Stanford | King Center on Global Development

Modernizing Retailers in an Emerging Market: Investigating Externally-focused and Internally-focused Approaches

Stephen J. Anderson

Leonardo Iacovone

Shreya Kankanhalli

Sridhar Narayanan

May, 2021

Working Paper No. 1088

Modernizing Retailers in an Emerging Market:

Investigating Externally-focused and Internally-focused Approaches

Stephen J. Anderson Leonardo Iacovone Shreya Kankanhalli Sridhar Narayanan

May 23, 2021

Abstract

This paper studies the impact of business modernization on the sales performance of traditional retailers. We define modernization as adopting physical structures and tangible practices ubiquitous in organized retail chains (for example, exterior signage with store name and logo, or a database to record product-level information). To address our research question, we implement a randomized field experiment in Mexico City with 1148 traditional retail firms. Our sample is randomized into three groups: 385 firms that we externally modernize in ways that are visible to customers; 383 firms that we internally modernize in ways that are not visible to customers; and 380 firms form a control group. We find a significant and persistent main effect of modernization on sales: firms in both treatment groups increase monthly sales by 15% to 19%, even 24 months after study recruitment. In terms of novel mechanism evidence, we find that externally-modernizing firms improve their store-level branding, while internally-modernizing firms strengthen their product management. These results have important implications for multinational managers who distribute products through traditional retail channels, and for policymakers interested in improving firm performance in the retail sector of emerging markets.

Keywords: Retail, emerging markets, branding, product management, small firm growth, field experiment

*Corresponding Author: Kankanhalli, Ph.D. Candidate in Marketing, Stanford University, T +1 (650) 285-7436, shreyak@stanford.edu. Anderson, Assistant Professor of Marketing, University of Texas at Austin, sjanderson@mccombs.utexas.edu; Iacovone, Lead Economist, World Bank and Hertie School, Liacovone@worldbank.org iacovone@hertie-school.org; Narayanan, Associate Professor of Marketing, Stanford University, sridhar.narayanan@stanford.edu. We thank seminar participants at Columbia Business School, Marketing Science Conference 2019, Notre Dame Mendoza College of Business, the QME Rossi Speaker Series, Penn State Smeal College of Business, Stanford GSB, the 11th Triennial Invitational Choice Symposium, UC Berkeley Haas School of Business, USC Marshall School of Business, and the UT McCombs School of Business for comments. Special thanks to Juan Espinosa, Marla Garcia, Florencia Leyson, Eduardo Roche, Merrell Guzman, and Jose Avalos for exceptional research assistance. We are grateful to Mario Romero Carsolio and Fundes for their excellent work. Additionally, we thank Mexico's Secretaría de Hacienda y Crédito Público, particularly the efforts of Vicente Garcia and Eduardo Piedra. Finally, we thank Luz Gomez and MasterCard Center for Inclusive Growth for outstanding insights and support. This research was supported by the UPS Foundation Endowment Fund (Stanford), the John A. and Cynthia Fry Gunn Faculty Scholar Award (Stanford), the Stanford King Center on Global Poverty and Development, Stanford GSB, and World Bank's Latin America and the Caribbean Gender Innovation Lab (LACGIL). LACGIL is supported by the Umbrella Facility for Gender Equality (UFGE), a World Bank multi-donor trust fund financing research, impact evaluations, and data to help policy makers and practitioners close gender gaps.

INTRODUCTION

In the retail sector globally, millions of traditional stand-alone shops — such as the *Kiranas* of India, *Tienditas* of Mexico and *Xiao Mai Bu* of China — co-exist alongside modern organized retail chains and compete with them to serve end-customers. Indeed, in emerging markets, traditional retailers dominate, accounting for 57% of annual retail sales (Bronnenberg & Ellickson 2015). But stark differences exist between traditional and modern retailers, both in the physical structures and tangible practices that are visible to customers, and those that are not. For example, modern retail outlets post clear signage with their business name and logo on the store exterior, display products and prices in a thoughtful manner, and present customers with promotions. Meanwhile, traditional outlets invest much less on their appearance and presentation to customers, as evident in the picture of a traditional retailer shown in Figure 1. These externally-focused differences are visible to customers. Other important differences between traditional and modern retailers are more internally-focused and less noticeable to customers. For instance, modern retailers often maintain robust internal systems to track product-level demand, flows of product stock, and record product margins, whereas traditional outlets usually operate without such systems.

Why do these differences exist? On the one hand, traditional retailers may lack the necessary information or capital to *modernize*, i.e., to adopt physical structures and tangible practices ubiquitous in organized retail chains (for example, exterior signage with store name and logo, or a database to record product-level information). In this case, relaxing constraints around modernization might help retailers become more competitive and grow their sales. In particular, externally-focused modernization could enhance the store-level brand of the retailer, while internally-focused modernization could improve their product management. On the other hand, remaining traditional might be a deliberate strategic choice of retailers. A no-frills presentation could help the traditional retailers signal lower prices or higher accessibility to their emerging market customers. Moreover, devoting time and resources to modernizing in a traditional environment — for example, by manually maintaining handwritten records of each customer transaction — could come at the cost of other business activities that are more critical for earning sales. In this scenario, modernizing a retailer may not

increase their sales. In fact, sales may plausibly decrease with modernization either because consumers who might otherwise have purchased at the shop think it is too expensive or inaccessible, or because owners substitute from other sales-enhancing activities to modernization-related activities.

Thus, the impact of modernization for a traditional retailer is an open empirical question, due to the competing explanations for why they remain traditional in the first place. We answer this question through a randomized controlled field experiment in Mexico City. We randomly assign a sample of 1148 traditional retail firms into three experimental groups: (i) a control group (N = 380) that receives no modernization intervention; (ii) an *external* treatment group (N = 385) that receives an intervention to modernize in ways that are visible to customers; and (iii) an *internal* treatment group (N = 383) that receives an intervention to modernize in ways that are not visible to customers. Firms in the two treatment groups receive 13 sessions from a *Modernization Agent*, where each session is dedicated to hands-on, direct implementation of structures and practices common to modern retailers. Thus, treated firms spend a total of 35 hours making tangible modernization changes with our assistance. Assuming they spend 1-2 additional hours on their own to complement the activities of each session, our interventions constitute at least 55 total hours spent on modernizing one's business. Subsequent to the interventions, we collect outcome data on all study firms 12 months and 24 months after they were recruited, to measure medium- and long-term effects. This study design allows us to causally identify the magnitude and direction of the effect of modernization on sales, as well as the differential impact of external versus internal retail modernization.

We find positive, statistically significant, and persistent effects of both external modernization and internal modernization on firm sales. Two years after they were recruited to participate in the program, firms in the external treatment group increased monthly sales by US $518 and firms in the internal treatment group increased monthly sales by US $430, relative to the control group. Given that an average control firm earns US $2776 in monthly sales, these effects represent an 18.7% and 15.5% performance improvement, respectively. This recurring monthly sales bump is also economically significant, corresponding to roughly 1.5 months in rental expenses for a small shop in our sample, or the monthly wage of 1.5 employees. The magnitude of this sales increase is

Figure 1: Traditional and Modern Retail in Neighbourhood of Roma, Mexico City

Traditional Retailer — Unnamed

Modern Retailer — Abarrotes Delirio

notable as our post-intervention qualitative fieldwork reveals that many traditional retailers do not modernize due to uncertainty in returns. We estimate that the cost of modernizing one's business (for 55 hours) can be recovered within six months. We also offer exploratory insights on how to modernize for the highest returns. Specifically, our analysis suggests that: (i) modernizing one's exterior appearance, customer engagement methods, demand analysis, and stock ordering processes are most useful for growing sales; and (ii) there are diminishing returns to modernization on sales.

Additionally, this study provides new mechanism evidence linking the process of modernization to sales gains. As a mechanism for their sales improvement, we find that retailers in the external treatment group significantly improved their store-level branding (versus the control group). We used novel methods to study the impact of modernization on store-level branding. In one approach, involving the analysis of photographs of the shops, independent raters assessed retailers in our sample on standard measures of branding from the marketing literature — such as brand loyalty, brand excitement, brand sophistication, and brand quality (D. A. Aaker 1991; Keller 1993; J. L. Aaker 1997). We also randomly surveyed actual customers of the businesses on branding measures. Under both approaches, external group firms were rated significantly higher on branding measures. In contrast, firms in the internal treatment group significantly improved their product management relative to the control group, as evaluated by independent auditors. With modern internal systems to

track product-level demand, flows of inventory, and product margins, retailers were able to gain a deeper understanding of their products and make more informed decisions using the product data. This included choosing an improved product assortment, improving choice of suppliers, and ordering optimal quantities of products to avoid spoilage and stock-outs.

The findings of this study can inform the practice of retail as well as policy. Multinational manufacturers recognize the importance of traditional retail firms for reaching end-customers in emerging markets, and view modernization as a potential means of enhancing product sales. This is evident from modernization initiatives already in place, such as efforts by Reliance and Grupo Bimbo to install POS systems in traditional shops of India (Abrams 2019) and Mexico (Grupo Bimbo Press Release 2013), respectively. Similarly, policymakers, like our study collaborators at Mexico's Ministry of Finance and the World Bank, are concerned with the performance of traditional retail firms due to the sheer size of this sector[1], and the vast number of livelihoods tied up in traditional retail[2]. In spite of this managerial and policy interest, no prior academic work has determined the direction of the modernization effect on the sales performance of traditional retailers, quantified the effect size, or analyzed mechanisms through which sales gains may occur.

Thus, this paper aims to make three contributions to the academic literature in marketing and economics. First, we provide causal evidence to address the empirical puzzle outlined above: traditional retailers are in fact able to increase sales when they adopt the structures and practices of modern retailers. A nascent literature in marketing has described the co-existence of traditional and modern retail, documenting economy-wide trends in their sales share (Bronnenberg & Ellickson 2015) and examining consumer choice between the two retail formats (Narayan *et al.* 2015). Our finding that modernization improves sales suggests that retailer decisions to remain traditional do not simply reflect strategic considerations, such as the preferences of low-income consumers. The large magnitude of our modernization effects show the *extent* of non-strategic behaviors that persist among traditional retailers — there is significant scope for these firms to improve upon their current sales and growth performance.

[1]Retailing accounts for around 20% of GDP in most global economies as per UN National Accounts Databases.

[2]Around 15% of Mexico's labour force is occupied in traditional retail (Labor Force Survey, INEGI-ENOE, 2017).

Second, we design and implement original interventions that directly change physical structures and tangible practices in traditional retailers so they resemble more modern retailers. The impact of on-site business support interventions is understudied for retailers with fewer than five employees (Bruhn *et al.* 2018). This context presents novel challenges as traditional retailers have little experience in trusting and receiving advice from external parties[3]. Moreover, the content and approach of our interventions is also distinct from the classroom-based training or consulting interventions previously studied in the development literature that aim to build business skills (such as Calderon *et al.* 2020; de Mel *et al.* 2014; Beaman *et al.* 2014; Bloom *et al.* 2013). We do not rely on the retailer to independently translate theoretical lessons (from an educator) or advice (from a consultant) into business skills and then apply those skills to make changes in their shop. This approach is not only lengthy, but also uncertain to materialize. By directly making physical changes at a retail shop, we leapfrog the theory-skill-application process of typical training and consulting programs. Our modernization interventions can therefore be implemented through local university students and intuitive graphical instruction manuals rather than through professional management consultants. This novel type of intervention can be added to the toolkit of policymakers or managerial stakeholders interested in scalable solutions for traditional retailers.

Finally, as our third contribution, we analyze novel marketing-related mechanisms for explaining sales improvements — enhanced store-level branding and improved product management. In the marketing literature, firm branding attempts have been studied in the context of large corporations in advanced markets, using descriptive or observational approaches due to data limitations (Keller & Lehmann 2006). No prior study has experimentally analyzed how small independent retailers (lacking specialized marketing teams and marketing budgets) can build a stronger brand, or the implications of such branding on firm performance. We show that an upgraded store-level brand matters in an understudied and unexpected context, i.e., among low-income customers who frequent traditional shops in emerging markets. Moreover, while marketing scholars have described the importance of product management through observational studies of large firms (Cosse & Swan 1983;

[3]Confirmed in our qualitative fieldwork, as outlined in Table 9.

Varley 2014), prior work has not empirically demonstrated how product management capacities can be built from scratch in low-technology settings through simple internal structures, such as handwritten notes tracking product-level information. Our study provides evidence on how both store-level brand building and improved product management are possible and beneficial for firm sales growth in emerging market retail.

The rest of the paper proceeds as follows. First, we develop hypotheses on modernization and sales performance. Next, we describe the experimental design and data collected. Subsequently, we present the results and finally, we conclude with implications.

CONCEPTUAL BACKGROUND AND HYPOTHESES

Across advanced and emerging markets, modern retail chains have been characterized as successful ventures driving increases in national productivity (Basker 2007; Bronnenberg & Ellickson 2015). This motivates the recommendation that traditional retailers should modernize. However, it is not clear that the physical structures and tangible practices of modern chains would yield the same (net) benefits when adapted to the context of small traditional shops. We see this for other business decisions. For instance, most modern retailers formally register their business with tax authorities and adopt information technology. Yet the calculus on these decisions is different for traditional retailers in emerging markets, often leading them to strategically make the opposite choices (De Mel *et al.* 2013; Sudhir & Talukdar 2015). Following a similar logic, the sign and size of the modernization effect for traditional retailers is (ex ante) unclear. Below, we describe mechanisms through which external and internal modernization might have a positive effect on sales. Subsequently, we discuss why differences in available technology, labour (both in terms of the number of employees and their skills), monetary resources, and types of customers may imply that modernization has negative effects for traditional retailers in emerging markets.

We focus on sales as the main outcome of interest for a number of reasons. For multinational product manufacturers who distribute products through traditional retailers and policymakers focused on retail sector growth, sales expansion is a primary objective. Marketing academics also recognize sales as a key performance indicator to be tracked and impacted by interventions (Farris *et*

al. 2010). Moreover, given that revenues are well-understood and salient to firm owners, sales is a key outcome used by researchers relying on primary data collected via firm audits in emerging markets.

External Modernization and Retail Firm Sales

To discuss how external modernization might impact sales, we first present a comprehensive set of *structures* — an umbrella term we henceforth use to describe physical objects, tangible systems, or observable practices present in modern retailers — that fall under the scope of external modernization. Figure 2 lists the 20 external modernization structures we propose as relevant to traditional retailers, all of which are visible to customers. Illustrative examples of external modernization structures include a large sign with business name and logo on the store exterior, an attractive product display, and customer communications via media. The 20 external structures are grouped under five modules: (i) exterior appearance; (ii) interior appearance; (iii) sales tactics; (iv) price labels and promotions; and (v) customer engagement. Each of these modules correspond to critical touchpoints (i.e. places of interface between businesses and their customers) discussed in the retailing literature (Stein & Ramaseshan 2016).

We propose that implementing these external modernization structures could help traditional retail firms increase their sales by enhancing their store-level branding. eOf these two concepts, store-level branding relates more closely to corporate branding, in that it pertains to the entire stand-alone retail shop rather than specific products stocked within. The corporate branding literature has suggested that various modern external structures contribute positively to brand building, such as displayed names (Kohli & LaBahn 1997), displayed logos (Simonson & Schmitt 1997; Grewal *et al.* 1998), interior cleanliness and attractiveness (Morales 2005), customer engagement materials (Bresciani & Eppler 2010), and sales staff appearance and knowledge (W. G. Kim & H.-B. Kim 2004). However, the corporations studied have large marketing departments, specialized marketing knowledge, and large associated budgets, to dedicate to brand-building. For instance, Madden *et al.* 2002 measure investments in brand-building using advertising expenditures as a proxy, which has a median value of US $9.1 million in the Compustat database they utilize. Firms in our sample, on the other hand, were willing to spend up to US $50 on modern external structures. To our knowl-

Figure 2: 20 External Modernization Structures

List of 20 external modernization structures, grouped by module	
Module 1: Exterior Appearance	*Module 4: Prices Labels and Promotions*
Clean and tidy storefront	Fixed and labelled prices
Large sign with business name, logo	Prices competitive with rivals
Posted store hours, open/close sign	Promotions for high-margin products
Signs of popular products	Promotions for low-rotation products
Module 2: Interior Appearance	*Module 5: Customer Engagement*
Clean and tidy shop interior	CRM database
Products unpacked and shelves filled	Customer communications via media
Well-painted and well-lit interior	Customer loyalty program
Attractive product display	Receipts provided to customers
Module 3: Sales Tactics	
Professional appearance of sales staff	
Well-informed sales staff	
Customer-service provided by sales staff	
Direct-selling techniques by sales staff	

edge, we are the first to test whether small shop owners lacking specialized marketing or business education, and with limited budgets, can engage in effective brand building using these modern external structures as levers. We suggest that they can, especially when they implement multiple structures in a coordinated and thoughtful manner as in our external modernization intervention.

The corporate branding literature also argues that brand building can be translated into brand equity, i.e., increases in the market valuations of these corporations (Keller 1993; Madden *et al.* 2002). In a similar vein, it may be possible for traditional retailers to leverage brand building to generate higher sales. In particular, brand building could help businesses drive repeat customer visits, attract new customers, and raise spending per visit (or share of wallet). Hence, we hypothesize:

Hypothesis 1 *Greater external modernization will increase the sales of traditional retail firms.*

Hypothesis 2 *External modernization improves the store-level brand of traditional retail firms as one mechanism for the increase in sales.*

On the other hand, brand building may not play out in the same way for traditional stand-alone retailers as it has for corporate chains investigated in the branding literature. Through extensive

fieldwork in Mexico City, we find that traditional retailers largely compete with other traditional retailers in their neighborhood. The lower-income consumers they compete for in these markets are often price-sensitive (Kamakura & Mazzon 2013) and priced out of modern retail (Narayan *et al.* 2015). Given this, they may find modern retailers intimidating and inaccessible. Thus, these customers may take their demand elsewhere if a traditional store starts to seem less accessible with modern upgrades, similar to the way lower income segments in the US have avoided shopping at the chain establishments replacing locally-owned stores (Monroe Sullivan & Shaw 2011). Moreover, these consumers may make price inferences from modern external structures that work against traditional retailers: they may perceive the store and its products as expensive, whether or not prices have actually risen in practice.

The existing branding literature has few causal studies on how firms build and leverage their brands, with none that focus on small independent retailers. This body of work also lacks a rigorous empirical study that directly links brand perceptions to firm sales as the outcome of interest (as opposed to stock market valuations, which do not apply to the millions of retailers that are not publicly listed). This is largely due to data limitations and identification challenges[4]. Our rigorous field experiment therefore contributes to this literature by causally examining store-level brand building and its relationship to firm sales in a novel context (i.e., emerging market neighborhoods with lower-income consumers) where the effects of brand building are theoretically ambiguous for the reasons above.

Internal Modernization and Retail Firm Sales

To discuss how internal modernization impacts sales, we first present in Figure 3 a set of structures whose adoption constitutes modernizing a retailer in ways that are internally-focused. All 20 structures listed are not visible to consumers, ubiquitous in the back-end of modern retail firms, and currently lacking in traditional retail. Key examples of internal modernization structures include a system to record sales by product, a system to record stock inflows, and an organized stock area. The 20 internal structures are grouped under five modules: (i) demand analysis; (ii) earnings analysis;

[4]Empirical studies of branding rely mostly on customer panel surveys of famous corporate brands or financial market responses to branding (e.g. Du *et al.* 2019; Madden *et al.* 2002)

(iii) stock ordering; (iv) stock quality; and (v) managing cashflow. Our internal modernization structures were designed to have a strong product focus. They pertain to knowing the demand for each product, knowing the margins earned on each product, tracking inflows and outflows of product stock, and storing products effectively, as well as other product-related activities.

Thus, implementing internal modernization structures could help traditional retail firms increase sales by improving their product management capacities. In the marketing strategy literature, product management is conceptualized as an important organizational function that "introduc[es] new products and manages existing products effectively through their life cycle" (Tyagi & Sawhney 2010). In retailing, product management involves assortment choice, maintenance of product quality, optimally selecting suppliers and maintaining good relationships with them, and procuring optimal quantities of products in a cost-effective manner (Varley 2014). We propose that internal modernization structures are important inputs to product management.

Figure 3: 20 Internal Modernization Structures

List of 20 internal modernization structures, grouped by module	
Module 1: Demand Analysis	*Module 4: Stock Quality*
System to record sales by product	Organized stock area
Daily sales recording implemented	Proper storage of stock
System to record expenses by product	Listing of current supplier information
Listing of COGS for each product	System to research offerings of new suppliers
Module 2: Earnings Analysis	*Module 5: Managing Cashflow*
Listing of profit margins for each product	Credit register for goods sold on credit
Listing of monthly operating profits by product	Separation of business and personal finances
Listing of sales targets by product	Savings plan for investment goals
Listing of monthly fixed costs	Bank account to save operating profits
Module 3: Stock Ordering	
System to record stock inflows	
System to record stock outflows	
Listing of stock re-order levels	
System for stock orders	

This is supported in the marketing literature. In the large firm context, Cossé and Swan (1983) argue that data on products, particularly demand and contribution margin estimates, are critical to

better performance by product managers as they enable strategic decisions on product assortment and targeting of sales efforts to specific products. Internal modernization structures related to demand analysis and earnings analysis (as per our intervention) correspond exactly to collecting such product data. Stock-related internal structures also facilitate improved product management — for example, records of stock flows allow retailers to order the right quantities of products (Corsten & Gruen 2003; Verhoef & Sloot 2006), while stock-area renovations improve product storage (and therefore, product quality). Finally, internal structures related to managing cash-flow can assist retailers with several aspects of product management. Savings generated from proper cash-flow management can be used by retailers to add to their product assortment, upgrade existing product quality, and maintain cash-on-hand to pay suppliers for all stock required in a timely manner.

We subsequently argue that improved product management can generate higher sales for retailers. Marketing studies have linked individual dimensions of product management, such as product assortment and product quality, positively to firm performance (Hoch *et al.* 1999; Mitra & Golder 2006). Moreover, procurement-related dimensions of product management can positively affect sales too, by helping retailers avoid sales losses from stock-outs or product spoilage, and increasing quantities of products stocked. Therefore, we hypothesize:

Hypothesis 3 *Greater internal modernization will increase the sales of traditional retail firms.*

Hypothesis 4 *Internal modernization improves the product management of traditional retail firms, as one mechanism for the increase in sales.*

However, our alternative hypothesis is that traditional retailers may find it difficult to leverage internal modernization to earn higher sales. Unlike large modern retailers, small traditional retailers do not typically have access to information technology that automates product record-keeping (Sudhir & Talukdar 2015), and so it remains a time-consuming, cognitively-challenging exercise that could detract from other important activities, such as interacting with consumers and advertising the firm. Moreover, while skilled specialists make product-related decisions at large modern retailers, in these traditional outlets the responsibility falls on non-specialist owners who often lack formal business education and are severely time-constrained as they manage daily operations with only a

handful of employees. These owners may not be able to analyze product records and translate the quantitative insights into better product management. For similar reasons, finance and accounting training has not been found in the development economics literature to lead to behavior changes among firm owners that grow sales (Drexler *et al.* 2014; Anderson *et al.* 2018).

