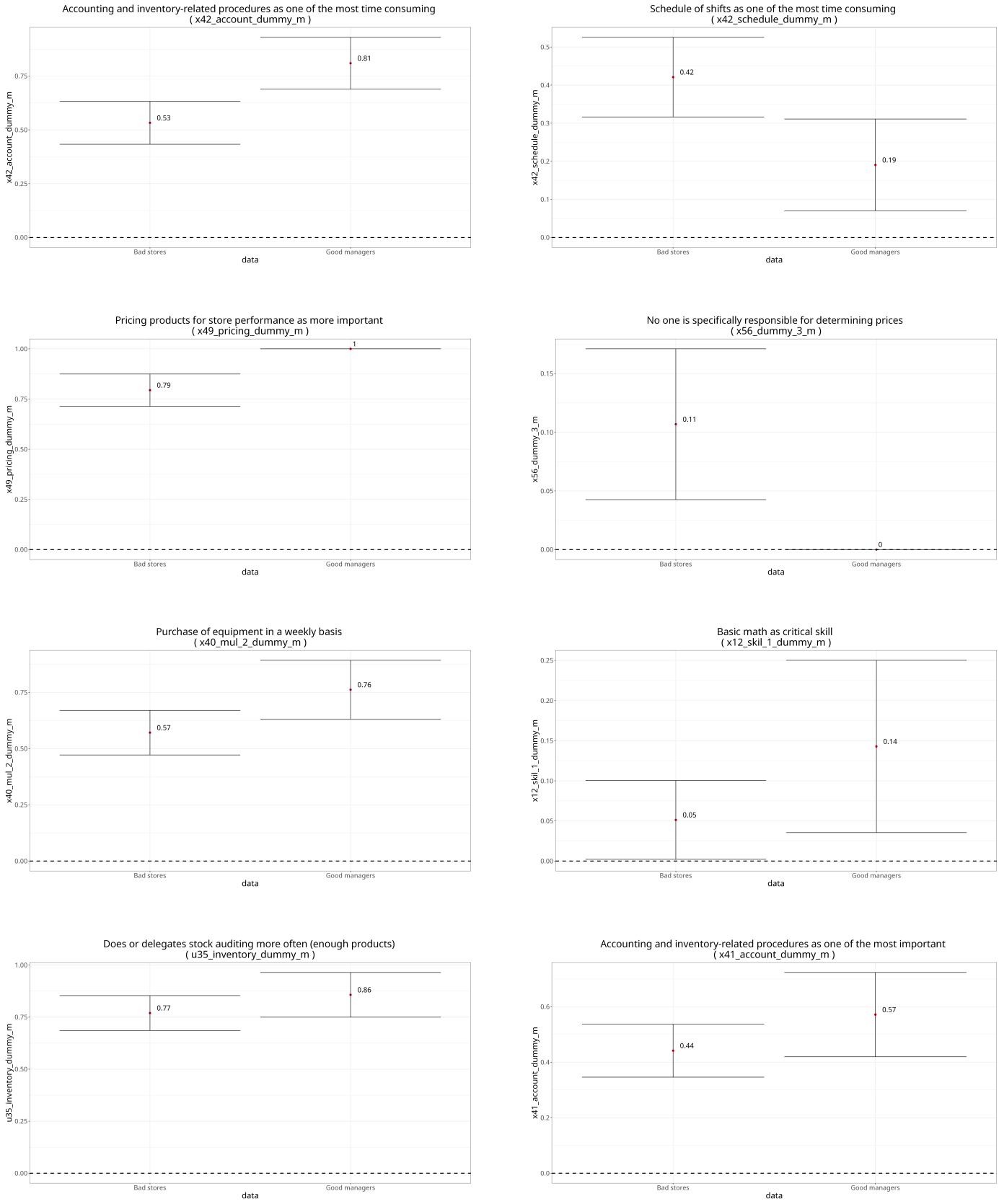


Figure A.20: Survey Results for Alternate Samples



Notes: Figure A.20 shows ... good movers v treated bad stores.

A.3 Additional Robustness Results

Table A.13: AKM based on different approaches

	Baseline	Andrews et al. (2008)	Leave-out Estimator
Var(θ)	0.024	0.018	0.018
Var(ψ)	0.205	0.199	0.120
Cov(ψ, θ)	-0.010	-0.005	-0.005
Corr(ψ, θ)	-0.138	-0.081	-0.078

Table A.13 reports the baseline model which comes from the estimation of equation (1) following Abowd et al. (1999), the bias correction of Andrews et al. (2008) and leave-out Estimator from Kline et al. (2020). The data spans over January 2017 to June 2020 through 220 stores. For each model we compute the variance of the manager fixed effects $Var(\theta)$, the variance of the stores fixed effects $Var(\psi)$, and the correlation $Corr(\psi, \theta)$ and covariance $Cov(\psi, \theta)$ of both type of fixed effects.