

Management Training and Firm Performance: Evidence from a Sponsored Program

By TITAS POVILAITIS

This paper estimates the causal effect of a sponsored managerial training program on the productivity of small firms. Using rich administrative data from 2010 to 2019 difference-in-differences is employed to compare firms that voluntarily enrolled in the program starting in January 2013 with eligible non-participants. It is found that the program is related to a substantial 41.8% increase in revenue per worker and a 24.2% increase in sales per worker. These positive effects are robust to multiple checks, hold consistently across the hospitality, retail, and fast food sectors, and are supported by visual evidence of parallel pre-treatment trends. The results provide evidence that managerial training can significantly enhance firm productivity suggesting high returns to scaling such interventions.

I. Introduction

Business outcomes such as productivity are fundamental for economic growth as they influence employment, wages, and overall competitiveness in both developed and developing economies. Enhancing productivity through targeted interventions like managerial training programs has been garnering attention as a means to improve business outcomes such as revenue, sales, and operational efficiency. Small-sized firms which often face resource constraints and skill gaps stand to benefit particularly from such programs as they can help lead to better management practices as well as higher labor utilization. However, establishing the causal impact of these interventions remains challenging due to selection biases, confounding factors, and data limitations.

Prior research underscores the potential of managerial training to drive productivity gains. For instance, Bloom et al. (2013) demonstrate through a randomized experiment in Indian textile firms that intensive management consulting improved productivity which highlights the role of structured practices in operation efficiency. Similarly, McKenzie and Woodruff (2014) review global evaluations of business training and find modest but positive effects on firm survival and profits. In rural Mexico, Calderon et al. (2020) show with a randomized trial that business literacy training increased revenues for microenterprises, hence, emphasizing the value of skill-building for underserved firms. Brooks et al. (2018) find in Kenya that mentorship-focused training boosted profits suggesting interpersonal guidance outperforms traditional classroom methods. Finally, Higuchi et al. (2015) document sustained productivity improvements from Kaizen training in Vietnamese firms with effects persisting in the long-term.

With this taken into account, the paper evaluates a government-sponsored managerial training program launched on January 1, 2013, aimed at increasing productivity among small-sized firms with 100 or fewer employees. Eligible firms could enroll voluntarily, thus, creating a natural quasi-experiment between participants (treated) and non-participants (control). Using firm-level monthly panel datasets from 2010 to 2019 to create a clean and cohesive dataset for analysis, the study focuses on productivity proxies such as revenue per worker and sales per worker. As it will be made clearer onwards, there were substantial positive effects with robustness checks confirming the validity of key assumptions.

The paper proceeds as follows: Section II looks at the dataset and its processing steps, Section III outlines the empirical strategy, Section IV presents the results, and Section V concludes.

II. Relevant Dataset

In order to ascertain the real effects of managers' training programs concerning firm productivity parameters, administrative data was utilized. The empirical analysis was conducted using a comprehensive firm-level monthly panel dataset spanning from 2010 to 2019. The initial raw dataset comprised approximately 56,600 observations. A series of data cleaning and processing steps were required to construct a reliable and analytically sound final sample. That is because a cohesive dataset had to be constructed out of 122 smaller datasets including one presenting firm-specific codes and their respective sectors (hospitality, retail, and fast food) and the other comprising monthly sales for each company. 120 separate datasets provided data on wage bills, revenues, employment, and the fact of enrolling into the training program (defined as 1 if a firm is training, 0 otherwise).

A significant issue surrounding data cleaning pertained to inconsistent date formats across the data files. In a substantial subset of the entries, the month and day values were reversed. This inconsistency posed a threat to the integrity of the panel structure and the accurate sequencing of pre- and post-treatment periods. To rectify this, a uniform date parsing procedure was implemented, which first identified the format ambiguity by checking for invalid dates (like a month value of 13) and then applied a consistent YYYY-MM-DD formatting standard across most of the dataset. Although the function was not uniformly successful at tackling all faulty date formats (there were still about 1,100 that were not changed, but were hence withdrawn from the dataset), this helped to ensure a more correct chronological order of all observations which is fundamental for the subsequent time-series and difference-in-differences analyses.