To the best of our knowledge, prior studies in marketing, strategy or economics have not causally analyzed whether the product management capacities of a small firm can be improved, nor found an empirical association between overall product management and firm sales. In advanced markets where modernized internal structures are ubiquitous among retail firms, it is difficult to obtain observational data on firms with substantial variation in product management. Through our field experiment on modernization in a setting where internal structures are (ex ante) largely not modernized in retail firms, we are able to causally test for the first time whether traditional retailers can improve their product management capacities and also link such improvements to firm sales.

EXPERIMENTAL DESIGN

In order to investigate the impact of modernization on sales performance, we implement a randomized controlled field experiment in Mexico City with 1148 small retailers. We do so because collecting and analyzing observational data from a cross-section of retailers would present obstacles to causal identification. First, omitted variables such as entrepreneurial ability (e.g., grit, cognitive capacity, self-control), improved access to capital (e.g., due to personal wealth or social connections), or time trends are likely to be correlated with modernization levels of firms, and are also likely to have a positive association with their sales. Second, as modernization can be seen as a costly investment, reverse causality concerns are present, in that firms with greater sales are more likely to have the resources to modernize. In contrast, randomly assigning some retail firms to receive interventions that modernize their business activities would allow us to identify causal effects.

We conducted the field experiment from January 2017 to December 2018 in partnership with Mexico's Ministry of Finance and the World Bank. 1148 retail firms in our sample were randomly assigned to one of three experimental groups: 385 external treatment firms; 383 internal treatment

firms; and 380 control firms. Importantly, control firms did not receive any modernization intervention but were recruited and surveyed in the same manner as the treated firms and, thus, the control group represents a valid counterfactual against which we can compare changes in business outcomes.

Sample Recruitment and Randomization

We recruited our sample of firms from May to December 2017 by going door-to-door and approaching around 10,000 small-scale retail businesses operating across the Mexico City metropolitan area in Mexico. Prior to recruitment, recruiters were trained to approach every eligible business by asking to speak to the owner and delivering a marketing pitch for participating in our "business support" program. The marketing steps and materials we used to recruit firms are outlined in Web Appendix A. In the recruitment material, the language we used was intentionally vague — the firms were not made aware that the objective of the program was modernization, let alone different types of modernization. We avoided specifying what the program entailed to prevent selection into the sample on the basis of modernization need, interest, or effectiveness. Recruiters systematically covered all "hot spots" where small retailers were known to operate.

Recruited firms had to meet certain eligibility criteria: (i) the firm must have been operational in the last 30 days (had at least one transaction); (ii) the business owner must have been available to participate in an 8-week program; and (iii) the business had to be operating from a permanent physical structure such as stand-alone shop, or a shop in large premise such as a mall or office complex. A permanent physical location can be thought of as a proxy for how established and committed an owner is to their business. We purposefully screened out firms without a permanent physical location from our sample for both conceptual and practical reasons. Conceptually, such non-established firms are unable to modernize significantly, in spite of interventions like ours (where many modernization changes cannot be implemented without a physical premise), and thus fall outside the scope of our study. From a practical standpoint, it would be very difficult to follow up with such non-established firms given their mobility and low survival rates. Moreover, to ensure ease in following up with our sample throughout a multi-year study, we confirmed that all firm owners actively consented to program participation (though program content was not specified to

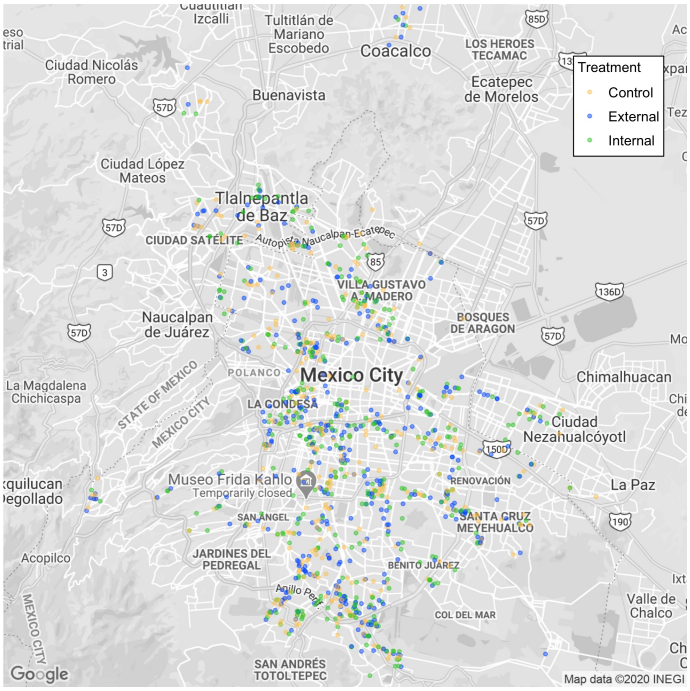
them). This systematic recruitment process left us with around 1600 firms who were invited for a baseline survey, conducted as an on-site 90-minute audit of the business.

At this stage, firm owners were told that due to popular demand there were more people interested in the program than available spots, so a lottery would determine who gets the program in 2018 versus during a future roll-out. This helped maintain commitment throughout the study period and minimize any systematic attrition from the control group. In the end, our study sample included the 1148 small retail firms in greater Mexico City whose owners willingly completed the baseline survey. To effectively manage logistics, we split the sample into four batches. We then randomized and delivered the modernization interventions batch-by-batch (timeline detailed in Table 1). Randomization happened at the individual firm-level, immediately prior to the scheduled intervention delivery for each batch. Randomization was stratified by location (five residential zones), baseline profits (above- and below- median profits), and sub-sector. The locations of our study firms are shown on the map in Figure 4. Given the comprehensive nature of our recruitment approach, we believe this sampling frame is fairly representative of the small traditional retail environment in Mexico City — although deliberately containing (for conceptual and practical purposes) retailers that expressed interest in receiving "business support" and operated out of a minimum physical structure required for modernization. Additionally, our consent requirements and eligibility criteria ensured that our sample recruitment closely resembles the process a policymaker or multinational firm manager might use to select firms for modernization support, which enhances the ecological validity of our study.

Table 1: Timeline for Experiment Steps

Experiment Activity	**Dates**
Recruitment Survey	June 2017 to Ocotober 2017
Baseline Survey	October 2018 to February 2018
Intervention delivered to Batch 1	March 2018 to May 2018
Intervention delivered to Batch 2	June 2018 to August 2018
Intervention delivered to Batch 3	September 2018 to October 2018
Intervention delivered to Batch 4	November 2018
Midline Survey	November 2018 to April 2019
Endline Survey	September 2019 to March 2020

Figure 4: Map of 1148 Firms in Sample, by Treatment Assignment



Modernization Interventions

We designed and implemented two distinct interventions — one that modernized businesses externally, and one that modernized businesses internally. The interventions had the same format and intensity. In both interventions, a firm was partnered with a *Modernization Agent* (i.e. a top university student majoring in business, economics or related disciplines) who visited the business 11 times, for sessions that lasted 2.5 hours. These sessions were exclusively dedicated to hands-on implementation of modernization structures. Agents were trained and supervised by senior managers from a well-known international NGO (FUNDES), who also made 2 site visits that were 2.5 hours in length. In addition, the business owners typically spent 1-2 additional hours on their own after each session ensuring modernizing structures were fully incorporated. Thus, retailers could spend $\sim$55 hours (or more) modernizing their business through our interventions.

What differed between the two interventions was the menu of modernization structures that agents were instructed to implement in the business. Within the external intervention, agents were

trained and instructed to implement the menu of 20 external modernization structures in Figure 2, all 20 of which were visible to customers. In order to illustrate the external intervention in practice, we show several modernizing structures implemented by firms receiving this intervention, in Figure 9 of Web Appendix B. By contrast, for the internal intervention, agents were trained and instructed to implement the menu of 20 internal modernization structures in Figure 3, all 20 of which were internally-focused and not visible to customers. Refer to Figure 10 of Web Appendix B for example structures implemented in the internal intervention. To prevent contamination between the two distinct treatment arms, each agent was trained to implement only one type of intervention and was not informed of the existence of the other intervention (or experimental groups).

A novel feature of the two interventions is that they directly focus on making concrete modernization changes in the business. This is distinct from the classroom-based approaches of standard business training programs reviewed in McKenzie and Woodruff (2014) and is understudied in the literature. Modernization Agents dedicated all their time with the business owner to direct physical implementation of modernization structures. Our approach is pedagogically motivated by research in entrepreneurship showing that learning-by-doing plays a more critical role than formal business education for lasting behavioral changes in entrepreneurs (Politis 2005).

We took three additional steps to ensure that the interventions were strong and likely to permanently modernize the treated firms. First, within the assigned menu of 20 possible modernization structures, we encouraged retail owners to choose what to implement based on their unique business needs. The Modernization Agent provided a diagnostic in their first visit, where they identified areas of strength and weaknesses. This helped the firm owner to decide which modernization structures were worth focusing on. Firm owners also paid for any materials required to implement structures, which ensured that the choice of structures reflected careful consideration. Second, we prioritized depth over breadth, and trained Modernization Agents accordingly to focus on making a few lasting modernization changes rather than skim across many areas. Third, we designed a novel pedagogical tool – a graphical instruction manual – and gave it to every treated firm owner to promote their modernization. A close analogy to our manual are do-it-yourself manuals that

homeware stores provide, in which each step of a self-assembling process has an associated diagram. This mostly-pictorial tool was helpful given the low levels of formal education in our sample and reinforced the action-oriented (rather than theoretically-oriented) theme of the intervention. Pages from the two intervention manuals are included in Figure 11 of Web Appendix B.

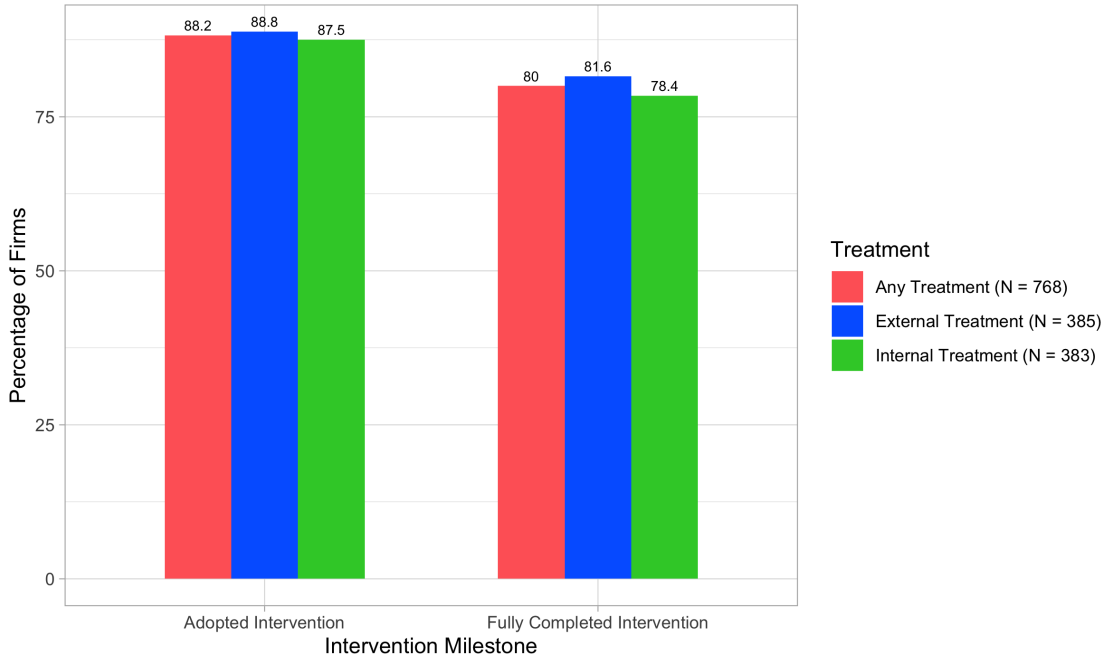
Intervention Checks

We developed numerous checks to ensure that the interventions were successful in modernizing businesses. Based on these checks, we present three sources of evidence on intervention strength: treatment adoption rates, compliance rates, and pre-to-post intervention changes in modernization structures for treatment compliers. First, we note that adoption of and compliance with assigned treatment was high, as reported in Figure 5. We define adoption by the business owner as completing the first session of the intervention. 88.8% of firms assigned to receive the external intervention adopted, as did 87.5% of firms assigned to receive the internal intervention. Moreover, there was little drop-off in intervention continuation after the first session, with 80% of all treatment group firms completing the full 13 sessions provided[5]. In other words, there was high compliance with completing our interventions. Business owners who complied with either treatment also reported a high degree of satisfaction with the intervention (detailed in Web Appendix B Figure 14).

Finally, the interventions resulted in expected improvements in modernization levels for treated firms. In the first and last sessions of the intervention, we measured successful implementation of external (internal) modernization structures by external (internal) treatment group firms. In the first session, the Modernization Agent went through the checklist of 20 external (internal) modernization structures in Figures 2 (Figure 3), requesting the firm owner to present evidence on the presence of each modernization structure. Subsequently, they privately gave a score of 0 to 5 on the quality of implementation of each modernization structure. We count a modernization structure as *verified* if scored 4 or above, a fairly strict threshold. In the last session, the monitoring Senior Manager (a different auditor) conducted the same measurement process to assess the firm's progress.

[5]44% of firms who dropped out of receiving the treatment were no longer eligible, 28% of drop-outs reported lack of trust as their reason, and 22% reported lack of time.

Figure 5: Firm Adoption and Compliance with Assigned Intervention



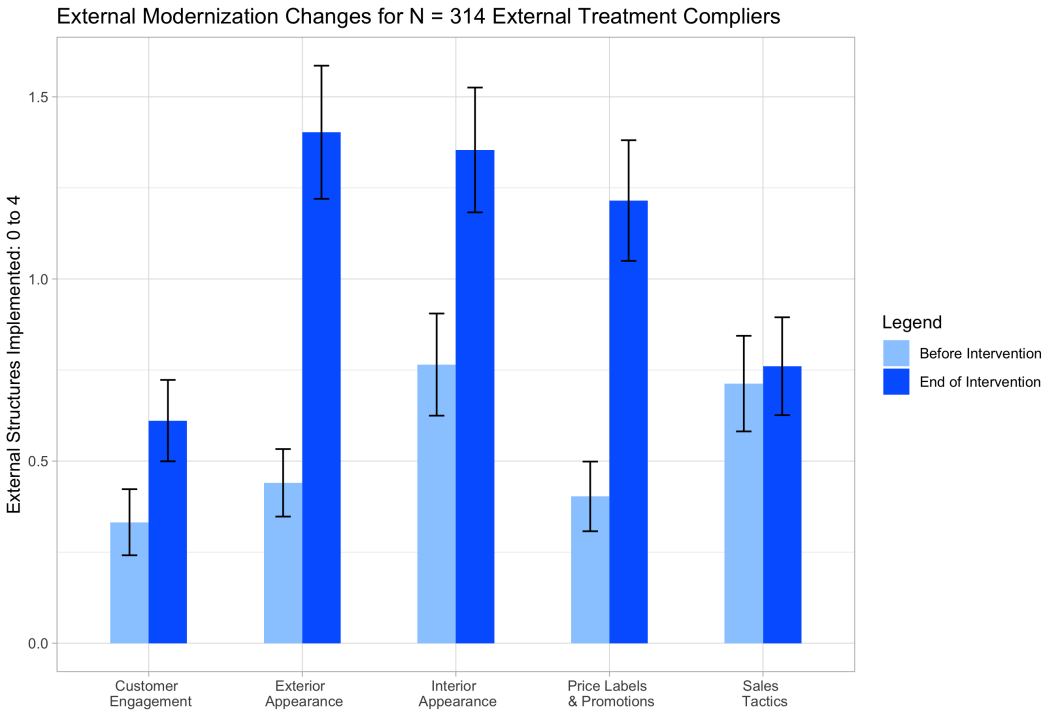
Note 1: Adopted refers to completing the first introductory session with the Modernization Agent

Note 2: Fully completed refers to finishing all 13 Modernization Sessions

If our interventions were successful, we would expect to see an improvement in verified external (internal) modernization structures for external (internal) treatment compliers from the first to last sessions. Indeed, we show in Figure 6a) that firms complying with the external treatment improved considerably on external modernization structures pre-to-post intervention. In particular, they made significant improvements on structures relating to their shop's exterior and interior appearance, areas highly salient to customers. Similarly, in Figure 6b), we show that firms complying with the internal treatment improved considerably on internal modernization structures. In particular, they had implemented systems to analyze demand and profit margins at the granular product level.

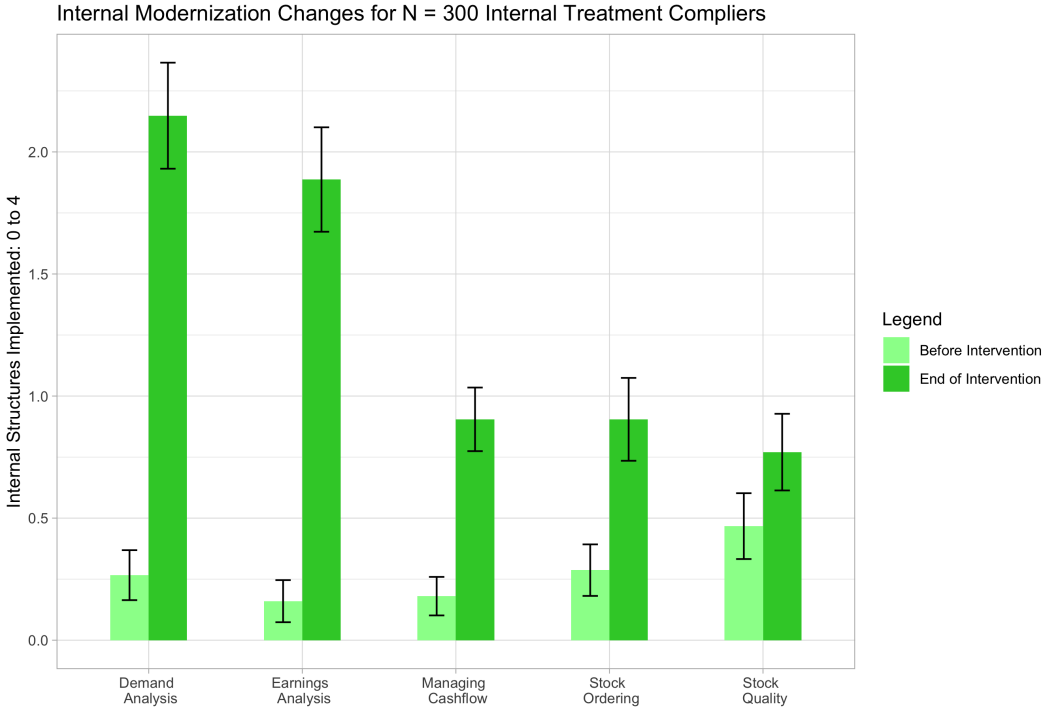
In Table 2, we summarize these changes in structures pre-to-post intervention. At the end of the intervention, firms complying with the external treatment had implemented 5.3 external modernization structures (out of the list of 20), almost doubling their average of 2.7 at the beginning of the intervention. The improvement was even larger for the internal treatment compliers, who went from having 1.4 internal modernization structures (out of list of 20) at the beginning of the

Figure 6: Modernization Changes by Treatment Group



Note 1: Audit of external modernization structures prior to external intervention was conducted by Modernization Agent assigned to the firm.

Note 2: Audit of external modernization structures after the external intervention was conducted by Senior Manager from NGO Fundes.



Note 1: Audit of internal modernization structures prior to internal intervention was conducted by Modernization Agent assigned to the firm.

Note 2: Audit of internal modernization structures after the internal intervention was conducted by Senior Manager from NGO Fundes.

intervention, to 6.6 at the end of the intervention. Both of these improvements are statistically distinguishable from zero at the 1% level[6].

Table 2: Modernization Changes due to Interventions

	Before Intervention	**At End of Intervention**
	External Group (N = 314)	External Group (N = 314)
External Structures: Count 0–20	2.653	**5.344**
Intervention Effect: Δ in Structures		**+2.691*****
	Internal Group (N = 300)	Internal Group (N = 300)
Internal Structures: Count 0–20	1.361	**6.615**
Intervention Effect: Δ in Structures		**+5.254*****

Notes: This table summarizes the change in modernization levels of firms before and after receiving interventions. The data comes from audits of modernization structures implemented by firms during their first and last intervention sessions. Data was collected for N = 614 firms (314 External Group and 300 Internal Group) who completed the intervention, thus receiving their last visit. The range for Δ in structures was [0,16] for the external group, and [0,18] for the internal group. P-values for the difference in pre-intervention vs. post-intervention means are highlighted as: p $< 0.1^{*}$ p $<0.05^{**}$ p $< 0.01^{***}$

Placebo Tests

In field experiments, the control group is often provided with a "light" treatment, analogous to placebo pills in medical trials. This serves two purposes: (i) to rule out that treatment effects are due to generic aspects of receiving a treatment such as the feeling of receiving a reward, receiving attention and assistance, or a motivation boost; and (ii) to ensure that control participants do not systematically attrite as a result of not receiving anything through their participation in the program. Our "light" treatment is a two-page business report created using each firm's pre-intervention data. We show an example of the report in Web Appendix B Figure 12. This report compares a firm's pre-intervention financial performance and business practices to other firms in their sector, and offers the generic advice: "Increase your sales and reduce your costs!". We presented this report (in hard-copy) to all firms after they completed the recruitment process. We also presented a certificate

[6]Additionally, in support of the idea that our intervention facilitated learning-by-doing, we find six months post-intervention that external (internal) treatment group firms not only continued to implement these external (internal) modernization structures, but also had adopted new structures.

of completion to all firms, including the control group, after the modernization interventions had been delivered. By doing so, we ensured that the control group also felt that they received a reward, or attention and assistance, or a motivation boost.

Additionally, incorporating two different treatment groups into our experimental design naturally builds in placebo tests for mechanism-related variables. Both modernization interventions are of equal length and identical format (thus implying the same attention, assistance, or motivation boost for firms), yet we do not expect externally-modernizing firms to improve on product management (the alternative mechanism channel). Likewise, we do not expect internally-modernizing firms to improve on store-level branding. Taken together, these steps help us ensure that the treatment effects presented later can be attributed to the process, and specific type, of modernization itself.

DATA COLLECTION AND CHECKS

We implemented four rounds of data collection to assess the impact of modernization on business performance, two rounds pre-intervention (recruitment, baseline) and two rounds post-intervention (midline, endline). The timeline of each data collection round is presented in Table 1. All survey rounds were conducted as an *audit* at the business location by an independent enumerator, blind to the experiment design and treatment status of the firm. The recruitment survey contained questions on owner and business characteristics to be used to screen eligible retail firms into the sample and as controls in the main analysis. The baseline survey also contained such questions on business characteristics, but mainly focused on collecting key performance data (e.g., monthly sales) prior to any intervention. The midline survey took place 12 months post-recruitment for each firm and concentrated on our mechanism measures of store-level branding and product management. Finally, the endline survey took place roughly 24 months post-recruitment and closely mirrored the baseline survey. The same performance data on monthly sales was collected in an identical manner as the baseline survey, to examine the long-run impact of the two modernization interventions.

Enumerators were supervised in the field by team leaders and a research manager to ensure high data quality. The daily review process was as follows. Team leaders accompanied groups

of enumerators going to a particular geographic zone on a given field day. At the end of the field day, team leaders reviewed the surveys stored on the electronic tools of enumerators (to check for outliers, anomalies or data entry mistakes). The team leader and enumerators remained close to the businesses they had just audited and thus could re-visit businesses to fix any errors detected on the same day. Corrections had to be made in around 10% of cases and typos on the order of magnitude were most common. Finally, the research manager used statistical software to review key variables of all new surveys at the end the day, and if no inconsistencies were detected, they authorized enumerators to upload the survey to the server (which was subsequently accessed by the researchers). Typically, the research manager detected errors in 20 to 30 surveys per data collection round (less than 5% of surveys) which were more conceptual in nature. These three levels of rigorous checks (enumerator audit at business site, team leader audit on field, research manager audit prior to survey upload) were implemented during every data collection round.

Characteristics of Retail Businesses in Sample (N = 1148)

In Table 3, we display summary statistics on business owner characteristics. We see that 45% of business owners in our sample are female — indeed in the retail sector, women own and operate businesses at almost the same rates as men do. The typical business owner is 44 years old, and has completed high school or received higher education. In terms of business experience, 37.5% of owners have had some form of business education (e.g. a course, program, or training), 78% had a salaried job at some point in the past and 72% of owners were the original founders of the business that they registered for our study.

In Table 4, we present the business characteristics at baseline for the sample of businesses. We see that 88% of businesses sell goods as opposed to services. The data in this table also indicate that firms in our sample are not subsistence-level businesses (in line with our eligibility criteria of sampling retailers operating out of a permanent physical structure). The average business in our sample has 2.07 paid employees, owns assets valued at US $21,394, and has a monthly sales turnover of US $2,449. The 1148 firms in our sample collectively accounted for US $33.7 million in annual retail sales, thus representing the retail backbone of Mexico City. Monthly profits accruing to

the owner were US $502 on average, which can be bench-marked against the $843 monthly income of the median Mexican household (National Survey of Household Income and Expenditure, 2018). In terms of formalization, 30% of retailers in our sample previously received a loan from a formal institution and 61.9% are formally registered with a tax authority. Overall, our sample consists of fairly established retail firms.

Table 3: Summary Statistics on Owner Characteristics for N = 1148 Firms in Final Sample

Variable	Mean	St. Dev.	Min	Max
Age	44.246	11.659	19	84
Age under 26	0.039	0.194	0	1
Age 26 to 45	0.504	0.500	0	1
Age over 46	0.457	0.498	0	1
Highest Education Level (1 to 13)	6.094	1.671	2	13
No Schooling	0.002	0.044	0	1
Primary Schooling Only	0.032	0.177	0	1
Secondary Schooling Only	0.440	0.497	0	1
College Education	0.526	0.500	0	1
Gender (Male=1)	0.546	0.498	0	1
Married	0.528	0.499	0	1
Have Dependent Children	0.531	0.499	0	1
Past Salaried Job	0.780	0.415	0	1
Owned Another Business	0.340	0.474	0	1
Founded the Business	0.721	0.449	0	1
Prior Business Education	0.375	0.484	0	1

Notes: This table presents recruitment summary statistics on all businesses in the sample. The data was collected for N = 1148 businesses (prior to randomization) at the business location, from June to October 2017.

Randomization Checks

Next, we present evidence that randomization of firms into experimental groups was successful. In Table 5, Columns (1) through (3) show the means of each characteristic for the control group (N = 380), external treatment group (N = 385), and internal treatment group (N = 383), respectively. In Column (4), we report the p-value from ANOVA F-tests of equality of the three means. We find that across the 18 tests (for equality of three means), we reject only one of the null hypotheses of mean equality at the 5% level, which is what would be expected by random chance. Additionally, the F-

Table 4: Summary Statistics on Business Characteristics for N = 1148 Firms in Final Sample

Variable	Mean	St. Dev.	Min	Max
Sell Goods vs. Services (Goods=1)	0.877	0.329	0	1
Number of Employees: Paid	2.070	2.230	0	20
Number of Employees: Unpaid and Paid	2.379	2.305	0	20
Total Assets (USD)	21,394	54,276	0	1,029,000
Monthly Sales Estimate (USD)	2,449	3,747	0	51,282
Monthly Profit Estimate (USD)	502	856	-1,128	10,256
Registered with Tax ID	0.619	0.486	0	1
Registered with Municipality	0.472	0.499	0	1
Obtained Formal Loan	0.301	0.459	0	1
Separates Business and Personal Finances	0.465	0.499	0	1
Percent of Profits Saved	11.528	20.530	0	100
Management Practices Count (0 to 11)	5.398	2.442	0	11
Technological Practices Count (0 to 9)	2.778	2.389	0	9

Notes: This table presents baseline summary statistics on all businesses in the sample. The data was collected for N = 1148 businesses (prior to randomization) at the business location, from January to February 2018. Where monetary values in US Dollars are reported, the exchange rate used is 19.5 MXN Pesos to 1 US Dollar.

test for joint equality of balance variables is not significant for the relevant three group comparisons. The experimental groups were balanced on observable owner and business characteristics.

Sample Attrition

We also show that attrition (i.e., non-response from a firm during any data collection round) does not pose a threat to causal identification in our experiment. First, attrition from our sample was low. From the sample of N = 1148 retail businesses randomized at baseline, we were able to reach 93.7% of the firm owners 12 months later at midline (N = 1081 completed surveys, including non-operational firms) and 92.2% of firm owners 24 months later at endline (N = 1059 completed surveys, including non-operational firms). Second, attrition from our sample was not systematically related to treatment assignment. In Web Appendix C, Table 11 presents linear and probit regression analysis to check for differential attrition between our three experimental groups. Columns (1) to (2) examine midline attrition, while columns (3) to (4) examine endline attrition. We do not find any evidence for differential attrition occurring in any of the treatment groups relative to the control group. Moreover, to demonstrate that the post-attrition sample is not imbalanced, in Web

Table 5: Balance Checks for N = 1148 Firms Randomized at Baseline

	Control Mean	External Mean	Internal Mean	P-Value (F-Test)
Number of Employees: Unpaid and Paid	2.541	2.310	2.450	0.558
Total Assets (USD)	19,454	17,929	26,840	0.245
Weekly Customers (1 to 12)	4.252	4.411	4.134	0.461
Monthly Sales Estimate (USD)	2,669	2,313	2,366	0.371
Monthly Profits Estimate (USD)	519	461	527	0.489
Registered with Tax ID	0.637	0.624	0.596	0.505
Obtained Formal Loan	0.307	0.284	0.313	0.643
Management Practices Count (0 to 11)	5.488	5.362	5.345	0.678
Technological Practices (0 to 9)	2.955	2.531	2.850	0.033**
Age	44.114	44.922	43.679	0.358
Highest Education Level (1 to 13)	6.214	5.971	6.098	0.161
Gender (Male=1)	0.587	0.546	0.506	0.117
Married	0.522	0.543	0.518	0.786
Have Dependent Children	0.506	0.527	0.560	0.385
Past Salaried Job	0.795	0.779	0.765	0.599
Owned Another Business	0.352	0.330	0.337	0.817
Founded the Business	0.706	0.721	0.738	0.599
Prior Business Education	0.378	0.353	0.395	0.530
Joint F-Test (Control v External)				0.839
Joint F-Test (Control v Internal)				0.573
Joint F-Test (Internal v External)				0.498

Notes: This table presents balance checks for the full sample of firms based on pre-intervention data on business and owner characteristics. The first three columns present average values by experimental group. The fourth column presents the equality of means F-test. The value displayed is the p-value for this F test where the null hypothesis is equality of three group means. Statistically significant p-values are highlighted by: * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$

Appendix C Table 10, we present balance checks for the sample of N = 1059 firms that did not attrite at endline. Across the 18 tests on balance variables (for equality of three means), we do not reject any of the null hypotheses of mean equality at the 5% level. In other words, all experimental groups had statistically the same (low) rate of attrition and even after excluding attriting firms, they were balanced on observable characteristics.

Beyond these checks, we provide a rigorous "attrition sensitivity analysis" to show that our main results are robust to attrition considerations. Even if we make conservative assumptions that control group attriting firms were higher performing than treatment group attriting firms, we find positive and statistically significant impacts of our modernization interventions on sales performance. This analysis is detailed in the results section.

Non-Survival in our Sample

We found that a small proportion of the 1148 firms randomized at baseline had not survived by the time we carried out our follow-up data collection rounds. In the early development economics literature, non-survival was treated as a key performance outcome – interventions were designed to raise the probability of survivorship of subsistence-focused microenterprises (McKenzie & Woodruff 2014). More recently, as established growth-focused firms have become the unit of analysis in this literature (rather than subsistence-focused microenterprises), studies have focused on sales or profitability as key performance outcomes instead of survivorship (McKenzie 2020). This is well-aligned with disciplines that often study established firms, such as marketing and industrial organization. Our study follows this approach. As a deliberate strategy, we sampled established firms operating out of a permanent physical structure. This is the segment of traditional retail firms for whom the construct of physical modernization is applicable. Non-survival is low for this segment and comparable to small business survival rates in advanced markets such as the US. In fact, non-survival for these firms may not even imply business failure, but that better professional opportunities arose for the entrepreneur. Two years after recruitment, only 23.7% of business in our sample had not survived, which is somewhat *lower* than the corresponding 30% statistic for small businesses in the US (Bureau of Labor Statistics' Business Employment Dynamics 2017). For these reasons, we did not pre-register survival as an outcome of interest in our field experiment. Rather, our interventions were hypothesized to improve demand-based performance indicators such as sales, via growth in the firm's customer base and spending per customer.

Nevertheless, in Web Appendix C, we examine in detail how our modernization treatments impacted the survivorship of firms. We do so to check that non-survival does not threaten the validity of our experimental inferences. For example, if small firms in the control group were less likely to survive than those in the treatment group, we might not observe relevant counterfactual sales outcomes for treated firms across the full distribution of firm size. Table 12 of Web Appendix C presents linear and probit regression analysis to compare non-survival rates between the experimental groups. Across Columns (1) to (4), we do not detect any differential effect of treatment

assignment on business non-survival. The statistically significant correlates of non-survival were asset value (negatively correlated) and formal registration status (negatively correlated) prior to study participation. However, we caution against interpreting these variables as drivers of non-survival since they are also correlated with other firm characteristics, and they were not exogenously shifted.

Having demonstrated that non-survival is not systematically related to treatment, our main analysis presents results for the sample of firms that survived at each survey round. However, in a robustness check for our main sales effect, we show that including non-survivors in the analysis with zeros on sales outcomes yields the same pattern of results. This analysis is detailed in the results section. Next, we describe how we measure key variables.

Measurement of Key Variables

Measurement of retailer sales. To precisely capture changes in retail firm sales due to our modernization interventions, we obtain two distinct measures of monthly sales in all survey rounds: (i) an *aided-recall* measure where the firm owner reports their last complete month's sales purely from memory; and (ii) an *anchor-adjusted* measure where we iterate and triangulate through daily and weekly sales estimates with the firm owner, while referencing any available financial records, to arrive at an average month's sales. Testing multiple independent measures separately increases the probability of a false positive and combining outcomes improves statistical power to detect effects that go in the same direction (Drexler *et al.* 2014). Thus, our main monthly sales variable is the average of the aided-recall and anchor-adjusted sales measures, each winsorized 1% on both tails and converted to USD. Refer to Web Appendix D for more details on the measurement of monthly sales.

Measurement of store-level branding. We devise an original approach to measuring the mechanism of store-level branding, as prior studies have not measured branding for hundreds of small, independent firms. A critical challenge is that we cannot measure branding by surveying retail firm owners, as we do with all other key variables in our study, since branding is perceived by parties external to the firm. In the corporate branding literature, empirical studies have surveyed random subjects on famous national brands (such as Levis, Hallmark, and Kodak in J. L. Aaker

1997). We attempt to mirror this process by collecting photographs of the firms in our sample and showing them to random subjects who then provide brand ratings. We solicit ratings from subjects on well-established branding dimensions: four of Aaker's (1997) brand personality dimensions (brand excitement, sincerity, sophistication and competence) and five dimensions based on Keller's (1993) customer-based brand equity (trust in brand, willingness to pay, willingness to recommend, brand quality signals, and brand attractiveness)[7].

First, we collect one exterior and one interior photo of all consenting businesses in our sample roughly 12 months post-recruitment. Subsequently, we solicit independent raters, fully blind to study design, from Amazon's Mechanical Turk platform. Raters view a photo of the firm and provide a score (on a 5-point Likert Scale) on nine branding statements corresponding to the nine dimensions above, for example: *This store and its brand signal high quality*. For each photograph, we obtain these scores from 5 different raters. Their brand scores form the basis of our two main branding variables:

Store Brand Index. We average the scores given to the firm's photograph by an independent rater across the nine branding dimensions. Next, for interpretation, we construct this variable by normalizing the score so that it ranges continuously from 0 to 1.

Store Brand Dimension Count. We generate a "high value" dummy for each dimension of branding if the independent rater scores a firm's photograph above-median on that dimension. Next, we construct this variable by adding up the nine binary variables, so it ranges from 0 to 9.

As an alternative measure for robustness, we solicit brand ratings on similar branding dimensions from (up to three) actual customers of the businesses in our sample. Refer to Web Appendix D for details on the construction of customer-based measures of store-level brand. In practice, the challenge with customer-based branding measures was that we had to sample customers already purchasing from the store, whose type could differ systematically between treatment and control due to our interventions. As a result we rely predominantly on the photo-based brand measures —

[7]We exclude the brand personality dimension, brand ruggedness. In pre-tests, subjects viewed ruggedness as a negative trait (akin to run-down or shabby), which is not aligned with Aaker's (1997) conception of ruggedness: a trait describing brands like Harley Davidson.

from more objective raters — in our mechanism analysis.

Measurement of Product Management. We measure product management by the retailer in order to test our proposed mechanism through which internal modernization can increase sales. We measure product management by auditing whether the firm owner consistently performs this function during the Midline (12 months post-recruitment) and Endline (24 months post-recruitment) data collection rounds. First, we develop a comprehensive set of six processes (or dimensions) underlying product management, from extensive field interviews with retail firm owners in Mexico City as well as the marketing literature on retail product management (Varley 2014). These six dimensions include: updating product assortment by launching new products and retiring old unsuccessful products; improving quality and packaging of products in the current assortment; targeting sales efforts to high-margin and/or popular products; selecting an optimal set of suppliers given business needs; aligning quantities of products ordered with business needs; and optimizing time and budget when procuring products.

We implement a double-audit of these six dimensions of product management. During the on-site audit of each data collection round, the enumerator (blind to study design and treatment status) reviews each dimension of product management listed and asks the firm owner to provide concrete examples of focus on that dimension in the last six months. The enumerator transcribes the open-ended text response of the owner, noting concrete evidence or examples (e.g. I introduced a new soda called "Jarritos" in February).

Next, we employ a second auditor to scrutinize the text responses that were transcribed by the enumerator and assign each firm with a score on the six dimensions of product management, ranging from 1 (indicating no focus on the product management dimension) to 7 (indicating excellent focus on the product management dimension). This second auditor was trained by our research team to specialize in assessing firms' product management. They were also blind to our study design and treatment status of firms. Their scores form the basis of our main product management variables:

Product Management Index. We average the scores given to the firm by the auditor across the six product management dimensions. Subsequently, we construct this variable by normalizing the

score so that it ranges continuously from 0 to 1.

Product Management Dimension Count. We generate a "high value" dummy for each dimension of product management when the auditor rates a firm as above-median on that dimension. Next, this variable is constructed by adding up these six binary variables, so it ranges in integers from 0 to 6.

As an alternative measure for robustness, we asked enumerators to count the number of concrete actions or analyses taken by the owner on each dimension of product management (details in Web Appendix D) .

ANALYSIS AND RESULTS

In this section, we discuss the impact of our modernization interventions on three retail dependent variables: firm sales (main performance outcome), store-level branding (key mechanism for external treatment) and product management (key mechanism for internal treatment). Apart from providing model-free evidence, we present results from the Intent-to-Treat (ITT) regression specified in Equation 1:

$$Y_i = \alpha + \beta_{ext} ExternalTreat_i + \beta_{int} InternalTreat_i + x_i'\gamma + \delta Y_{i,base} + \varepsilon_i \quad (1)$$

Y_i is the dependent variable of interest (e.g. sales, store-level branding, product management capacity), for firm i. Our main explanatory variables are $ExternalTreat_i$ and $InternalTreat_i$ which are dummy variables indicating whether firm i was (randomly) assigned to the external or internal treatment, respectively. x_i represents a vector of control variables measured pre-intervention, including: 8 baseline controls for owner characteristics (gender; age; marital status; children status; education; prior salaried job; prior business ownership; prior business training), 8 baseline controls for business characteristics (founder; assets; weekly customers; employees; tax registration status; formal loans; profit savings rate; separation of business-personal affairs; business practices), 5 sub-sector fixed effects (set of two digit SIC codes) and 3 randomization strata dummies indicating which batch the firm was part of. We also control for the baseline value of the dependent variable,

$Y_{i,base}$. We report robust standard errors throughout.

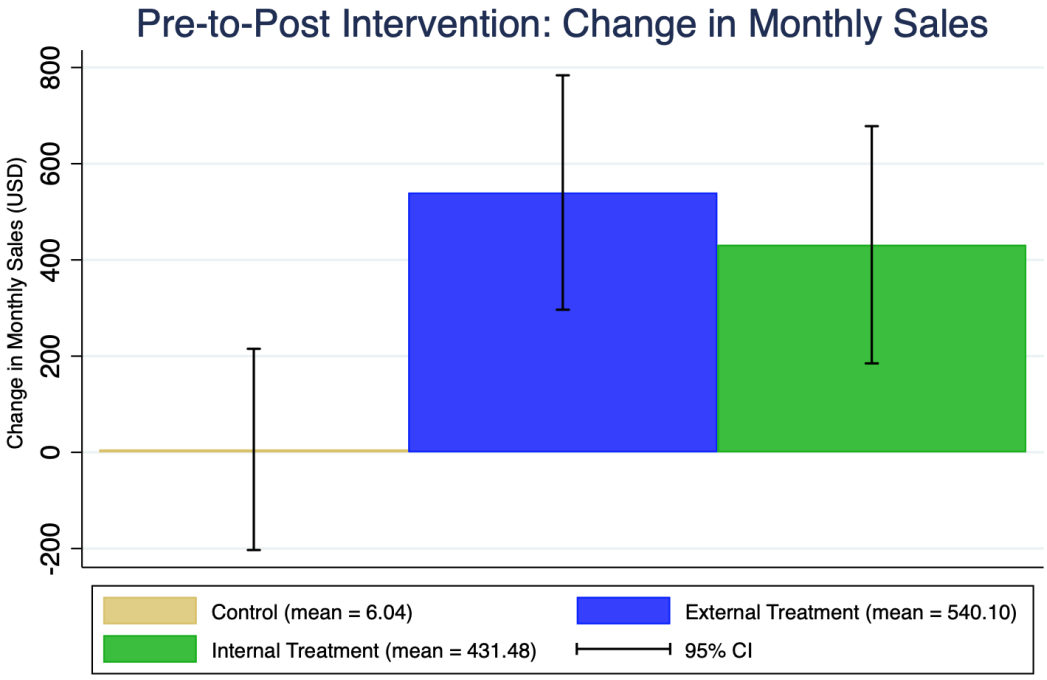
This specification confers numerous advantages, and is recommended in the development economics literature when there are multiple pre-and-post intervention data collection rounds (McKenzie 2012). First, by using the exogenous treatment assignment variables (rather than endogenous treatment compliance variables), we provide ITT estimates that are unbiased for the average treatment effect. Second, by including the value of the dependent variable at baseline, as well as control variables measured at baseline, we improve the precision of estimates and account for any group imbalances due to attrition or non-survival.