Another critical challenge involved the firm identification codes. A notable number of firms had codes that contained trailing zeros (that were not present in the firm information dataset) which created a mismatch when attempting to merge records with the master firm database that contained the canonical identifiers. This could have resulted in a failure to correctly link financial and employment data leading to erroneous or missing observations. The solution was to programmatically standardize all firm codes by removing any zeros in a firm code that may exist in an observation. This thereby automatically removed any leading or trailing zeros. The process provided a consistent and accurate identifier that enabled a successful and reliable merge across all data tables and preserved the panel structure for each firm.

The handling of missing data was guided by a principle of balancing between data retention and safeguarding against the introduction of bias. However, given the randomness of registered

NAs across variables and firms, a blanket deletion of any row with a missing value was deemed appropriate. This conservative strategy helped to ensure that firms' missing variables for a given period may not be linked to observations of that same firm in other periods which would compromise the cohesiveness of time-specific data that is crucial for differences-in-differences. By retaining only contributing data, this aided in enhancing the statistical reliability of the analysis.

To align the study's focus with its framework concerning the eligible-for-training enterprises, a sample restriction was necessary. The raw dataset included firms of all sizes. The analysis was deliberately restricted to firms with 100 or fewer employees. This filter was applied to isolate the effects of the training program on the specific firm segment it was designed to target as larger corporations were not subject to the training program. This step reduced the sample size but increased the relevance of the estimated treatment effects.

Finally, the assignment of treatment status required careful attention. A handful of firms in the dataset lacked an explicit employment number. However, these firms were present among the entities treated for the program (they had 1 as their dummy variable for training in January 2013). A decision was made to classify these firms as part of the eligible treatment group. This approach is justified under the intention-to-treat principle as these firms were part of the population of treated firms and their inclusion directly indicates that their employment number was 100 or less.

The culmination of these data processing steps resulted in a final and clean analytic sample of about 26,420 firm-period observations further dissected into three different types like hospitality, retail, and fast food. The descriptive statistics can be found in Table 1. It forms the basis for all subsequent empirical models.

Variable	N	Mean	Std. Dev.	Min	Pctl. 25	Pctl. 75	Max
Revenue per Worker	26417	4327	3187	268	2183	5475	40106
Sales per Worker	26417	75	28	19	56	89	259
Revenue per Wage	26417	14	11	1	7.4	18	173
Sales per Wage	26417	0.25	0.11	0.042	0.18	0.3	1.5
Employment	26417	66	24	20	46	85	178
Revenue	26417	328764	317225	8159	104743	440473	3664818
Wage Bill	26417	21653	11993	1948	12680	27740	103253
Sales	26417	5301	3304	504	2716	7068	26371
Treated Firm	26417	0.68	0.47	0	0	1	1
<i>Firm Sector:</i>							
Fast Food	6795	26%					
Hospitality	8098	31%					
Retail	11542	44%					

Note: All statistics are computed over the full sample.

TABLE 1: SUMMARY STATISTICS

III. Empirical Strategy

Given the ultimate aim to estimate the causal effect of a government-sponsored training program on firm productivity and outcomes, it is key to credibly identify this impact. Thus, the empirical strategy should, first, define and operationalize robust measures of firm productivity and, second, employ a quasi-experimental research design that can isolate the program's impact from other confounding factors.

Firm productivity measures: the latter seems to be one of the most fundamental challenges, especially in the case in small-sized firms where detailed data on inputs and outputs are often unavailable or unreliable. It is addressed by employing four distinct yet complementary proxy variables for productivity. They each capture a slightly different facet of a firm's operational efficiency.