Impact of Modernization on Firm Sales

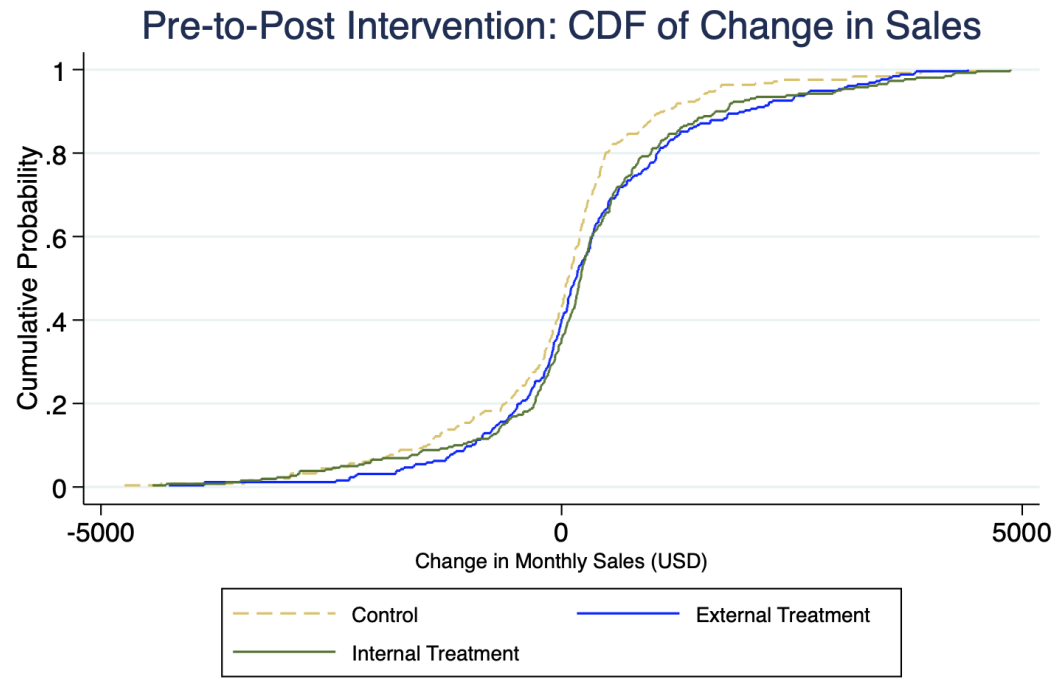
In the top panel of Figure 7, we provide initial model-free evidence that both modernization interventions improved the sales performance of retail firms. While monthly sales did not differ between groups at baseline, the average change in monthly sales from baseline to endline was significantly larger for external and internal treatment group firms than for control group firms. This improvement in average monthly sales from baseline to endline is not an artefact of outliers. As previously mentioned, all sales variables were winsorized 1% on both tails. Additionally, in the bottom panel of Figure 7, we plot the empirical CDF for the change in monthly sales, by treatment assignment. We show that the CDF for the external and internal treatment group is rightward shifted, indicating that across the distribution of firms, treatment group firms realized greater increases in sales (from baseline to endline) compared to control group firms.

This model-free evidence is supported by our ITT regression analysis. In Table 6, we report estimates of β_{ext} and β_{int} from Equation 1. The estimates indicate a positive and statistically significant effect of external and internal modernization treatments on monthly sales. We first interpret the estimates in Columns (1) to (3), where the dependent variable is the monetary monthly sales variable. Firms assigned to external treatment earned US $576.5 more than the control group at midline (12 months post-recruitment) and US $545.7 more than the control group at endline (24 months post-recruitment), which represents an average monthly sales improvement of 18.7% over the control group as in Column (3). Similarly, firms assigned to the internal treatment earned US

Figure 7: Model-free Evidence of Treatment Effects on Sales



Note 1: Monthly sales is the average of the aided-recall and anchor-adjusted monetary estimates, each winsorized 1% on both tails
Note 2: Means are significantly different for External vs. Control (p = 0.0012) and Internal vs. Control (p = 0.0101)



Note 1: Monthly sales is the average of the aided-recall and anchor-adjusted monetary estimates, each winsorized 1% on both tails
Note 2: CDF of pre-to-post intervention changes in monthly sales (range restricted from -5000 to 5000) for N = 787 firms operational at Endline

$466.3 more than control firms at midline (12 months post-recruitment) and US $476.0 more than control firms at endline (24 months post-recruitment), which represents an average monthly sales improvement of 15.5% over the control group as in Column (3)[8]. In Columns (4) through (6), we show that these results are robust to an alternative sales dependent variable: instead of winsorizing, we take the average of the log of aided-recall sales and anchor-adjusted sales. This specification handles outliers differently, but shows comparable effect sizes of the modernization treatments on sales.

These effect sizes are statistically significant, persistent at least up to 24 months post-recruitment and economically substantial. In order to shed light on economic magnitude, we subsequently contextualize the US $519 and US $430 improvement in monthly sales for external and internal treatment firms, respectively. At baseline, the average reported salary of a full-time employee was US $280 per month, and monthly rent for the shop premise was US $297. Therefore, the sales improvement due to our modernization interventions is equivalent to hiring roughly 1.5 more full-time employees or covering 1.5 months rent, both meaningful changes for small retail owners.

Robustness checks. We conduct multiple tests to ensure these reported sales effects are robust. The robustness tests for the main sales effect are summarized in Web Appendix F. First, in Table 14, we show that results are similar — in statistical significance and economic interpretation — when we alter the regression sample to include non-operational firms (coded as earning zeros for sales) or to exclude all firms except those who kept sales records at baseline. The latter test provides evidence that our results are not driven by treatment firms systematically reporting more accurate sales figures due to greater incidence of sales record-keeping post-treatment. Next, we show how our sales effects would change with different assumptions on sales growth for attriting firms in Table 15. We highlight in Column (4) of Table 15 that even under the most conservative attrition assumptions, where all control attritors are assigned the average sales growth of treatment group firms and all treatment attritors are assigned a sales growth of zero since baseline, we would still obtain a positive sales effect for the external and internal treatment group of 12.6% and 9.8%, respectively. Finally, we show that

[8]We show ATT effects of our modernization interventions in Table 13 of Web Appendix E, which is relevant to managers or policymakers allocating scarce resources to modernizing traditional retailers. They can identify firms in their distribution network or economy who will comply with modernization.

Table 6: Impact of Modernization Interventions on Firm Sales

	DV: Monthly Sales (USD)			DV: Log of Monthly Sales		
	(1)	(2)	(3)	(4)	(5)	(6)
	Midline	Endline	Midline and Endline Avg.	Midline	Endline	Midline and Endline Avg.
External Treat	576.5***	545.7***	518.9***	0.119*	0.198***	0.190***
	(222.7)	(183.2)	(163.8)	(0.0679)	(0.0745)	(0.0694)
Internal Treat	466.3**	476.0***	430.0**	0.175***	0.226***	0.237***
	(216.6)	(177.4)	(167.8)	(0.0639)	(0.0757)	(0.0703)
Baseline Value of DV	Yes	Yes	Yes	Yes	Yes	Yes
Sub-sector FE	Yes	Yes	Yes	Yes	Yes	Yes
Biz/Owner Controls	Yes	Yes	Yes	Yes	Yes	Yes
Strata FE	Yes	Yes	Yes	Yes	Yes	Yes
Mean of DV: Control	2439.5	2954.7	2776.8	10.22	10.33	10.27
SD of DV: Control	3161.3	3754.1	3409.5	1.084	1.378	1.305
Effect Size in SD: Ext	0.182	0.145	0.152	0.109	0.144	0.146
Effect Size in %: Ext	23.63	18.47	18.69	12.59	21.95	20.96
Effect Size in SD: Int	0.148	0.127	0.126	0.162	0.164	0.182
Effect Size in %: Int	19.12	16.11	15.48	19.14	25.35	26.75
Obs.	883	791	791	883	791	791
P-Value: $\beta\_ext = \beta\_int$	0.624	0.714	0.595	0.360	0.640	0.363

Notes: Data underlying these regressions were collected in two survey rounds – Midline (6 months post-intervention) and Endline (18 months post-intervention). The DV in Columns (1)-(2) is the monetary measure of monthly sales, i.e. the average of the recall-based and anchored-adjusted monetary sales estimates, both winsorized 1% on each tail. The DV in column (3) is the average of the DVs in (1) and (2), thus representing the average monthly sales in USD over the two survey rounds. The DV in Columns (4)-(5) is the log transformed measure of monthly sales, i.e. the average of the log of the recall-based and anchored-adjusted sales estimates. The DV in column (6) is the average of the DVs in (4) and (5), thus representing the average of the log of monthly sales over the two survey rounds. The indicated regressions include: the baseline value of the dependent variable, 8 baseline controls for owner characteristics (gender; age; marital status; children status; education; prior salaried job; prior business ownership; prior business training), 8 baseline controls for business characteristics (founder; assets; weekly customers; employees; tax registration status; formal loans; profit savings rate; separation of business-personal affairs; business practices), 5 sub-sector fixed effects (set of two digit SIC codes) and 3 strata dummies indicating which randomization/implementation batch the firm was part of. All regressions exclude firms that were found non-operational or attrited during the survey round. Robust standard errors are in parentheses. P-values are highlighted as: $p < 0.1^{*}$ $p < 0.05^{**}$ $p < 0.01^{***}$

our results do not depend on particular details of the empirical specification. In Table 16, we find statistically significant sales effects of similar magnitudes when excluding all control variables (in Columns 1 to 3) and when using a double-selection lasso to choose control variables (in Columns 4 to 6). The same applies for when we use panel data estimates of treatment effects in Table 17.

Spillover checks. In addition to these robustness tests, we provide evidence that the treatment effect on sales is not driven by sales losses for control group firms as they compete with more modernized treated firms. We use GPS data of the firms to investigate the possibility of competitive spillovers to the control group. First, our recruitment strategy ensured that firms were spread apart: the average distance between firms in our sample was 2.26 miles. Next, for control group firms, the *closest* treated firm in the same business line was 1.32 miles away (27 minutes by walking). Thus, the likelihood of competitive spillovers are minimal given we rarely sample firms in the same consideration set for consumers. Finally, sales increased more for control group firms with a nearby treated firm in their same business line relative to those that did not, though this difference was not significantly significant at conventional levels ($t = 1.23$). Control group firms who had a treated firm in their same business line (defined by 3-digit SIC codes) within 15 minutes walking distance saw average sales growth of US $158 over the two post-treatment survey rounds. The remaining control group firms saw an average sales decline of US $112. This pattern of results persists even when we compare similar control group firms by including 19 neighborhood fixed effects and the standard controls used in all our regression analyses[9] These checks suggest that our sales effect was not simply due to treated firms negatively impacting the sales of control firms.

In totality, our analyses of firm sales support Hypotheses 1 and 3. Additionally, we hypothesized that external modernization enhances the store-level brand of retailers (Hypothesis 2), while internal modernization improves product management (Hypothesis 4). We examine these mechanisms next.

[9]Due to brevity, this regression analysis is omitted from the paper, but can be provided upon request.

Effect of External Modernization on Store-Level Branding

In Columns (1) and (2) of Table 7, we show the ITT effects of external and internal modernization on our two photo-based store-level branding variables, with the sample of firms operational and not attriting at endline. N = 784 of these firms consented to provide photographs. Each observation in these regressions is at the photograph-rating level (5 independent ratings per photograph, 2 photographs per firm). Thus in all specifications, we include rater fixed effects for the 1425 raters in our sample and cluster standard errors at the level of the firm (the unit of randomization)[10]. Focusing on the Store Brand Index in Column (1) first, we see a positive and statistically significant effect of the external modernization treatment. Compared to control group firms, external treatment firms received a 6.2% (0.132 SD) higher normalized average score across the various branding dimensions. In Column (2), we show the ITT effects of modernization treatments on the Store Brand Dimensions Count. Firms in the control group on average performed highly on 3.29 out of 9 branding dimensions. In contrast, firms in the external treatment group performed highly on 3.72 dimensions on average, an improvement of 13.2% (0.144 SD). Again this effect is statistically significant at the 1% level.

A critical point to note is that we observe no improvement in branding scores for firms that were assigned to receive the internal modernization treatment. This helps us test whether it is truly the external modernization structures that assisted firm owners in strengthening their store-level brand. If branding improvements were driven by other generic aspects of the external treatment (such as the feeling of receiving a reward, receiving attention and assistance, a motivation boost, being observed by respected third-parties), we would expect to see the same improvements in branding among firms assigned to receive the internal treatment too. However, we do not observe this — coefficients for internal treatment are roughly a quarter of the size of coefficients on external treatment, while standard errors are similar. In support of Hypothesis 2, this analysis shows that externally-modernizing firms in Mexico City were able to build a stronger brand.

[10]We prefer this specification to one where observations are at the firm-level as estimates are more precise when we include standard controls for rating data (rater FEs, time taken to rate photo). Results are similar when observations are at the firm-level.

Table 7: Mechanism Analysis: Store-level Brand and Firm Sales

	Impact of Treatments on Store Branding		Correlation of Store Branding and Sales: OLS		Impact of Store Branding on Sales: IV	
	(1) Store Brand Index (Continuous: 0 to 1)	(2) Store Brand Dimensions (Count: 0 to 9)	(3) Post-Treat Sales (USD)	(4) Post-Treat Sales (USD)	(5) Post-Treat Sales (USD)	(6) Post-Treat Sales (USD)
External Treat	0.0330***	0.435***				
	(0.0123)	(0.145)				
Internal Treat	0.00871	0.189				
	(0.0122)	(0.141)				
Store Brand Index (Cont.: 0 to 1)			250.1**		9962.0***	
			(100.4)		(3107.3)	
Store Brand Dims. (Count: 0 to 9)				21.29**		865.9***
				(8.511)		(282.4)
Baseline Value of DV	N/A	N/A	Yes	Yes	Yes	Yes
Biz/Owner Controls	Yes	Yes	No	No	Yes	Yes
Sub-sector/Strata FE	Yes	Yes	Yes	Yes	Yes	Yes
Mean of DV: Control	0.532	3.288	2792.6	2792.6	2792.6	2792.6
SD of DV: Control	0.250	3.013	3451.8	3451.8	3451.8	3451.8
Effect Size in % (SD): Ext	6.2 (.132)	13.2 (.144)	—	—-	—	—
Effect Size in % (SD): Int	Not Sig.	Not Sig.	—	—	—	—
First Stage F Stat.	—	—	—	—	16.9	16.42
Obs.	7184	7184	7050	7050	7050	7050
No. of Firms	784	784	769	769	769	769

Notes: This table summarizes the relationship between our modernization interventions, store-level branding and retail firm sales. There are 2 photographs corresponding to every included firm, each rated by approximately 5 MTurk raters. The DV in Column (1) is the normalized average score given to a firm's photograph by an MTurk Rater on nine dimensions of store branding: attractiveness, willingness to reccommend, willingness to pay, trust in brand, high quality signals, brand excitement, brand sophistication, brand sincerity, and brand competence. The DV in Column (2) is a count of these nine dimensions of store branding on which the firm received an above-median score from the MTurk rater. In Columns (3) to (6) the DV is the average monthly sales in the two post-treatment data collection rounds. In the 2SLS regressions of Columns (5) and (6), store branding variables are instrumented by assignment to external treatment. The indicated regressions include: the baseline value of the dependent variable; 8 baseline controls for owner characteristics (gender; age; marital status; children status; education; prior salaried job; prior business ownership; prior business training), 8 baseline controls for business characteristics (founder; assets; weekly customers; employees; tax registration status; formal loans; profit savings rate; separation of business-personal affairs; business practices), 5 sub-sector fixed effects (set of two digit SIC codes) and 1425 MTurk Worker FE. All regressions exclude firms that were found non-operational or attrited during the Endline survey round. Clustered robust standard errors (by firm) are in parentheses. P-values are highlighted as: $p < 0.1^{*}$ $p < 0.05^{**}$ $p < 0.01^{***}$

We solicited similar brand ratings from actual customers of the businesses in our sample, to test the robustness of treatment effects estimated using the photo-based measures. In Table 18 of Web Appendix G, we show that the external treatment also had a positive and statistically significant effect on the Customer-based Brand Index. The effect size found is comparable to the effects found using the photo-based measures: firms in the external treatment group received a 4.8% (0.159 SD) higher score from their customers on branding dimensions relative to the control group. Again, no significant effect of treatment on branding was found for firms in the internal treatment.

To understand factors driving stronger brand ratings for firms in the external treatment group, we additionally analyze words provided by independent raters to describe the store-level brand of firms in our sample. For each photo in our sample, we asked the five independent raters to provide up to three keywords describing the brand of the business. In Table 18 of Web Appendix G (bottom panel), we summarize the ten most frequently-occurring words in our sample that had a positive valence (as per Warriner *et al.* 2013), presenting their average occurrence per photo by treatment group. Overall, raters more frequently described external treatment group firms using these positive words. Per photo, firms in the internal and control group received 2.9 positive descriptor words on their brand, while external treatment group firms received 3.6 positive descriptor words. This represents an improvement of 24% and it is statistically significant at the 1% level. In particular, external treatment group firms were more often described as clean, organized, friendly, professional, and modern, which could be key traits underlying their higher brand assessments.

Having demonstrated that external modernization positively impacts store-level branding, the next step of our mechanism analysis is to show that the enhanced store-level brand drives increases in sales. The key challenge is that we do not randomize the branding variable. Thus, we provide correlational evidence as well as an instrumental variable analysis relying on the exogenous variation in treatment assignment. In Columns (3) and (4) of Table 7, we regress post-treatment average monthly sales on the two store-level branding variables, respectively, while controlling for average monthly sales at baseline. We see a positive and statistically significant association: going from the minimum to maximum possible score on branding is associated with a US $250 jump in sales as per Column

(3), while having a high performance on all nine dimensions of branding is associated with a US $191 increase in monthly sales as per Column (4). These estimates do not have a causal interpretation, and so in Table 7 Columns (5) and (6), we conduct IV regressions of monthly sales on the two store-branding variables, instrumenting for them with assignment to external treatment. This instrument is relevant, with first-stage F Statistics of 16.90 and 16.42 in Columns (5) and (6) respectively.

The IV regression coefficients are larger than the OLS coefficients, and statistically significant at the 1% level. They suggest that going from the minimum to maximum possible score on branding drives a US $9960 increase in monthly sales, while moving from below-median to above-median on one dimension of branding results in a US $866 increase in monthly sales. These are upper-bound estimates as the exclusion restriction may not be met; the external treatment could affect sales through other channels. In the last subsection, we attempt to rule out some other channels such as price impressions (relative to competitors), owner confidence in making business improvements, time spent on making business improvements (vs. routine operations) and seriousness of business purpose, which we measured for exploratory analysis. Overall, this pattern of results suggests a positive relationship between store-level branding and firm sales, supporting Hypothesis 2 that external modernization drives improvements in sales by enhancing the store-level brand of traditional retailers.

Effect of Internal Modernization on Product Management

Next, we assess Hypothesis 4: that internal modernization improves sales by enhancing the product management capacity in retail firms. The first step here is to show treatment effects of our internal modernization intervention on product management. In Columns (1) and (2) of Table 8, we show the ITT effects of external and internal modernization treatments on our two product management variables, with the sample of firms operational and not attriting at endline. Focusing on the Product Management Index in Column (1) first, we see a positive and statistically significant effect of the internal modernization treatment. Compared to control group firms, internal treatment firms received a 13.3% (0.317 SD) higher normalized average score across the various product management dimensions. In Column (2), we show the ITT effects of modernization treatments on

the Product Management Dimension Count. Firms in the control group on average performed highly on 2.17 out of 6 product management dimensions. In contrast, firms in the internal treatment group performed highly on 2.77 dimensions on average, an improvement of 27.3% (0.335 SD). Again this effect is statistically significant at the 1% level. Analogous to our analysis on store-level branding, we find that the external treatment group did not significantly improve product management relative to the control group. This provides evidence that it is indeed internal modernization structures that assisted firm owners in strengthening their product management capacity, rather than any common aspects of both modernization interventions.

We solicited similar product management ratings from enumerators, to test the robustness of these treatment effects estimated. In Table 19 of Web Appendix H, we show that the internal treatment also had a positive and statistically significant effect on the Enumerator-based Product Management Index. The effect size found is comparable to the effects found using the auditor-based measures: firms in the internal treatment group were verified by enumerators as completing 8.7% (0.174 SD) more product management actions or analyses relative to the control group. In this table (bottom panel), we also summarize the ten frequently-occurring proactive product management words in our survey text responses, presenting their average usage per firm by treatment group. As expected, firm owners in the internal treatment group used more proactive product management words. Firm owners in the external and control group used 1.89 proactive product management words, while internal treatment group firm owners used 2.18 proactive product management words. This is an improvement of 13% which is statistically significant at the 5% level.

Next, we provide evidence that enhanced product management drives increases in sales. In Columns (3) and (4) of Table 8, we regress post-treatment average monthly sales on the two product management variables, respectively, while controlling for average monthly sales at baseline. We see a positive and statistically significant association: going from the minimum to maximum possible score on product management is associated with a US $1405 jump in sales as per Column (3), while having a high performance on all six dimensions of product management is associated with a US $810 increase in monthly sales as per Column (4).

Table 8: Mechanism Analysis: Product Management and Firm Sales

	Impact of Treatments on Prod. Management		Correlation of Prod. Management and Sales: OLS		Impact of Prod. Management on Sales: IV	
	(1) Prod. Mgmt. Index (Continuous: 0 to 1)	(2) Prod. Mgmt. Dims. (Count: 0 to 6)	(3) Post-Treat Sales (USD)	(4) Post-Treat Sales (USD)	(5) Post-Treat Sales (USD)	(6) Post-Treat Sales (USD)
External Treat	0.0246	0.214				
	(0.0150)	(0.145)				
Internal Treat	0.0584***	0.593***				
	(0.0161)	(0.151)				
Prod. Mgmt. Index (Cont.: 0 to 1)			1405.5***		7125.3**	
			(355.4)		(3126.5)	
Prod. Mgmt. Dims. (Count: 0 to 6)				135.6***		692.1**
				(36.24)		(301.7)
Baseline Value of DV	N/A	N/A	Yes	Yes	Yes	Yes
Biz/Owner Controls	Yes	Yes	No	No	Yes	Yes
Sub-sector/Strata FE	Yes	Yes	Yes	Yes	Yes	Yes
Mean of DV: Control	0.438	2.174	2776.8	2776.8	2776.8	2776.8
SD of DV: Control	0.185	1.767	3409.5	3409.5	3409.5	3409.5
Effect Size in % (SD): Ext	Not Sig.	Not Sig.	—	—	—	—
Effect Size in % (SD): Int	13.3 (0.317)	27.3 (0.335)	—	—	—	—
First Stage F Stat.	—	—	—	—	13.80	16.73
Obs.	793	793	781	781	781	781

Notes: This table summarizes the relationship between our modernization interventions, product management and retail firm sales. The DV in Column (1) is the normalized average score given to the firm by an auditor on six dimensions of product management: improving product assortment, improving product quality, sales effort on popular/profitable products, improving supplier selection, aligning stock orders with demand, and saving time and money in procurement. The DV in Column (2) is a count of these six dimensions of product management on which the firm received an above-median score from the auditor. In Columns (3) to (6) the DV is the average monthly sales in the two post-treatment data collection rounds. In the 2SLS regressions of Columns (5) and (6), product management variables are instrumented by assignment to internal treatment. The indicated regressions include: the baseline value of the dependent variable, 8 baseline controls for owner characteristics (gender; age; marital status; children status; education; prior salaried job; prior business ownership; prior business training), 8 baseline controls for business characteristics (founder; assets; weekly customers; employees; tax registration status; formal loans; profit savings rate; separation of business-personal affairs; business practices), 5 sub-sector fixed effects (set of two digit SIC codes) and 3 strata dummies indicating which randomization/implementation batch the firm was part of. All regressions exclude firms that were found non-operational or attrited during the Endline survey round. Robust standard errors are in parentheses. P-values are highlighted as: $p < 0.1^{*}$ $p < 0.05^{**}$ $p < 0.01^{***}$

Finally, in Table 8 Columns (5) and (6), we conduct IV regressions of monthly sales on the two product management variables, instrumenting for them with assignment to internal treatment. This instrument is relevant, with first-stage F Statistics of 13.80 and 16.73 in Columns (5) and (6) respectively. The IV regression coefficients are larger than the OLS coefficients, and statistically significant at the 1% level. They suggest that going from the minimum to maximum possible score on product management drives a US $7125 increase, while moving from below-median to above-median on one dimension of product management results in a US $692 increase in monthly sales. Thus, the results we obtain support Hypothesis 4 that internal modernization improves sales by enhancing the product management capacity of traditional retailers[11].

Ruling out Alternative Mechanisms

Our main mechanism variables of store-level branding and product management were pre-registered (as was our main outcome of sales), and the bulk of our data collection focused on measuring these three variables. To complement this analysis, we provide some exploratory evidence on alternative mechanisms. First, we examine whether customers reported paying higher prices (relative to competitors) at more modernized businesses. In our conceptual development, we suggested that external modernization may not have a positive impact on sales because customers might perceive visibly modernized firms as more expensive. In practice, we do find a positive impact of external modernization on firm sales, which suggests this price mechanism is not at play. We now provide direct evidence of this using price data collected from consumers.

Specifically, we asked customers surveyed at each business to report the price they paid for a single product that they just purchased. To ensure prices were comparable across firms selling different products, we asked customers to additionally report the price they would expect to pay for the same product if they could not purchase it at the focal firm but instead had to purchase it from other firms in their consideration set. Then, for each firm we constructed a price index, which is the ratio of price paid at the focal firm relative to the price a customer would have to pay to obtain

[11] Formal mediation analyses support our mechanism analysis. We omit these analyses for brevity.

the product elsewhere, averaged across customers and normalized. We report ITT effects of our modernization interventions on this variable. In Column (1) of Table 20 of Web Appendix I, we find that there are no treatment effects of external or internal modernization on the price index.

Additionally, we analyze whether our treatments impacted owner confidence in making business improvements, seriousness of business purpose, and time spent on making business improvements. These are common mechanisms considered in the development economics literature on small firm growth (e.g. in McKenzie & Sansone 2017; Fafchamps & Woodruff 2016). In Columns (2) to (4) of Table 20 of Web Appendix I, we show that there are no treatment effects of external or internal modernization on these alternative mechanisms. Overall, this supports our Hypotheses 2 and 4 that store-level branding and product management are key mechanisms driving sales gains here.

Exploring Returns to Modernization

The causal effects of modernization that we have estimated and presented so far confirm our hypotheses that both external and internal modernization can help retailers grow their sales. We therefore return to the puzzle of why so many retailers in emerging markets fail to modernize. Our post-intervention qualitative fieldwork in Mexico City, summarized in Table 9, reveals two key constraints: (i) retailers in our sample lack information on services to help them modernize; and (ii) they are uncertain on the returns to modernization (i.e., the associated costs relative to benefits). The first set of informational constraints can be tackled through targeted campaigns in traditional retail zones on the availability of business support services. Regarding the second set of constraints – uncertainty on returns to modernization – further exploratory analyses is provided in this subsection to guide policymakers, managerial stakeholders, and traditional retailers themselves.

First, we argue that implementing modernization programs like ours can yield a viable return on investment, through a cost-benefit analysis in Web Appendix J Table 21. In the top panel, we show that assignment to our modernization treatments led to a US $116 monthly increase in net income or profits, and these gains persisted even 24 months post-recruitment. This analysis utilizes the same ITT regression specification outlined previously and thus, these profit effects do have

Table 9: Why don't more retail firms utilize services (by companies, NGOs) to help them modernize?

	External Treatment	Internal Treatment
Pure Information Constraints		
Don't know that services to help them modernize exist	0.43	0.57
Don't know which areas of the business they need to modernize	**0.85**	**0.86**
Aware that services exist, but don't know how to find information on them	**1**	**1**
Trust Frictions		
Find it hard to judge quality of services to help them modernize	0.14	0.05
Don't trust that services will modernize them successfully	0.28	0.19
Don't think their employees have skills to implement modern practices	0.24	0.33
Search Frictions or Inertia		
Worried it will take too much time or hassle to find a service provider	0.28	0.29
Worried it will take too much time or hassle to receive support to modernize	0.33	0.47
Not interested in changing what they do in their business	0.05	0.33
Cost-Benefit Considerations		
Don't think modernization services will bring any benefits	0.14	0.24
Think modernization services could benefit them, but not by too much	0.38	0.25
Think modernization services will benefit them, but they are too costly	**0.95**	**0.95**

Notes: Numbers shown are the proportion of firms who strongly agree that this is one reason why small retailers do not utilize business support services to modernize. The top three reasons are highlighted in bold. Data was collected in November 2020 through structured interviews of N = 42 randomly-selected firms who were assigned to either treatment group (21 to the external treatment group, 21 to the internal treatment group). We thank the reviewers for suggesting this analysis.

a causal interpretation. Next, in the bottom panel, we estimate costs to a policy stakeholder of implementing a modernization program similar to ours. Under a low-cost scenario that we were able achieve through partnerships with NGOs and universities, the complete modernization program cost approximately US $675 per firm. This means it might take the typical retail owner approximately six months to start realizing a positive return on the modernization investment. In other settings, it may be difficult to achieve these low program costs, so we additionally estimate costs when a professional management consultancy implements the program. The complete program would then cost US $1365 per firm, and so positive returns on the investment would be realized after 12 months. While the individual retailer may find this investment worthwhile, policy stakeholders additionally may be concerned about the general equilibrium effects of modernization programs. Our study was not designed to explicitly estimate general equilibrium effects, yet we show that our

modernization interventions did not lead to sales losses for nearby control group firms in the same business sector. This is potentially reassuring to a policymaker given that retailers in our sample reported they largely compete with other traditional retailers in their vicinity.

Second, for the individual retailer, we provide exploratory analyses on how much a marginal modernization structure adds to firm sales and whether there are diminishing returns to the number of modernization structures implemented. We use data collected in the midline survey round where we measured whether a firm had implemented each of the 40 modernization structures outlined in Figure 2 and Figure 3. In Web Appendix J Table 22, we link this continuous measure — the count of modernization structures implemented out of a possible 40 — with firm sales. In Column (1), we find through a linear regression that each marginal modernization structure implemented is positively and significantly associated with a US $26 increase in monthly sales. In Columns (3) and (4), we see that both external and internal modernization structures have a similar relationship with sales.

Subsequently in Column (5), we relax the assumption of a linear relationship between modernization structures and sales. In particular, we test for threshold effects suggesting that the largest gains in sales performance occur when the retailer moves from being very traditional to having a handful of modern structures. There may be diminishing returns to further modernization beyond a certain minimum level. We therefore sort retailers into quartiles based on how many modernization structures they have implemented and regress monthly sales on the quartile membership. Predictably, we find the best sales performance among the top quartile of retailers (i.e., those with over 20 modernization structures). Their gain in sales relative to the bottom quartile (i.e., those with 0 to 3 modernization structures) is US $650 per month. However, the second lowest quartile of retailers (i.e., those with 4 to 10 modernization structures) *earn more in sales* relative to the bottom quartile (US $383), than retailers moving from the second quartile to the third quartile (incremental gain of US $36), and retailers moving from the second quartile to the top quartile (incremental gain of US $267). In other words, the greatest sales growth occurs when a retailer moves from having 0 to 3 modernization structures to having 4 to 10 modernization structures. This is the type of change we focused on achieving through our interventions, as evident in Table 2. There are diminishing returns

after 10 modernization structures are implemented, though further modernization is still beneficial.

Finally, in Web Appendix J Figure 15, we show which *types* of modernization structures have the strongest association with sales. We regress post-intervention monthly sales on the number of modernization structures implemented within a particular module. The coefficients from these regressions are plotted in Figure 15. We find that modernizing one's external appearance and customer engagement strategies matter the most as far as changes visible to customers are concerned. We also find that demand analysis, cashflow management, and systematic stock ordering are most critical when it comes to internal modernization. We again caveat this analysis by noting that selection of which modernization structures to implement are non-random. Overall, responding to retailers concerns on returns to modernizing, we provide analyses showing that investments in modernization equivalent to our program can be recovered within a relatively short period. We additionally show that there are significant returns to gain from initial attempts at modernizing, and that returns can vary according to type of modernization structure implemented.

CONCLUSION

In this study, we investigate the impact of modernization on the sales of traditional retailers through a rigorous randomized field experiment. Our analysis finds that externally-modernizing retailers significantly increase monthly sales by US $518 (19% improvement relative to the control group) over the two-year study period, while internally-modernized retailers also significantly increase monthly sales by US $430 (15% improvement relative to the control group). These effects are robust to alternative model specifications, different measures of the dependent variable, systematic measurement checks for treatment versus control retailers, and sensitivity tests on attriting retailers. Moreover, these sales gains occurred through different mechanisms for externally-modernizing retailers versus internally-modernizing retailers. Retailers in the external modernization treatment group significantly improved their store-level branding relative to the control group, as evaluated by independent raters and customers in Mexico City. Retailers in the internal modernization treatment group did not improve their store-level brand. Instead, they significantly improved their product

management relative to the control group, as evaluated by independent auditors of the businesses.

Implications for marketing research. In this study, we generate new, causal evidence to resolve the puzzle of whether it benefits traditional retailers to modernize in emerging markets. We show that remaining structurally different from modern retailers does not serve a strategic purpose, in spite of the types of customers traditional retailers serve and the resource-starved environment they operate in. This underscores the importance of studying barriers to modernization and future research can rigorously address whether these barriers are informational, financial, or behavioral.

This paper also takes a first look at store-level branding and product management capacity as theoretical mechanisms driving retailer performance. Scholars and practitioners alike have identified product management as an important function for (large, advanced-market) retailers to strengthen. Yet, no study prior to ours has exogenously varied internal systems and databases in any type of retail business then causally linked this variation to improved product management and sales, more than a year after the experimental intervention. Building further on the previous literature, we demonstrate that product management can be honed even in low-technology settings, and beneficially utilized by small shop owners lacking business knowledge specialists or the help of multiple employees. Similarly, while there is a rich literature in marketing on corporate branding that has relied on data from consumer surveys and financial markets, this is the first study to experimentally modernize the physical appearance of hundreds of retail businesses and link these changes to improved store-level brand perceptions as well as increases in sales more than a year later. We show that brand building can occur and be beneficial for small independent shops serving low-income consumers in emerging markets. One limitation of our study is that we do not randomly manipulate these mechanisms and hence cannot make causal conclusions about their impact. This is another avenue for future research.

Implications for policy. This research has additional implications for policy stakeholders who are motivated to help traditional retail firms grow their businesses, such as our collaborators at the World Bank and Mexico's Ministry of Finance. We propose business modernization as a new policy objective for improving retail sector outcomes. Many pre-existing policy programs (including an early pilot version of our program in Mexico) aim to *formalize* traditional retail businesses

by encouraging owners to register the business with a municipal or federal government authority. Yet, there is little evidence that retail firms benefit from formalization in terms of their economic performance (De Mel *et al.* 2013; McKenzie & Sakho 2010; De Andrade *et al.* 2013). In contrast, we find that modernization has a sizeable, positive impact on retailer sales. Another insight from our study, for policy stakeholders designing modernization programs, is that directly making tangible changes in a business is effective. Through our interventions, firms were able to adopt three to five new modernization structures, a substantial addition to the two they had in place prior to any intervention. Other training programs rarely have this impact, a typical change is a 5 to 10 percentage point improvement in business practices (McKenzie & Woodruff 2014).

Implications for marketing practice. Insights from our study can also be leveraged by managers at multinational firms that distribute their products through traditional retail channels. Many of these firms, such as the Bimbo Group in Mexico and Reliance Industries in India, have developed initiatives to modernize traditional retail businesses. Our results suggest some key mechanisms through which modernization improves firm performance, and this can be incorporated into the design of modernization initiatives. For instance, suppliers often enhance the appearance of traditional shops by providing them with materials that showcase famous product brands. We find that helping firms cultivate their *own store-level brand* can be fruitful too for increasing foot traffic to these businesses and expanding product sales. Additionally, as multinational firms attempt to change internal systems in traditional retailers, our research suggests that new systems introduced should provide clear product-level insights for the retailer to improve their product management. This can have positive upstream consequences for manufacturers too as it implies higher product quality at the point-of-sale and more efficient distribution interactions with traditional retailers.

Finally, this paper highlights opportunities that exist for future research on retail modernization, an understudied yet important area given the dominance of traditional firms in the retail sector of emerging markets and the number of livelihoods that are occupied in the practice of traditional retail. Many open research questions on retail modernization are relevant to policy and practice, including questions on how traditional retail firms can start to digitize their business processes.

References

1. Aaker, D. A. *Managing Brand Equity* (Simon and Schuster, 1991).
2. Aaker, J. L. Dimensions of Brand Personality. *Journal of Marketing Research* **34,** 347–356 (1997).
3. Abrams, C. India's Richest Man Battles Amazon, Walmart to Work With Mom-and-Pop Shops. *Wall Street Journal.* `https://www.wsj.com/articles/indias-richest-man-battles-amazon-walmart-to-work-with-mom-and-pop-shops-11570885202` (2019).
4. Anderson, S. J., Chandy, R. & Zia, B. Pathways to Profits: The Impact of Marketing vs. Finance Skills on Business Performance. *Management Science* **64,** 5559–5583 (2018).
5. Basker, E. The Causes and Consequences of Wal-Mart's Growth. *Journal of Economic Perspectives* **21,** 177–198 (2007).
6. Beaman, L., Magruder, J. & Robinson, J. Minding Small Change Among Small Firms in Kenya. *Journal of Development Economics* **108,** 69–86 (2014).
7. Bloom, N., Eifert, B., Mahajan, A., McKenzie, D. & Roberts, J. Does Management Matter? Evidence from India. *The Quarterly Journal of Economics* **128,** 1–51 (2013).
8. Bresciani, S. & Eppler, M. J. Brand New Ventures? Insights on Start-Ups' Branding Practices. *Journal of Product & Brand Management* (2010).
9. Bronnenberg, B. J. & Ellickson, P. B. Adolescence and the Path to Maturity in Global Retail. *Journal of Economic Perspectives* **29,** 113–34 (2015).
10. Bruhn, M., Karlan, D. & Schoar, A. The Impact of Consulting Services on Small and Medium Enterprises: Evidence From a Randomized Trial in Mexico. *Journal of Political Economy* **126,** 635–687 (2018).
11. Calderon, G., Cunha, J. M. & Giorgi, G. D. Business Literacy and Development: Evidence from a Randomized Controlled Trial in Rural Mexico. *Economic Development and Cultural Change* **68,** 507–540 (2020).
12. Corsten, D. & Gruen, T. Desperately Seeking Shelf Availability: An Examination of the Extent, the Causes, and the Efforts to Address Retail Out-of-Stocks. *International Journal of Retail & Distribution Management* (2003).
13. Cosse, T. J. & Swan, J. E. Strategic Marketing Planning by Product Managers—Room for Improvement? *Journal of Marketing* **47,** 92–102 (1983).
14. De Andrade, G. H., Bruhn, M. & McKenzie, D. *A Helping Hand or the Long Arm of the Law? Experimental Evidence on What Governments Can Do to Formalize Firms* (The World Bank, 2013).
15. De Mel, S., McKenzie, D. & Woodruff, C. The Demand for, and Consequences of, Formalization Among Informal Firms in Sri Lanka. *American Economic Journal: Applied Economics* **5,** 122–50 (2013).
16. De Mel, S., McKenzie, D. & Woodruff, C. Business Training and Female Enterprise Start-up, Growth, and Fynamics: Experimental Evidence from Sri Lanka. *Journal of Development Economics* **106,** 199–210 (2014).
17. Drexler, A., Fischer, G. & Schoar, A. Keeping it Simple: Financial Literacy and Rules of Thumb. *American Economic Journal: Applied Economics* **6,** 1–31 (2014).
18. Du, R. Y., Joo, M. & Wilbur, K. C. Advertising and Brand Attitudes: Evidence from 575 Brands Over Five Years. *Quantitative Marketing and Economics* **17,** 257–323 (2019).
19. Fafchamps, M. & Woodruff, C. *Identifying Gazelles: Expert Panels vs. Surveys as a Means to Identify Firms With Rapid Growth Potential* (The World Bank, 2016).
20. Farris, P. W., Bendle, N., Pfeifer, P. E. & Reibstein, D. *Marketing Metrics: The Definitive Guide to Measuring Marketing Performance* (Pearson Education, 2010).
21. Grewal, D., Krishnan, R., Baker, J. & Borin, N. A. The Effect of Store Name, Brand Name and Price Discounts on Consumers' Evaluations and Purchase Intentions. *Journal of Retailing* **74,** 331 (1998).
22. Hoch, S. J., Bradlow, E. T. & Wansink, B. The Variety of an Assortment. *Marketing Science* **18,** 527–546 (1999).

23. Kamakura, W. A. & Mazzon, J. A. Socioeconomic Status and Consumption in an Emerging Economy. *International Journal of Research in Marketing* **30,** 4–18 (2013).

24. Keller, K. L. Conceptualizing, Measuring, and Managing Customer-Based Brand Equity. *Journal of Marketing* **57,** 1–22 (1993).

25. Keller, K. L. & Lehmann, D. R. Brands and Branding: Research Findings and Future Priorities. *Marketing Science* **25,** 740–759 (2006).

26. Kim, W. G. & Kim, H.-B. Measuring Customer-Based Restaurant Brand Equity. *Cornell Hotel and Restaurant Administration Quarterly* **45,** 115–131 (2004).

27. Kohli, C. & LaBahn, D. W. Creating Effective Brand Names: A Study of the Naming Process. *Journal of Advertising Research* **37,** 67–75 (1997).

28. Madden, T. J., Fehle, F. R. & Fournier, S. *Brands Matter: An Empirical Investigation of Brand-Building Activities and the Creation of Shareholder Value* (Division of Research, Harvard Business School, 2002).

29. McKenzie, D. Beyond Baseline and Follow-Up: The Case for More T in Experiments. *Journal of Development Economics* **99,** 210–221 (2012).

30. McKenzie, D. Small Business Training to Improve Management Practices in Developing Countries: Reassessing the Evidence for "Training Doesn't Work" (2020).

31. McKenzie, D. & Sakho, Y. S. Does it Pay Firms to Register for Taxes? The Impact of Formality on Firm Profitability. *Journal of Development Economics* **91,** 15–24 (2010).

32. McKenzie, D. & Sansone, D. *Man vs. Machine in Predicting Successful Entrepreneurs: Evidence From a Business Plan Competition in Nigeria* (The World Bank, 2017).

33. McKenzie, D. & Woodruff, C. What are We Learning From Business Training and Entrepreneurship Evaluations Around the Developing World? *The World Bank Research Observer* **29,** 48–82 (2014).

34. Mitra, D. & Golder, P. N. How Does Objective Quality Affect Perceived Quality? Short-Term Effects, Long-Term Effects, and Asymmetries. *Marketing Science* **25,** 230–247 (2006).

35. Monroe Sullivan, D. & Shaw, S. C. Retail Gentrification and Race: The Case of Alberta Street in Portland, Oregon. *Urban Affairs Review* **47,** 413–432 (2011).

36. Morales, A. C. Giving Firms an "E" for Effort: Consumer Responses to High-Effort Firms. *Journal of Consumer Research* **31,** 806–812 (2005).

37. Narayan, V., Rao, V. R. & Sudhir, K. Early Adoption of Modern Grocery Retail in an Emerging Market: Evidence From India. *Marketing Science* **34,** 825–842 (2015).

38. Politis, D. The Process of Entrepreneurial Learning: A Conceptual Framework. *Entrepreneurship Theory and Practice* **29,** 399–424 (2005).

39. Simonson, A. & Schmitt, B. H. *Marketing Aesthetics: The Strategic Management of Brands, Identity, and Image* (Simon and Schuster, 1997).

40. Stein, A. & Ramaseshan, B. Towards the Identification of Customer Experience Touch Point Elements. *Journal of Retailing and Consumer Services* **30,** 8–19 (2016).

41. Sudhir, K. & Talukdar, D. The "Peter Pan syndrome" in Emerging Markets: The Productivity-Transparency Trade-off in IT adoption. *Marketing Science* **34,** 500–521 (2015).

42. Tyagi, R. K. & Sawhney, M. S. High-Performance Product Management: The Impact of Structure, Process, Competencies, and Role Definition. *Journal of Product Innovation Management* **27,** 83–96 (2010).

43. Varley, R. *Retail Product Management: Buying and Merchandising* (Routledge, 2014).

44. Verhoef, P. C. & Sloot, L. M. in *Retailing in the 21st Century* 239–253 (Springer, 2006).

45. Warriner, A. B., Kuperman, V. & Brysbaert, M. Norms of Valence, Arousal, and Dominance for 13,915 English Lemmas. *Behavior Research Methods* **45,** 1191–1207 (2013).

Web Appendix A: Recruitment Materials

Figure 8: Marketing Flyer Circulated to Recruit Firms

Pitch to Recruit Firms

Good morning, my name is <name> and I would like to help you improve your business through the free and specialized Mi Asesor program. This program has been developed by several institutions, such as Business School of Stanford University and the National Institute of Entrepreneurs (INADEM), and consists of provision of custom help to small companies in order for them to increase their sales, increase the number of their customers, and improve the management of their company. This is a unique opportunity for you and your business as you will benefit from customized help by specialized consultants. Throughout 13 sessions, they will work alongside with you. It is a free program with potential to improve the performance of your business. There are limited spaces, so I encourage you to apply as soon as possible.

Web Appendix B: Intervention Materials

Large Sign for Shop Exterior

Price Labels

Price Promotions

Exterior Stand with Menu

Walls Painted

Impulse Purchase Display

Figure 9: Six examples of modernization structures implemented during external intervention

Renovated Stock Area

System to Record Stock Flows

System to Record Product-level Sales

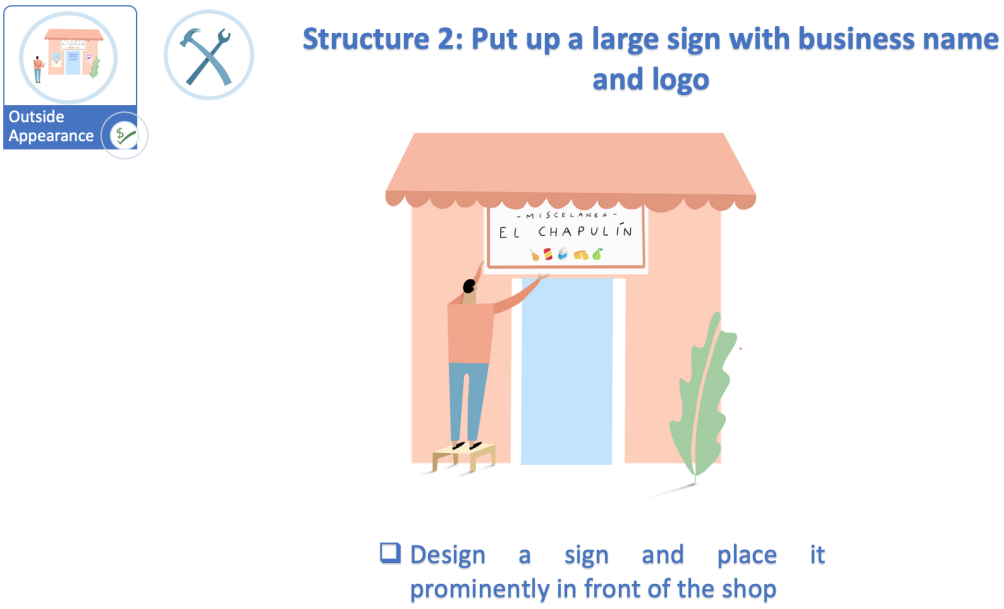
Database of Potential Suppliers

Listing of Margins of All Products

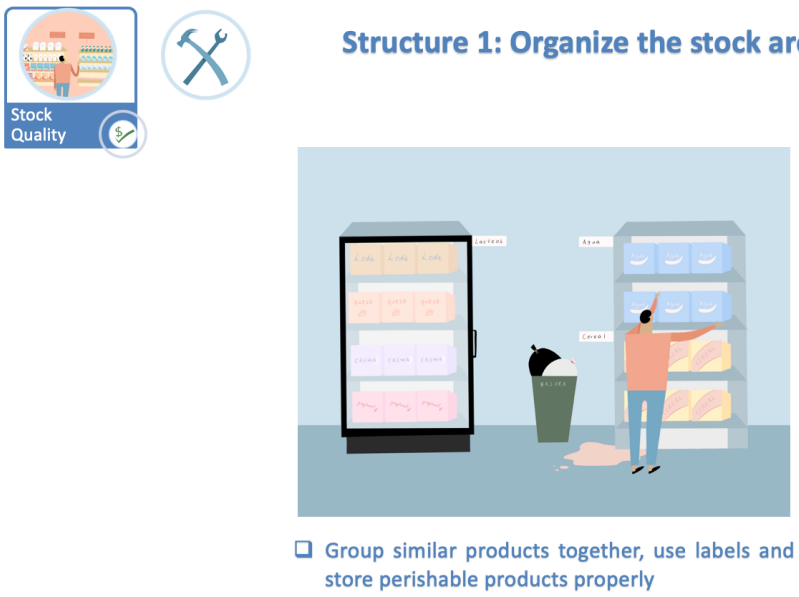
System to Record Expenses

Figure 10: Six examples of modernization structures implemented during internal intervention

Do-It-Yourself Intervention Manual



External Modernization Manual: Post Sign with Business Name and Logo



Internal Modernization Manual: Organize Stock Area

Figure 11: Excerpts from Do-It-Yourself Intervention Manuals

Business Report

Diagnóstico Sobre Tu Negocio

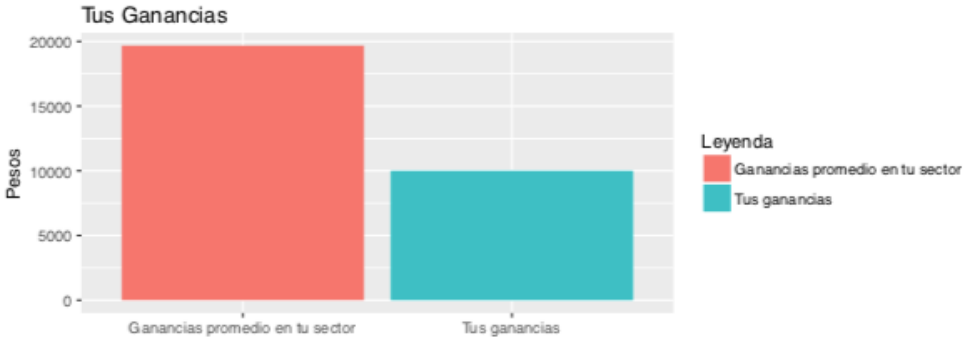
Mi Asesor

13/10/2017

Nombre del negocio: [redacted]

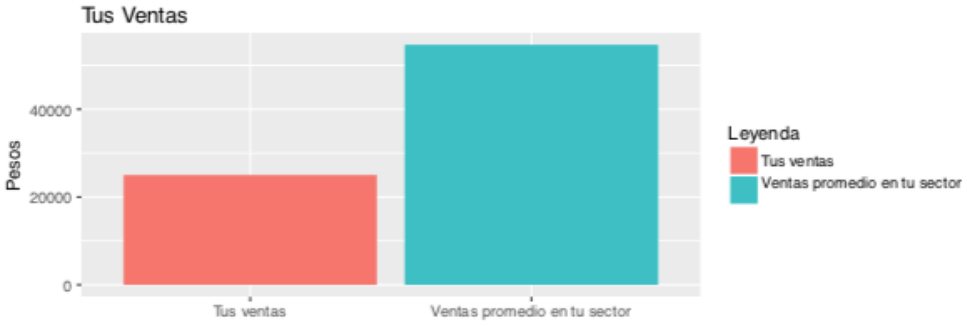
Resumen de tus ganancias

¡En la figura que se muestra a continuación puedes ver tus ganancias en el mes! Además, se muestran las ganancias promedio de los otros negocios de tu sector que también son parte de Mi Asesor.



Resumen tus ventas

¡En la figura que se muestra a continuación puedes ver tus ventas en el mes! Además, se muestran las ventas promedio de los otros negocios de tu sector que también son parte de Mi Asesor.



Resumen de tus prácticas de negocio

A continuación, te mostramos informaión sobre las prácticas que llevas a cabo en tu negocio. Además, puedes ver qué tipo de prácticas implementan los otros negocios de tu sector que también son parte de Mi Asesor.

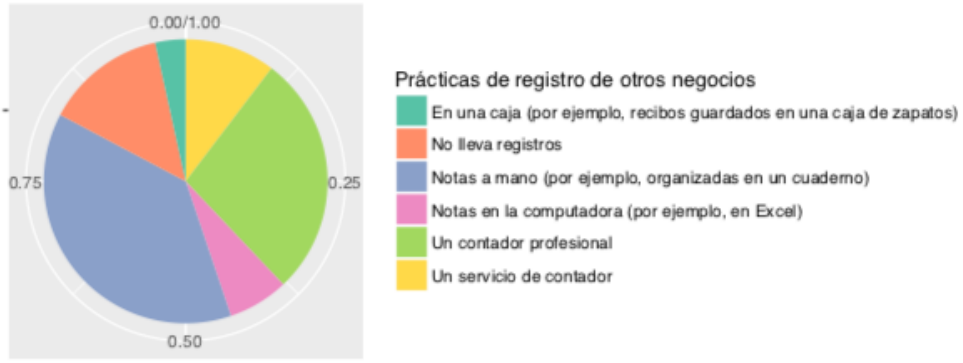
1

Figure 12: Page 1 of Business Report Provided as "Light" Treatment to Control Group

Seguimiento a las finanzas de la empresa

¿Utilizas algún método para administrar las finanzas del negocio?: Notas a mano (por ejemplo, organizadas en un cuaderno).

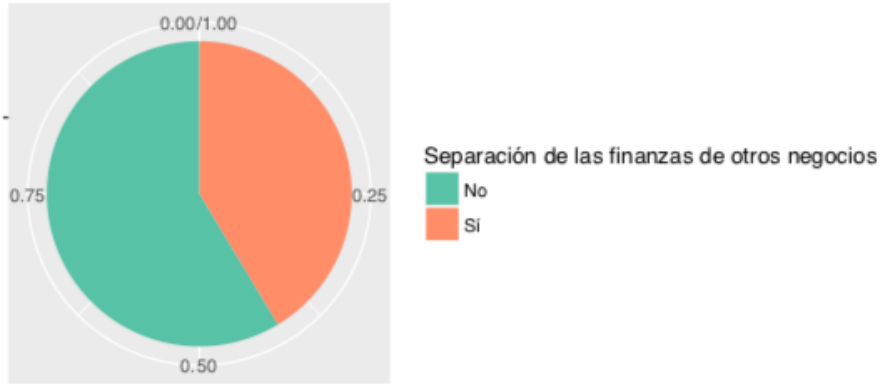
Otros negocios de de tu sector hacen lo siguiente:



Separación de las finanzas del negocio y las personales

¿Separas tus finanzas del negocio de las personales?: No.

Otros negocios de tu sector hacen lo siguiente:



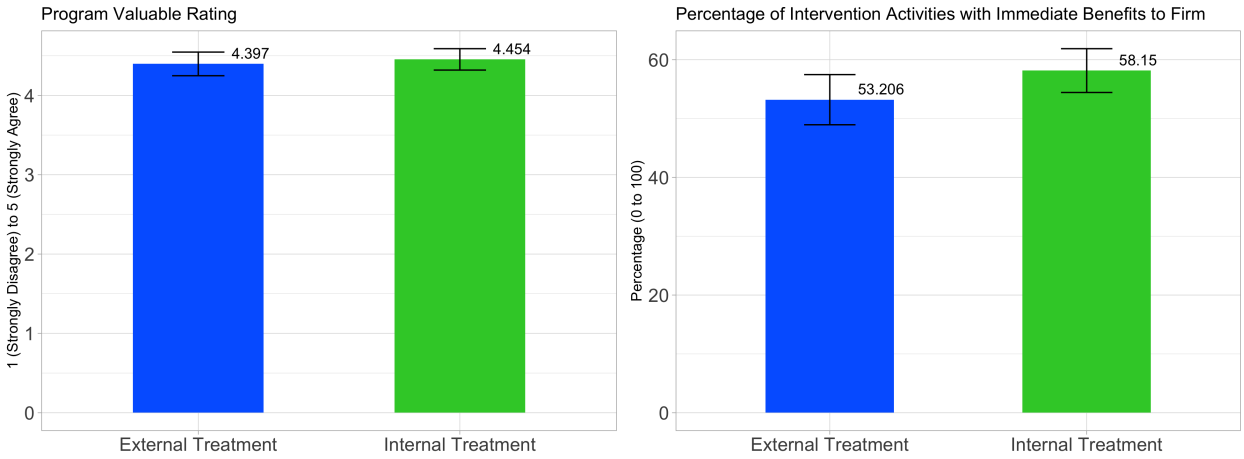
Nuestra recomendación:

¡Mejora al AUMENTAR tus ventas y REDUCIR tus costos!

Figure 13: Page 2 of Business Report Provided as "Light" Treatment to Control Group

Business Owner Satisfaction with Intervention

Figure 14: Firm Satisfaction with Assigned Intervention



Web Appendix C: Attrition and Non-Survival Checks

Table 10: Balance Checks for N = 1059 Firms Responding at Endline

	Control Mean	External Mean	Internal Mean	ANOVA P-Value
Number of Employees: Unpaid and Paid	2.471	2.326	2.451	0.782
Total Assets (USD)	19,024	17,692	21,840	0.334
Weekly Customers (1 to 12)	4.109	4.361	4.040	0.322
Monthly Sales Estimate (USD)	2,626	2,346	2,169	0.323
Monthly Profits Estimate (USD)	529	461	510	0.550
Registered with Tax ID	0.608	0.616	0.577	0.538
Obtained Formal Loan	0.297	0.293	0.325	0.613
Management Practices Count (0 to 11)	5.434	5.385	5.327	0.848
Technological Practices (0 to 9)	2.945	2.557	2.873	0.070*
Age	40.346	41.342	40.192	0.563
Highest Education Level (1 to 13)	5.714	5.517	5.626	0.522
Gender (Male=1)	0.534	0.499	0.462	0.167
Married	0.491	0.501	0.471	0.717
Have Dependent Children	0.411	0.444	0.460	0.423
Past Salaried Job	0.729	0.723	0.704	0.750
Owned Another Business	0.325	0.310	0.296	0.697
Founded the Business	0.654	0.652	0.684	0.609
Prior Business Training	0.343	0.326	0.356	0.694

Notes: This table presents balance checks for the sample of firms that responded to our survey 24 months post-recruitment, using their baseline data of business and owner characteristics. The first three columns present average values by experimental group. The fourth column presents the equality of means F-test. The value displayed is the p-value for this F test where the null hypothesis is equality of three group means. Statistically significant p-values are highlighted by: * p<0.10; ** p<0.05; *** p < 0.01

Table 11: Attrition at Each Follow-Up Round, by Treatment Assignment

	DV: Attrition at Midline (Yes = 1)		DV: Attrition at Endline (Yes =1)	
	(1) OLS	(2) Probit	(3) OLS	(4) Probit
External Treat	-0.0181	-0.1947	-0.0232	-0.1667
	(0.0158)	(0.1534)	(0.0189)	(0.1336)
Internal Treat	0.0141	0.1338	-0.00128	0.0075
	(0.0181)	(0.1423)	(0.0202)	(0.1279)
Sub-sector FE	Yes	Yes	Yes	Yes
Biz/Owner Controls	Yes	Yes	Yes	Yes
Strata FE	Yes	Yes	Yes	Yes
Mean of DV: Control	0.0605	0.0605	0.0868	0.0868
Obs.	1148	1148	1148	1148

Notes: This table analyzes attrition status by treatment assignment. Data underlying these regressions were collected in two survey rounds – Midline (12 months post-recruitment) and Endline (24 months post-recruitment). The DV in Columns (1)-(2) is a binary indicator coded '0' if the firm responded to the midline survey (data obtained or non-operational status confirmed) and '1' for attriter if the firm did not respond at midline (no data obtained and could not reach in any way to confirm operating status). The DV in Columns (3)-(4) is a binary indicator coded '0' if the firm responded to the endline survey and '1' for attriter if the firm did not respond at endline.The indicated regressions include: 8 baseline controls for owner characteristics (gender; age; marital status; children status; education; prior salaried job; prior business ownership; prior business training), 8 baseline controls for business characteristics (founder; assets; weekly customers; employees; tax registration status; formal loans; profit savings rate; separation of business-personal affairs; business practices), 5 sub-sector fixed effects (set of two digit SIC codes) and 3 strata dummies indicating which randomization/implementation batch the firm was part of. Robust standard errors are in parentheses. P-values are highlighted as: $p < 0.1^{*}$ $p < 0.05^{**}$ $p < 0.01^{***}$

Table 12: Intent-to-Treat Analysis of Non-Survival at Each Follow-Up Round

	DV: Non-Operational at Midline (Yes = 1)		DV: Non-Operational at Endline (Yes = 1)	
	(1) Probit	(2) OLS	(3) Probit	(4) OLS
External Treat	0.0632	0.0198	-0.0229	-0.00175
	(0.111)	(0.0279)	(0.106)	(0.0314)
Internal Treat	-0.0938	-0.0197	-0.0986	-0.0211
	(0.117)	(0.0279)	(0.107)	(0.0311)
Sub-sector FE	Yes	Yes	Yes	Yes
Biz/Owner Controls	Yes	Yes	Yes	Yes
Strata FE	Yes	Yes	Yes	Yes
Mean of DV: Control	0.171	0.171	0.251	0.251
Effect Size: Ext	Not Sig.	Not Sig.	Not Sig.	Not Sig.
Effect Size: Int	Not Sig.	Not Sig.	Not Sig.	Not Sig.
Obs.	1081	1081	1059	1059

Notes: Data underlying these regressions were collected in two survey rounds – Midline (6 months post-intervention) and Endline (18 months post-intervention). The DV in Columns (1)-(2) is a binary indicator coded '0' if the firm was operational at the time the midline survey was conducted and '1' if the firm had failed at midline (non-operational with zero monthly sales). The DV in Columns (3)-(4) is a binary indicator coded '0' if the firm was operational at the time the endline survey was conducted and '1' if the firm had failed at endline (non-operational with zero monthly sales). The indicated regressions include, where possible: 8 baseline controls for owner characteristics (gender; age; marital status; children status; education; prior salaried job; prior business ownership; prior business training), 8 baseline controls for business characteristics (founder; assets; weekly customers; employees; tax registration status; formal loans; profit savings rate; separation of business-personal affairs; business practices), 5 sub-sector fixed effects (set of two digit SIC codes) and 3 strata dummies indicating which randomization/implementation batch the firm was part of. Robust standard errors are in parentheses. P-values are highlighted as: $p < 0.1^{*}$ $p < 0.05^{**}$ $p < 0.01^{***}$

Web Appendix D: Measurement of Key Variables

Firm Sales

FIRM SALES: MAIN EFFECT			
Measure	**Variables**	**Questions and Steps**	**Responses**
Firm Sales Estimate #1	Aided-Recall Estimate	Q1: How much did you sell in the LAST MONTH (past 30 days)? This is the total amount of money that you got in sales. It includes all the money you bring IN to your business BEFORE you pay any bills, expenses, or salaries. Please give your first guess.	Numeric Response
Firm Sales Estimate #2	Averaged estimate (based on a typical week)	Q2A: Think of the BEST WEEK in your business last month. This is the week where you sold the most. How much money came into your business that week?	Numeric Response
		Q2B: Think of the WORST WEEK in your business last month. This is the week where you sold the least. How much money came into your business that week?	Numeric Response
		Q2C: Based on what you just told me about the total sales in your Best Week and in your Worst Week, we can estimate your sales to be approximately: [average(2A, 2B) * 4.25] for LAST MONTH (the past 30 days). WE GOT THIS SECOND GUESS FOR YOUR MONTHLY SALES BY TAKING THE AVERAGE OF YOUR BEST AND WORST WEEKS, THEN MULTIPLYING IT BY THE NUMBER OF WEEKS IN A MONTH. Does this amount seem correct to you?	0 = No 1 = Yes
Firm Sales Estimate #3	Averaged estimate (based on a typical day)	Q3A: In a typical week, how many DAYS is your business OPEN?	Numeric Response
		Q3B: What were your total sales YESTERDAY?Remember to include all the money you brought IN to your business yesterday before you paid any bills, expenses, or salaries.	Numeric Response
		Q3C: Think of the BEST DAY in your business last month. This is the day where you sold the most. How much were your sales on your best day last month?	Numeric Response
		Q3D: Think of the WORST DAY in your business last month. This is the day where you sold the least. How much were your sales on your worst day last month?	Numeric Response
		Q3E: Let me confirm what you have told me about your DAILY SALES during the last month: Sales on your Best Day (highest daily sales) = [3C]; Sales total for Yesterday = [3B]; Sales on your Worst Day (lowest daily sales) = [3D] NOW THINK ABOUT A TYPICAL DAY FOR YOUR BUSINESS LAST MONTH. USING THESE 3 NUMBERS TO GUIDE YOU, WHAT IS YOUR BEST GUESS OF HOW MUCH MONEY YOU USUALLY GET FROM SELLING YOUR PRODUCTS/SERVICES TO CUSTOMERS ON AN AVERAGE OR TYPICAL DAY?	Numeric Response
		Q3F: Based on what you just told me about the total sales on your Typical Day, we can estimate your sales to be approximately: [(3A) * (3E) * 4.25] for LAST MONTH (the past 30 days). WE GOT THIS THIRD GUESS FOR YOUR MONTHLY SALES BY TAKING YOUR TYPICAL DAY AND MULTIPLYING IT BY THE NUMBER OF DAYS YOU'RE OPEN PER WEEK, THEN MULTIPLYING IT BY 4 WEEKS IN A MONTH. Does this amount seem correct to you?	0 = No 1 = Yes
Firm Sales Estimate #3	Anchor-Adjusted Estimate (based on all previous estimates and/or official accounting records)	Q4A: HERE'S WHAT YOU TOLD ME ABOUT YOUR SALES FOR THE PAST 30 DAYS: Guess #3 (Based on Typical Daily Sales) = [3E]; Guess #2 (Based on Best and Worst Weeks) = [2C]; Guess #1 (Based on First Guess) = [1] USING THESE NUMBERS TO GUIDE YOU, WHAT IS YOUR FINAL 'BEST' GUESS FOR THE TOTAL SALES LAST MONTH? We mean how much money did you get from selling your products/services to customers last month (past 30 days) before you paid any bills, expenses, or salaries?	Numeric Response
		Q4B: So your TOTAL SALES LAST MONTH (the past 30 days) is: [4A]? Does this amount seem correct to you?	0 = No 1 = Yes
		(Q4C: Enumerator Check) Once the business owner finishes providing the sales estimates, the enumerator checks: Is the last month's sales figure - [display value from 4A] - above zero? The survey tool does not proceed if the enumerator does not confirm that the sales figure is above zero.	0 = No 1 = Yes
		(Q4D: Enumerator Check) Once the business owner finishes providing the sales estimates, the enumerator checks: Is the last month's sales figure - [display value from 4A] - below 200,000 MXN Pesos? If the answer is 'No,' the enumerator discusses all the estimates again in Q1-Q3 with the business owner to ensure that the high value of the sales figure provided is not a mistake.	0 = No 1 = Yes

Store-level Brand

STORE-LEVEL BRAND: MECHANISM FOR EXTERNAL MODERNIZATION TREATMENT			
Measure	**Variables**	**Questions/Steps for Enumerator Taking Photograph**	**Responses**
Photographs	Photograph Instruction	In the data collection round, enumerators are instructed to take standardized photos of the business. We provide the instructions given to the enumerator: Take a picture of the exterior of the store, facing it from the MIDDLE. You should walk about 15 metres out from the store entrace to take this picture. Walk more if you cannot capture the ENTIRE STOREFRONT in your picture. Take a picture of the exterior of the store, facing it from the MIDDLE. You should walk about 15 metres out from the store entrace to take this picture. Walk more if you cannot capture the ENTIRE STOREFRONT in your picture.	Photo upload
		Take a picture of the interior of the store, facing it from the entrance. Take two steps in from the entrance and do not change your angle to take this picture.	Photo upload
		Questions/Steps for Independent Rater Rating Firm's Store-level Brand	
Store-level Brand	Rater Instruction	(Intro): In this task, we are going to show you an image of a small businesses in Mexico and ask you to rate it. There are a range of businesses in the market. They vary on quality, branding, price and products. Below are some examples of small businesses in Mexico [Business #1: Internet Café, Business #2: Grocery store, Business #3: Restaurant]. When you answer the questions, please think about the Mexican business landscape, rather than any other context. Now that you know what kind of small businesses we are considering, we will move on to the task at hand. There are no risks involved in completing the questions. Your answers will remain confidential and well-protected. While answering the questions, look closely at the image displayed. You should be as honest and accurate with your answers as possible. Do NOT base your answers on the quality of the images or the photographic skills that went into taking the images. Look at the image of the business below: [Insert Image]	
	Brand Personality: Exciting	Q1: This store and its brand are EXCITING (e.g. daring, spirited, imaginative, modern).	1 = Strongly Disagree..to.. 5 = Strongly Agree
	Brand Personality: Sophisticated	Q2: This store and its brand are SOPHISTICATED (e.g. classy, exclusive)	1 = Strongly Disagree..to.. 5 = Strongly Agree
	Brand Personality: Sincere	Q3: This store and its brand are SINCERE (e.g. down to earth, honest)	1 = Strongly Disagree..to.. 5 = Strongly Agree
	Brand Personality: Competent	Q4: This store and its brand are COMPETENT (e.g. reliable, successful in meeting needs)	1 = Strongly Disagree..to.. 5 = Strongly Agree
	Brand Strength: Attractive	Q5: This store looks appealing and attractive to shop at	1 = Strongly Disagree..to.. 5 = Strongly Agree
	Brand Strength: Recommend	Q6: I would recommend people I know to shop at this store	1 = Strongly Disagree..to.. 5 = Strongly Agree
	Brand Strength: Quality	Q7: This store and its brand signal high quality	1 = Strongly Disagree..to.. 5 = Strongly Agree
	Brand Strength: Trust	Q8: I would trust the business associated with this store and brand	1 = Strongly Disagree..to.. 5 = Strongly Agree
	Brand Strength: WTP	Q9: I would be willing to shop at this store even if its prices are high	1 = Strongly Disagree..to.. 5 = Strongly Agree
	Unstructured Brand Description	Q10: What are some words you would use to describe this store and brand? Provide up to three keywords.	Text Response

Product Management

PRODUCT MANAGEMENT: MECHANISM FOR INTERNAL MODERNIZATION TREATMENT			
Measure	**Variables**	**Questions and Steps**	**Responses**
Product Management	Improving Product Assortment	Q1A: Over the last 6 months, have you carried out any analysis or actions to introduce a new product or product(s) in your business? By this we mean, have you used your resources (time and money) to add a new product or products(s) to your product assortment?	Text Response
		(Q1B: This queston is completed privately by enumerator) Over the last 6 months, how many actions did you take to introduce a new product or product(s) in your business?	0 = Never took any deliberate actions or analysis ..to.. 5 = Took deliberate actions or analyses 5 (or more) different times
		Q1C: Over the last 6 months, have you carried out any analysis or actions to remove a product or product(s) from your business offerings? By this we mean, have you planned to remove a product or products(s) from assortment that you were previously offering?	Text Response
		(Q1D: This queston is completed privately by enumerator) Over the last 6 months, how many actions did you take to remove a product or product(s) from your assortment in your business?	0 = Never took any deliberate actions or analysis ..to.. 5 = Took deliberate actions or analyses 5 (or more) different times
		(Q1E: This queston is completed privately by the independent product management auditor) Assess how focused the respondent is on deciding the product assortment in their business. How much effort do you think the respondent has dedicated to trying to choose the best possible set of products to offer to the market, removing suboptimal products from the assortment and adding products strategically to their assortment?	1 = Not focused on this dimension of product management at all ..to.. 7 = Extremely focused on this dimension of product management
	Improving Product Quality	Q2A: Over the last 6 months, have you carried out any analysis or actions to improve the quality of product or product(s) in your business? By this we mean, have you used your resources (time and money) to offer improved versions of the existing product or products(s) in your assortment?	Text Response
		(Q2B: This queston is completed privately by enumerator) Over the last 6 months, how many actions did you take to introduce improvements to product or product(s) in your business?	0 = Never took any deliberate actions or analysis ..to.. 5 = Took deliberate actions or analyses 5 (or more) different times
		(Q2C: This queston is completed privately by the independent product management auditor) Assess how focused the respondent is on improving the products offered in their business. How much effort do you think the respondent has dedicated to ensuring the products they offer are of high quality, well-packaged and presented?	1 = Not focused on this dimension of product management at all ..to.. 7 = Extremely focused on this dimension of product management
	Sales Effort on Popular/Profitable Products	Q3A: Over the last six months, have you carried out any analysis or actions to identify high-margin products in your business and sell them in greater quantities? In other words, have you tried to figure out which products earn you the most margins, and then taken deliberate and specific actions to sell more of those products?	Text Response
		(Q3B: This queston is completed privately by enumerator) Over the last 6 months, how many actions did you take to sell a greater quantity of high-margin product(s) in your business?	0 = Never took any deliberate actions or analysis ..to.. 5 = Took deliberate actions or analyses 5 (or more) different times
		Q3C: Over the last six months, have you taken actions to change the volume of sales of your two most popular products? In other words, have you tried to sell a higher (or lower) quantity of your two most popular products, for example through promoting the product, advertising it more/less or changing it's price?	Text Response
		(Q3D: This queston is completed privately by enumerator) Over the last 6 months, how many actions did you take to sell a greater quantity of popular product(s) in your business?	0 = Never took any deliberate actions or analysis ..to.. 5 = Took deliberate actions or analyses 5 (or more) different times
		(Q3E: This queston is completed privately by the independent product management auditor) Assess how focused the respondent is on managing margins in their business. How much effort do you think the respondent has applied to pushing products that earn the highest profit margins or have high demand?	1 = Not focused on this dimension of product management at all ..to.. 7 = Extremely focused on this dimension of product management

	Improving Supplier Selection	Q4A: Over the last 6 months, have you carried out any analysis or actions to change the set of suppliers from whom you order stock or raw materials? By this we mean, have you used your resources (time and money) to investigate new suppliers and start procuring from them? Or have you stopped procuring from a supplier you previously had? Please describe why, when and how you took steps to change the set of suppliers you order from.	Text Response
		(Q4B: This queston is completed privately by enumerator) Over the last 6 months, how many actions did you take to change the set of suppliers you order from?	0 = Never took any deliberate actions or analysis ..to.. 5 = Took deliberate actions or analyses 5 (or more) different times
		Q4C: Over the last 6 months, have you carried out any analysis or actions to improve the relationship you have with a supplier or suppliers? By this we mean, have you tried to enhance your relationships with suppliers through actions like: being more organized and thus reducing delivery time, repaying supplier credit on time, demonstrating loyalty to suppliers, taking adventage of trade promotions they offer you, etc.? Please describe why, when and how you took steps to improve your relationship with a supplier or suppliers.	Text Response
		(Q4D: This queston is completed privately by enumerator) Over the last 6 months, how many actions did you take to improve your relationship with suppliers?	0 = Never took any deliberate actions or analysis ..to.. 5 = Took deliberate actions or analyses 5 (or more) different times
		(Q4E: This queston is completed privately by the independent product management auditor) Assess how focused the respondent is on deciding the right mix of suppliers for their business and having good relationships with them.	1 = Not focused on this dimension of product management at all ..to.. 7 = Extremely focused on this dimension of product management
	Aligning Stock Orders with Demand	Q5A: How do you typically plan or make decisions to buy new stock for your business? For example, do you order stock on an "as needed" basis i.e. order whenever you're running low or do you have a system in place for ordering fixed quantities of stock/materials at regular frequencies? Please describe how, when and why you plan to order stock for your business.	Text Response
		Q5B: Over the last 6 months, have you carried out any analysis or actions to align the amount of stock/materials ordered with your business needs? In other words, have you carried out any analysis or actions to change the quantities of stock you buy to fulfill a specific business need such as: A) Matching customer demand for the products; B) Avoid being sold out of some products; C) Avoid spoilage or wastage of stock	Text Response
		(Q5C: This queston is completed privately by enumerator) Over the last 6 months, how many actions did you take to align the amount of stock/materials ordered with your business needs	0 = Never took any deliberate actions or analysis ..to.. 5 = Took deliberate actions or analyses 5 (or more) different times
		(Q5D: This queston is completed privately by the independent product management auditor) Assess how focused the respondent is on deciding the right mix of suppliers for their business and having good relationships with them.	1 = Not focused on this dimension of product management at all ..to.. 7 = Extremely focused on this dimension of product management
	Saving Time and Money During Procurement	Q6A: Over the last 6 months, have you carried out any analysis or actions to change the way you order supplies to save time or money? For example, have you grouped orders together to make fewer trips to order supplies, or did you switch to a different supplier (or sources of raw material) who has cheaper prices? Please describe why, when and how you took steps to save time or money in ordering supplies for products.	Text Response
		(Q6B: This queston is completed privately by enumerator) Over the last 6 months, how many actions did you take to save time or money in ordering stock for your business?	0 = Never took any deliberate actions or analysis ..to.. 5 = Took deliberate actions or analyses 5 (or more) different times
		(Q6E: This queston is completed privately by the independent product management auditor) Assess how focused the respondent is on saving time and money while procuring products/materials for their business.	1 = Not focused on this dimension of product management at all ..to.. 7 = Extremely focused on this dimension of product management

Description of Customer-based Branding Measures

During each data collection round, enumerators randomly sampled the first three customers who purchased at the business. Customers were asked to rate the store on a set of branding dimensions based on Keller 1993 that they would uniquely be able to speak to: willingness to recommend, willingness to pay, brand quality, brand consistency, trust in brand and shopping satisfaction. We were not able to ask customers to rate the firm on brand personality dimensions. Our pilot tests revealed that it took too long to explain the concept to customers and thus they became reluctant to answer our survey fully. Ratings were provided by customers on a 7-point Likert scale. We then constructed the dependent variables in Table 18 as follows:

Customer-based Brand Index. We average the scores given to the firm by their customers across the six branding dimensions and across customers. Subsequently, we construct this variable by normalizing the score so that it ranges continuously from 0 to 1.

Customer-based Brand Dimensions. We average the scores given to the firm on each branding dimension across customers. We generate a "high value" dummy for each dimension of branding if the average customer score was above the sample median on that dimension. Next, we construct this variable by adding up the six binary variables, so it ranges from 0 to 6.

Description of Enumerator-based Product Management Measures

We asked enumerators to count the number of concrete actions or analyses taken by the owner on each dimension of product management during the on-site audit. Note that this measure does not account for the quality of those actions or analyses, which was assessed by the specialized product management auditor. Hence, we use this measure purely for robustness. We constructed the dependent variables in Table 19 as follows:

Enumerator-based Product Management Index. We sum the number of product management actions or analyses taken by the firm owner and verified by the enumerator, across the six dimensions of product management. Subsequently, we normalize the sum so that it ranges continuously from 0 to 1.

Enumerator-based Product Management Dimensions. We generate a "high value" dummy for each dimension of product management if the enumerator verified that the owner had taken more

actions or analyses than the sample median owner on that dimension. Next, we construct this variable by adding up the six binary variables, so it ranges from 0 to 6.

Web Appendix E: ATT on Sales

Table 13: ATT Estimate for Impact of Modernization Interventions on Firm Sales

	DV: Monthly Sales (USD)					
	(1) Midline	(2) Endline	(3) Midline and Endline Avg.	(4) Midline	(5) Endline	(6) Midline and Endline Avg.
External Treatment Adopter (Fitted)	618.4***	574.8***	546.6***			
	(235.3)	(190.0)	(169.9)			
Internal Treatment Adopter (Fitted)	498.3**	517.9***	467.8***			
	(226.6)	(188.9)	(178.9)			
External Treatment Complier (Fitted)				647.7***	606.3***	576.5***
				(245.4)	(199.9)	(178.6)
Internal Treatment Complier (Fitted)				521.6**	551.3***	498.0***
				(237.6)	(201.3)	(190.8)
Baseline Value of DV	Yes	Yes	Yes	Yes	Yes	Yes
Sub-sector FE	Yes	Yes	Yes	Yes	Yes	Yes
Biz/Owner Controls	Yes	Yes	Yes	Yes	Yes	Yes
Strata FE	Yes	Yes	Yes	Yes	Yes	Yes
Mean of DV: Control	2439.5	2954.7	2776.8	2439.5	2954.7	2776.8
SD of DV: Control	3161.3	3754.1	3409.5	3161.3	3754.1	3409.5
Effect Size in SD: Ext	0.196	0.153	0.160	0.205	0.162	0.169
Effect Size in %: Ext	25.35	19.45	19.69	26.55	20.52	20.76
Effect Size in SD: Int	0.158	0.138	0.137	0.165	0.147	0.146
Effect Size in %: Int	20.43	17.53	16.85	21.38	18.66	17.94
Obs.	883	791	791	883	791	791

Notes: Estimates presented are the ATT, computed via 2SLS, with treatment adoption (defined as participation in first modernization session of the intervention) and compliance (defined as participation in at least 10 modernization sessions of the intervention) instrumented by the random treatment assignment. Data underlying these regressions were collected in two survey rounds – Midline (6 months post-intervention) and Endline (18 months post-intervention). The DV in Columns (1), (2), (4) and (5) is the monetary measure of monthly sales, i.e. the average of the recall-based and anchored-adjusted monetary sales estimates, both winsorized 1% on each tail. The DV in Columns (3) and (6) is the average of the DVs in the two previous columns, thus representing the average monthly sales in USD over the two survey rounds. The indicated regressions include: the baseline value of the dependent variable, 8 baseline controls for owner characteristics (gender; age; marital status; children status; education; prior salaried job; prior business ownership; prior business training), 8 baseline controls for business characteristics (founder; assets; weekly customers; employees; tax registration status; formal loans; profit savings rate; separation of business-personal affairs; business practices), 5 sub-sector fixed effects (set of two digit SIC codes) and 3 strata dummies indicating which randomization/implementation batch the firm was part of. All regressions exclude firms that were found non-operational or attrited during the survey round. Robust standard errors are in parentheses. P-values are highlighted as: $p < 0.1^{*}$ $p < 0.05^{**}$ $p < 0.01^{***}$

Overall, the ATT effect is positive and significant across all regressions and larger than the ITT effects reported in the main analysis. Moreover, as expected, the ATT effect increases when we define compliance as receiving at least ten modernization sessions versus one modernization session.

Web Appendix F: Robustness of Main Sales Effect

Table 14: Sales Effect Robustness Check – Different sub-samples

	All Firms inc. Non-Operational DV: Monthly Sales (USD)			Firms with Sales Records at Baseline DV: Monthly Sales (USD)		
	(1)	(2)	(3)	(4)	(5)	(6)
	Midline	Endline	Midline and Endline Avg.	Midline	Endline	Midline and Endline Avg.
External Treat	466.1**	466.7***	411.1***	723.3**	639.0***	635.6***
	(191.9)	(170.8)	(155.1)	(296.8)	(229.3)	(213.0)
Internal Treat	396.5**	343.0*	288.4*	620.2**	543.0**	512.3**
	(195.9)	(178.8)	(167.3)	(275.5)	(226.2)	(213.6)
Baseline Value of DV	Yes	Yes	Yes	Yes	Yes	Yes
Sub-sector FE	Yes	Yes	Yes	Yes	Yes	Yes
Biz/Owner Controls	Yes	Yes	Yes	Yes	Yes	Yes
Strata FE	Yes	Yes	Yes	Yes	Yes	Yes
Mean of DV: Control	2019.1	2203.1	2070.4	2591.9	3216.9	2999.0
SD of DV: Control	3019.6	3486.9	3182.1	3443.4	4133.9	3732.7
Effect Size in SD: Ext	0.154	0.134	0.129	0.210	0.155	0.170
Effect Size in %: Ext	23.09	21.19	19.86	27.91	19.86	21.19
Effect Size in SD: Int	0.131	0.0984	0.0906	0.180	0.131	0.137
Effect Size in %: Int	19.64	15.57	13.93	23.93	16.88	17.08
Obs.	1071	1046	1046	658	593	593
F-Stat.	7.900	13.48	14.65	6.518	18.47	16.16

Notes: Data underlying these regressions were collected in two survey rounds – Midline (6 months post-intervention) and Endline (18 months post-intervention). The DV in all columns is the monetary measure of monthly sales, i.e. the average of the recall-based and anchored-adjusted monetary sales estimates, both winsorized 1% on each tail. In Columns (1)–(3), the sample of firms are those found operational and non-operational (but willing to answer the survey) in each survey round, with sales for non-operational firms coded as zero. In Columns (4)–(6), the sample of firms are those found operational in each survey round who also reported keeping sales records at baseline. All regressions include: the baseline value of the DV, 8 baseline controls for owner characteristics (gender; age; marital status; children status; education; prior salaried job; prior business ownership; prior business training), 8 baseline controls for business characteristics (founder; assets; weekly customers; employees; tax registration status; formal loans; profit savings rate; separation of business-personal affairs; business practices), 5 sub-sector fixed effects (set of two digit SIC codes) and 3 strata dummies indicating which randomization/implementation batch the firm was part of. Robust standard errors are in parentheses. P-values are highlighted as: $p < 0.1^{*}$ $p < 0.05^{**}$ $p < 0.01^{***}$

Table 15: Sales Effect Robustness Check – Attrition Bounding

	DV: Monthly Sales (USD) at Endline			
	(1) Sales Growth = 0 for all attriters	(2) Sales Growth = Control Growth for all attriters	(3) Sales Growth = Control Growth for C attriters; Sales Growth = 0 for T attriters	(4) Sales Growth = Treatment Growth for C attriters; Sales Growth = 0 for T attriters
External Treat	516.4***	510.1***	489.7***	395.6**
	(166.3)	(165.8)	(166.1)	(167.2)
Internal Treat	432.8***	437.2***	405.2**	307.9*
	(157.3)	(156.7)	(157.1)	(158.6)
Baseline Value of DV	Yes	Yes	Yes	Yes
Sub-sector FE	Yes	Yes	Yes	Yes
Biz/Owner Controls	Yes	Yes	Yes	Yes
Strata FE	Yes	Yes	Yes	Yes
Mean of DV: Control	3008.6	3037.4	3037.4	3139.2
SD of DV: Control	3756.3	3792.9	3792.9	3946.9
Effect Size in SD: Ext	0.137	0.134	0.129	0.100
Effect Size in %: Ext	17.16	16.80	16.12	12.60
Effect Size in SD: Int	0.115	0.115	0.107	0.0780
Effect Size in %: Int	14.39	14.39	13.34	9.807
Obs.	893	893	893	893
F-Stat.	30.51	32.90	31.12	30.84

Notes: Column (1) assigns all attriters a sales growth of zero. Column (2) assigns all attriters the average sales growth of the control group. Column (3) assigns all control attiters the sales growth rate of the control group and all treatment attriters a sales growth of zero. Column (4) assigns all control attriters the average sales growth of the treatment groups and all treatment attriters a sales growth of zero. Data underlying these regressions were collected at Endline (18 months post-intervention). The DV in all columns is the monetary measure of monthly sales, i.e. the average of the recall-based and anchored-adjusted monetary sales estimates, both winsorized 1% on each tail. All regressions include: the baseline value of the DV, 8 baseline controls for owner characteristics (gender; age; marital status; children status; education; prior salaried job; prior business ownership; prior business training), 8 baseline controls for business characteristics (founder; assets; weekly customers; employees; tax registration status; formal loans; profit savings rate; separation of business-personal affairs; business practices), 5 sub-sector fixed effects (set of two digit SIC codes) and 3 strata dummies indicating which randomization/implementation batch the firm was part of. Robust standard errors are in parentheses. P-values are highlighted as: $p < 0.1^{*}$ $p < 0.05^{**}$ $p < 0.01^{***}$

Table 16: Sales Effect Robustness Check – Different Sets of Controls

	No Control Variables Included DV: Monthly Sales (USD)			Control Variables Selected by Lasso DV: Monthly Sales (USD)		
	(1)	(2)	(3)	(4)	(5)	(6)
	Midline	Endline	Midline and Endline Avg.	Midline	Endline	Midline and Endline Avg.
External Treat	522.3**	594.7***	538.4***	549.2**	555.3***	515.7***
	(218.8)	(187.1)	(165.8)	(218.5)	(188.6)	(162.5)
Internal Treat	420.6*	509.8***	432.0**	427.7**	501.9***	428.7***
	(221.1)	(180.4)	(169.2)	(218.1)	(179.3)	(166.4)
Baseline Value of DV	Yes	Yes	Yes	Yes	Yes	Yes
Sub-sector FE	No	No	No	1 selected	No	No
Biz/Owner Controls	No	No	No	3 selected	1 selected	2 selected
Strata FE	No	No	No	No	No	No
Mean of DV: Control	2439.5	2954.7	2776.8	2439.5	2954.7	2776.8
SD of DV: Control	3161.3	3754.1	3409.5	3161.3	3754.1	3409.5
Effect Size in SD: Ext	0.165	0.158	0.158	0.174	0.148	0.151
Effect Size in %: Ext	21.41	20.13	19.39	22.51	18.79	18.57
Effect Size in SD: Int	0.133	0.136	0.127	0.135	0.134	0.126
Effect Size in %: Int	17.24	17.25	15.56	17.53	16.99	15.44
Obs.	883	791	791	883	791	791

Notes: This table demonstrates the robustness of our sales effect — we show treatment effects with different sets of control variables. Data underlying these regressions were collected in two survey rounds – Midline (12 months post-recruitment) and Endline (24 months post-recruitment). The DV in all columns is the monetary measure of monthly sales, i.e. the average of the recall-based and anchored-adjusted monetary sales estimates, both winsorized 1% on each tail. The lasso regressions select control variables from the following set: the baseline value of the dependent variable, 8 baseline controls for owner characteristics (gender; age; marital status; children status; education; prior salaried job; prior business ownership; prior business training), 8 baseline controls for business characteristics (founder; assets; weekly customers; employees; tax registration status; formal loans; profit savings rate; separation of business-personal affairs; business practices), 5 sub-sector fixed effects (set of two digit SIC codes) and 3 strata dummies indicating which randomization/implementation batch the firm was part of. All regressions exclude firms that were found non-operational or attrited during the survey round. Robust standard errors are in parentheses. P-values are highlighted as: $p < 0.1^{*}$ $p < 0.05^{**}$ $p < 0.01^{***}$

Table 17: Panel Diff-in-Diff Estimate for Impact of Modernization Interventions on Firm Sales

	DV: Monthly Sales (USD)			
	(1) Fixed Effect; Unbalanced Panel	(2) Fixed Effect; Balanced Panel	(3) Random Effect; Unbalanced Panel	(4) Random Effect; Balanced Panel
External Treatment × Post-Treat	617.6***	431.4***	608.1***	431.4***
	(160.8)	(162.0)	(160.9)	(163.1)
Internal Treatment × Post-Treat	461.5***	478.6***	440.9***	478.6***
	(157.2)	(174.5)	(158.5)	(175.7)
Time Period FE	Yes	Yes	Yes	Yes
Sub-sector FE	No	No	Yes	Yes
Biz/Owner Controls	No	No	Yes	Yes
Strata FE	No	No	Yes	Yes
Mean of DV: Control	2679.2	2805.0	2679.2	2805.0
SD of DV: Control	3456.2	3521.9	3456.2	3521.9
Effect Size in SD: Ext	0.179	0.122	0.176	0.122
Effect Size in %: Ext	23.05	15.38	22.70	15.38
Effect Size in SD: Int	0.134	0.136	0.128	0.136
Effect Size in %: Int	17.23	17.06	16.46	17.06
Obs.	2627	2139	2627	2139

Notes: Estimates presented are from panel regressions of the DV on treatment assignment, interacted with a post-intervention dummy variable to indicate post-intervention time periods. T=3 in all regressions, corresponding to 3 survey rounds: Baseline (pre-intervention), Midline (6 months post-intervention) and Endline (18 months post-intervention). The DV in all columns is the monetary measure of monthly sales, i.e. the average of the recall-based and anchored-adjusted monetary sales estimates, both winsorized 1% on each tail. The indicated regressions include: 3 time (survey-round) fixed effects, 8 baseline controls for owner characteristics (gender; age; marital status; children status; education; prior salaried job; prior business ownership; prior business training), 8 baseline controls for business characteristics (founder; assets; weekly customers; employees; tax registration status; formal loans; profit savings rate; separation of business-personal affairs; business practices), 5 sub-sector fixed effects (set of two digit SIC codes) and 3 strata dummies indicating which randomization/implementation batch the firm was part of. Regressions in Columns (1) and (3) include N = 957 firms who reported sales estimates during any post-treatment survey round. Regressions in Columns (2) and (4) include only N = 713 firms who reported sales estimates during all 3 pre- and post-treatment survey rounds. Robust standard errors, clustered by firm, are in parentheses. P-values are highlighted as: $p < 0.1^{*}$ $p < 0.05^{**}$ $p < 0.01^{***}$

Web Appendix G: Robustness of Store-Level Branding Effects

Table 18: Robustness Tests: Effect of External Modernization on Store-level Brand

	(1) Customer-based Brand Index (Continuous: 0 to 1)	(2) Customer-based Brand Dimensions (Count: 0 to 6)
External Treat	0.0306**	0.262*
	(0.0152)	(0.155)
Internal Treat	0.0223	0.198
	(0.0147)	(0.153)
Sub-sector FE	Yes	Yes
Biz/Owner Controls	Yes	Yes
Strata FE	Yes	Yes
Mean of DV: Control	0.683	3.119
SD of DV: Control	0.193	1.874
Effect Size in % (SD): Ext	4.48 (.159)	8.41 (.140)
Effect Size in % (SD): Int	Not Sig.	Not Sig.
Obs.	843	843

Notes: This table summarizes treatment effects of our modernization interventions on the customer-based measure of store-level branding. The DV in Column (1) is the normalized average score given to the firm by randomly sampled customers across six dimensions of store-level brand: shopping satisfaction, loyalty, willingness to recommend, store quality, store consistency/reliability, and willingness to pay. The DV in Column (2) is a count of these six dimensions of store-level brand on which the firm received an above-median score from their randomly sampled customers. The indicated regressions include: 8 baseline controls for owner characteristics (gender; age; marital status; children status; education; prior salaried job; prior business ownership; prior business training), 8 baseline controls for business characteristics (founder; assets; weekly customers; employees; tax registration status; formal loans; profit savings rate; separation of business-personal affairs; business practices), 5 sub-sector fixed effects (set of two digit SIC codes) and 3 strata dummies indicating which randomization/implementation batch the firm was part of. Regressions exclude firms that were found non-operational or attrited during the Midline survey round. Robust standard errors are in parentheses. P-values are highlighted as: $p < 0.1^{*}$ $p < 0.05^{**}$ $p < 0.01^{***}$

	Mean Occurrence per Photo			
	Control/Internal Group (N = 512)	External Group (N = 272)	Difference in Means	Student's *t* Statistic
Top 10 Positive Store-Brand Words	2.913	3.603	0.689	4.850***
1. Clean	0.792	0.960	0.167	3.071***
2. Organized	0.496	0.746	0.250	5.381***
3. Friendly	0.418	0.501	0.083	2.012**
4. Colorful	0.259	0.276	0.016	0.502
5. Professional	0.175	0.250	0.076	2.753***
6. Modern	0.180	0.237	0.057	2.043**
7. Fun	0.191	0.169	-0.022	-0.840
8. Inviting	0.158	0.186	0.028	1.287
9. Neat	0.120	0.134	0.015	0.756
10. Welcoming	0.123	0.143	0.019	0.963

Notes: This table summarizes the store-level brand descriptor words for firms in our sample, by treatment group. We include N = 784 firms operational at Endline (24 months post-recruitment) who consented to providing us photos. There are 2 photographs corresponding to every included firm, each rated by approximately 5 MTurk raters who were requested to provide up to three keywords. P-values are highlighted as: $p < 0.1^{*}$ $p < 0.05^{**}$ $p < 0.01^{***}$

Web Appendix H: Robustness of Product Management Effects

Table 19: Robustness Test: Effect of Internal Modernization on Product-level Management

	(1) Enumerator-based Prod. Mgmt. Index (Continuous: 0 to 1)	(2) Enumerator-based Prod. Mgmt. Dimensions (Count: 0 to 6)
External Treat	0.0143	0.0837
	(0.0119)	(0.133)
Internal Treat	0.0262**	0.154
	(0.0132)	(0.134)
Sub-sector FE	Yes	Yes
Biz/Owner Controls	Yes	Yes
Strata FE	Yes	Yes
Mean of DV: Control	0.300	1.735
SD of DV: Control	0.151	1.583
Effect Size in % (SD): Ext	Not Sig.	Not Sig.
Effect Size in % (SD): Int	8.70 (.174)	Not Sig.
Obs.	793	793

Notes: This table summarizes treatment effects of our modernization interventions on the enumerator-based measures of product-level management. The DV in Column (1) is the normalized number of product-level management actions or analyses taken by the business owner and verified by the enumerator. The actions or analyses pertain to six dimensions of product-level management: improving product assortment, improving product quality, sales effort on popular/profitable products, improving supplier selection, aligning stock orders with demand, and saving time and money in procurement. The DV in Column (2) is the count (out of these six) dimensions of product-level management on which the owner performed at least one action or analysis. The indicated regressions include: 8 baseline controls for owner characteristics (gender; age; marital status; children status; education; prior salaried job; prior business ownership; prior business training), 8 baseline controls for business characteristics (founder; assets; weekly customers; employees; tax registration status; formal loans; profit savings rate; separation of business-personal affairs; business practices), 5 sub-sector fixed effects (set of two digit SIC codes) and 3 strata dummies indicating which randomization/implementation batch the firm was part of. Regressions exclude firms that were found non-operational or attrited during the Endline survey round. Robust standard errors are in parentheses. P-values are highlighted as: $p < 0.1^{*}$ $p < 0.05^{**}$ $p < 0.01^{***}$

	Mean Occurrence per Firm			
	Internal Group (N = 269)	Control/External Group (N = 531)	Difference in Means	Student's *t* Statistic
Top 10 Product Management Words	2.178	1.887	0.291	1.990**
1. Promotions	0.301	0.299	0.001	0.035
2. New	0.279	0.242	0.036	0.867
3. Changes	0.256	0.188	0.211	1.695*
4. Sales	0.312	0.284	0.028	0.581
5. Brands	0.212	0.214	-0.003	-0.060
6. Demand	0.207	0.204	0.002	0.077
7. Costs	0.171	0.154	0.017	0.522
8. Seasonal	0.156	0.096	0.060	1.722*
9. Deal/Bargain	0.156	0.105	0.051	1.749*
10. Variety	0.130	0.094	0.036	1.359

Notes: This table summarizes the product management descriptor words for firms in our sample, by treatment group. We include N = 800 firms operational at Endline (24 months post-recruitment) who responded to product management questions. P-values are highlighted as: $p < 0.1^{*}$ $p < 0.05^{**}$ $p < 0.01^{***}$

Web Appendix I: Robustness to Alternative Mechanisms

Table 20: Effect of Modernization Interventions on Alternative Mechanisms

	(1) Price Index (0 to 1)	(2) Confidence Score (0 to 1)	(3) Seriousness Score (0 to 1)	(4) Time Spent Improving Biz. (0 to 1)
External Treat	0.0049	0.0001	0.0211	-0.0047
	(0.0110)	(0.0202)	(0.0336)	(0.0150)
Internal Treat	0.0129	0.0070	0.0428	-0.0185
	(0.0102)	(0.0201)	(0.0338)	(0.0156)
Sub-sector FE	Yes	Yes	Yes	Yes
Biz/Owner Controls	Yes	Yes	Yes	Yes
Strata FE	Yes	Yes	Yes	Yes
Mean of DV: Control	0.389	0.637	0.457	0.265
SD of DV: Control	0.125	0.240	0.387	0.195
Effect Size: Ext	Not Sig.	Not Sig.	Not Sig.	Not Sig.
Effect Size: Int	Not Sig.	Not Sig.	Not Sig.	Not Sig.
Obs.	767	806	806	806

Notes: This table summarizes treatment effects of our modernization interventions on alternative mechanisms. The DV in Column (1) is the normalized price ratio (i.e., ratio of price paid for a product purchased at the firm relative to the price customers would have to pay to obtain the same the product outside of the firm), reported by customers of the firm, and averaged across customers. The DV in Column (2) is the normalized self-reported agreement rating by the firm owner to the statement "I am better at making business improvements than a typical business owner in the same industry as me". The DV in Column (2) is the normalized self-reported rating by the firm owner on their business purpose: where 0 represents low seriousness (i.e. the business is a means of economic sustenance for me and my family) and 1 represents high seriousness (i.e. the business is a profit-making venture to invest my time and resources). The DV in Column (3) is the proportion of time in a typical business week spent on making big-picture improvements (as opposed to managing day-to-day operations). The indicated regressions include: 8 baseline controls for owner characteristics (gender; age; marital status; children status; education; prior salaried job; prior business ownership; prior business training), 8 baseline controls for business characteristics (founder; assets; weekly customers; employees; tax registration status; formal loans; profit savings rate; separation of business-personal affairs; business practices), 5 sub-sector fixed effects (set of two digit SIC codes) and 3 strata dummies indicating which randomization/implementation batch the firm was part of. Regressions exclude firms that were found non-operational or attrited during the Midline survey round. Robust standard errors are in parentheses. P-values are highlighted as: $p < 0.1^{*}$ $p < 0.05^{**}$ $p < 0.01^{***}$

Three alternative mechanism variables in Columns (2) through (4) were measured as self-reports from business owners during the Endline data collection round. We measured owner confidence in making business changes by initially priming them to think about the knowledge they have acquired on business activities as well as resources available to them to make business improvements. Then, we ask owners to rate the extent to which they agree that "I am better at making business improvements than a typical business owner in the same industry as me", on a 5-point Likert scale. This normalized rating is our dependent variable measuring owner confidence. To measure seriousness of business purpose, we ask another 1 to 5 scale question on how the owner views the purpose of their business, where 1 represents low seriousness ("A means of economic sustenance for me and my family"). The scale options progressively increase in seriousness, up to 5 which represents high seriousness ("A profit-making venture to invest my time and resources"). This normalized rating is our dependent variable measuring seriousness of business purpose. Finally, we ask the business owner to consider a typical business week and report the proportion of time they spend on making big-picture improvements versus routine business operational activities. This variable, which ranges from 0 to 1, is our dependent variable measuring time spent on making business improvements.

Web Appendix J: Returns to Modernization

Table 21: ROI for our Modernization Program

Panel A: Profit Gain due to Modernization Program (ITT Estimates)

	(1) Monthly Profits (USD)	(2) IHS of Monthly Profits
Assigned to Treatment	116.1*	0.701*
	(60.98)	(0.369)
Baseline Value of DV	Yes	Yes
Sub-sector FE	Yes	Yes
Biz/Owner Controls	Yes	Yes
Strata FE	Yes	Yes
Mean of DV: Control	649.4	6.882
SD of DV: Control	1078.5	5.270
Effect Size in % (SD)	17.88 (.108)	10.18 (.133)
Obs.	791	791

Notes: This table summarizes the intent-to-treat (ITT) effects of our modernization interventions on firm profits. Data underlying these regressions were collected at Endline (24 months post-recruitment). The DV in Column (1) is the monetary measure of monthly profits, i.e. the average of the recall-based and anchored-adjusted monetary profits estimates, both winsorized 1% on each tail. The DV in Column (2) is the inverse-hyperbolic-sine transformed measure of monthly profits, i.e. the average of the IHS of the recall-based and anchored-adjusted profit estimates. All regressions include: the baseline value of the dependent variable, 8 baseline controls for owner characteristics (gender; age; marital status; children status; education; prior salaried job; prior business ownership; prior business training), 7 baseline controls for business characteristics (founder; assets; weekly customers; employees; tax registration status; formal loans; separation of business-personal affairs; business practices), 5 sub-sector fixed effects (set of two digit SIC codes) and 3 strata dummies indicating which randomization/implementation batch the firm was part of. Robust standard errors are in parentheses. P-values are highlighted as: $p < 0.1^{*}$ $p < 0.05^{**}$ $p < 0.01^{***}$

Panel B: Cost per Firm (USD) of Modernization Treatment

	Low-Cost Scenario	High-Cost Scenario
1. Personnel Costs for 35 hrs of Sessions	360	1050
Modernization Agent	*75*	–
Supervisor	*225*	–
Professional Management Consultant	–	*875*
Organizational Overheads (20%)	*60*	*175*
2. Owner Opportunity Cost for 50 hrs	215	215
3. Estimated Materials Cost	100	100
Total Cost in USD	675	1365
Net Profit Gain per Month in USD	116	116
Time to Realize Positive ROI	5.8 months	11.8 months

Notes: In Panel B, we chart out costs of a modernization program under two scenarios: a low-cost scenario that corresponds to the intervention we implemented in partnership with universities and NGOs, and a high-cost scenario where professional management consultants deliver the intervention. Professional management consultant hourly wages were estimated from a Glassdoor search of management consultant salaries in Mexico City. We present costs for the external modernization intervention, as this intervention was more expensive. We estimated cost of the following materials, typically implemented during an external modernization intervention: large banner for store (USD 50), flyers or newsletters (USD 25), stationary (USD 5) and paint (USD 20). Opportunity cost of owners time is the federal minimum wage in Mexico: USD 4.30. We conclude the table with an estimate of the time required to recuperate program costs, given the $116 gain in profits that operational firm owners experience at least 24 months post-recruitment (from Panel A).

Table 22: Modernization Structures and Firm Sales

	DV: Monthly Sales (USD) at Midline				
	(1)	(2)	(3)	(4)	(5)
	OLS	2SLS: Treatment as IV for Structures	OLS	OLS	OLS
Count of Total Modernization Structures: 0 to 40	26.30**	175.2**			
	(11.05)	(73.08)			
Count of External Modernization Structures: 0 to 20			39.54**		
			(19.69)		
Count of Internal Modernization Structures: 0 to 20				43.49**	
				(20.83)	
2nd Quartile of Modernization Structures: 4 to 10					383.0**
					(189.3)
3rd Quartile of Modernization Structures: 11 to 20					419.2*
					(237.7)
4th Quartile of Modernization Structures: 21 to 40					649.5**
					(266.7)
Baseline Value of DV	Yes	Yes	Yes	Yes	Yes
Sub-sector FE	Yes	Yes	Yes	Yes	Yes
Biz/Owner Controls	Yes	Yes	Yes	Yes	Yes
Strata FE	Yes	Yes	Yes	Yes	Yes
Obs.	759	759	759	759	759
First Stage F	–	10.47	–	–	–

Notes: Data underlying these regressions were collected in the Midline survey round (12 months post-recruitment). The DV in all columns is the monetary measure of monthly sales, i.e. the average of the recall-based and anchored-adjusted monetary sales estimates, both winsorized 1% on each tail. The main explanatory variables are counts of the number of modernization structures implemented by the firm (out of a maximum possible 20 external modernization structures and 20 internal modernization structures). The indicated regressions include: the baseline value of the dependent variable, 8 baseline controls for owner characteristics (gender; age; marital status; children status; education; prior salaried job; prior business ownership; prior business training), 8 baseline controls for business characteristics (founder; assets; weekly customers; employees; tax registration status; formal loans; profit savings rate; separation of business-personal affairs; business practices), 5 sub-sector fixed effects (set of two digit SIC codes) and 3 strata dummies indicating which randomization/implementation batch the firm was part of. All regressions exclude firms that were found non-operational or attrited during the survey round. Robust standard errors are in parentheses. P-values are highlighted as: $p < 0.1^{*}$ $p < 0.05^{**}$ $p < 0.01^{***}$

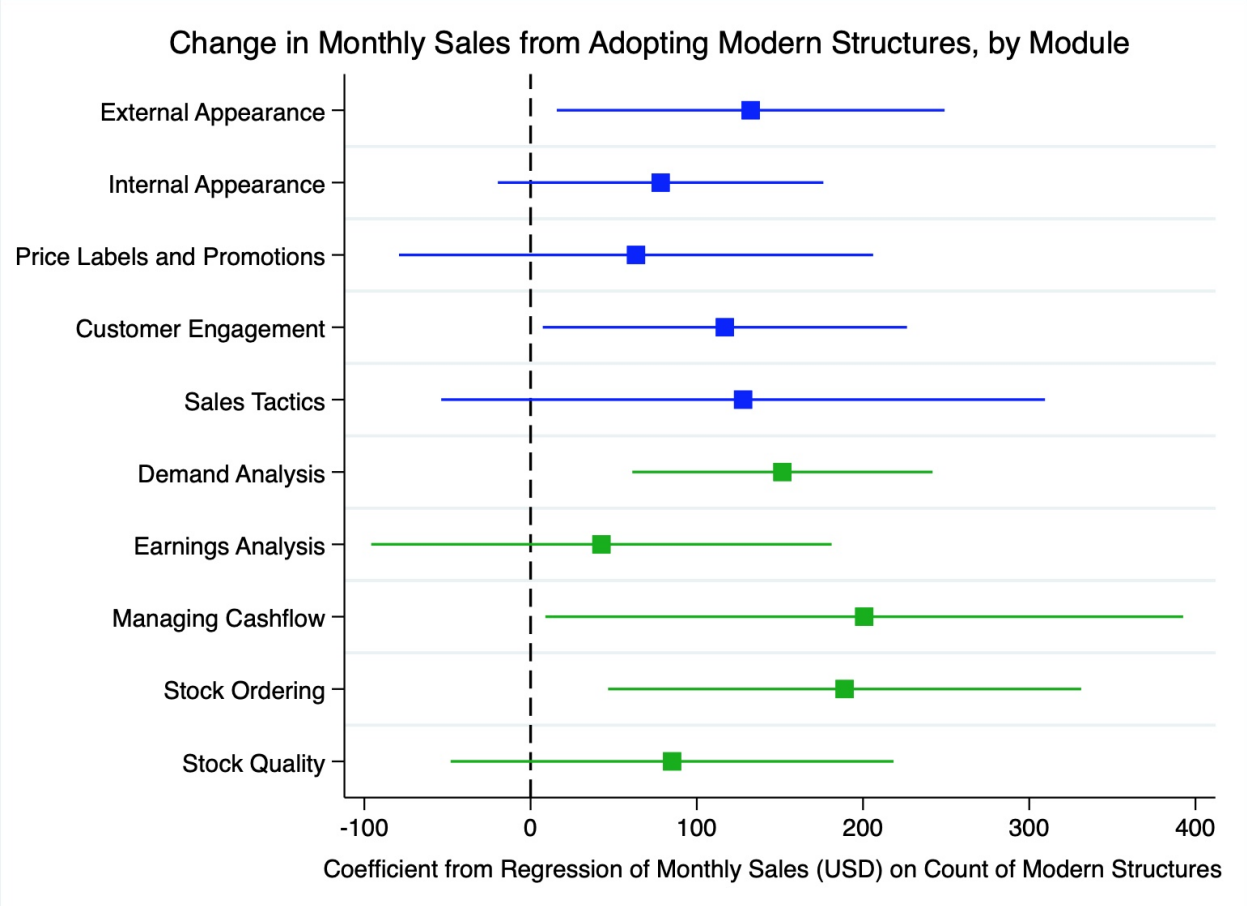


Figure 15: This figure plots the coefficient estimates from regressions of Monthly Sales at Midline (12 months post-recruitment) on the count of modern structures within each module. Each of the ten modules consist of four modern structures, detailed in Figure 2 and 3. Modules that correspond to external modernization are reported in blue, while modules that correspond to internal modernization are reported in green. The regressions include: the baseline value of the dependent variable, 8 baseline controls for owner characteristics (gender; age; marital status; children status; education; prior salaried job; prior business ownership; prior business training), 8 baseline controls for business characteristics (founder; assets; weekly customers; employees; tax registration status; formal loans; profit savings rate; separation of business-personal affairs; business practices), 5 sub-sector fixed effects (set of two digit SIC codes) and 3 strata dummies indicating which randomization/implementation batch the firm was part of. The 90% confidence interval around each coefficient estimate, based on robust standard errors, is reported. Finally, we report the fraction of treatment compliers that implemented structures within each module: external appearance – 76%; internal appearance – 83%; price labels and promotions – 61%; customer engagement – 72%; sales tactics – 51%; demand analysis – 62%; earnings analysis – 48%; managing cashflow – 57%; stock ordering – 55%; stock quality – 69%.