The first as well as, it will be argued, primary measure used in the analysis is **revenue per worker**. It captures the average revenue generated by each employee and serves as a direct indicator of labor productivity. Utilizing the log transformation aids in normalizing the distribution of this variable and in allowing for a percentage interpretation of the treatment effect. This measure is plausibly the most sensitive to value-added activities and is a common metric in the literature for assessing the efficiency of labor utilization. That is why it will be the primary estimate in this paper and it will be the only one used in the robustness check part of the study.

Secondly, **sales per worker** measure the volume of goods or services sold per employee. A positive effect on this measure would probably suggest that the training program improved the operational throughput or sales efficiency of the workforce independent of price changes. That is why such a variable can be an appropriate indicator of firm outcomes and their effectiveness.

Revenue per wage bill assesses the return on a firm's labor expenditure. It indicates how much revenue is generated for every currency unit spent on employee wages. Likewise, its increase could entail that the firm is extracting more value from its wage investments that could be due to either higher output or more efficient compensation structures post-training.

Finally, **sales per wage bill** is similar to the previous measure. It is a pure measure of operational efficiency relative to labor costs that may isolate the physical output from revenue-based fluctuations. It is particularly useful for assessing whether the training led to a more cost-effective workforce in terms of direct sales volume.

These four outcomes plausibly provide a holistic view of whether the training program made firms more efficient in their use of labor both in absolute terms (per worker) and in cost-efficiency terms (per wage bill). To estimate the causal effect of the training program the study will employ a research design that is particularly convenient for the setting of the project.

Difference-in-Differences estimation: this methodology, inspired by Cunningham (2021), leverages the variation in small-sized firms who were not training (control group) and who engaged in the program since January 2013 (treated group). The baseline two-way fixed effects (TWFE) regression model used to estimate impacts on different productivity variables is as follows:

$$Y_{it} = \beta_0 + \beta_1 \cdot \text{adopt_t}_{it} + \alpha_i + \lambda_t + \varepsilon_{it} \quad (1)$$

Where:

- Y_{it} : outcome variable for firm i in month t .
- adopt_t_{it} : binary treatment indicator that equals 1 for all months after a firm has been treated (so, from January 2013 onwards for the treated group) and 0 otherwise.
- α_i : firm fixed effects.
- λ_t : time fixed effects.
- ε_{it} : idiosyncratic error term.
- β_1 : average treatment effect on the treated firms.

Standard errors are clustered at the firm level to account for arbitrary serial correlation within firms over time. The key identifying assumption for a causal interpretation of β_1 is the parallel trends assumption. In the absence of the treatment the trends in productivity outcomes for the treatment and control groups would have evolved in parallel. This assumption will be tested in the next section and will be visible in figures illustrating tendencies three years before (2010) and after (2016) the start of the treatment.

IV. Results of the Model

Baseline results: Table 2 provides results from the two-way fixed effects regression. It seems that there is strong evidence that the training program had a significant positive effect on several dimensions of productivity. The coefficient on the treatment dummy adopt_t is positive and statistically significant at the 1% level for all four outcome variables.

The primary outcome, log revenue per worker, shows a substantial treatment effect of 0.4176. It indicates that, on average, firms that participated in the training program experienced a 41.8% increase in revenue generated per employee compared to the control group after accounting for firm-specific and time-specific fixed effects. This is a large and economically meaningful result which may suggest the program successfully enhanced labor productivity.

Similarly, sales per worker increased by about 24.2% indicating that the program improved the physical volume of sales handled by each employee independent of price changes. The effect of revenue per wage bill is also positive, but smaller (8.9%) implying that firms generated more revenue for every unit of currency spent on wages, hence, pointing toward improved efficiency of labor expenditure.

Nevertheless, the effect on sales per wage bill is a notable exception. Even though small (1.36%), its coefficient suggests that the cost-efficiency in terms of physical sales volume did not improve and may have slightly decreased. A plausible explanation for this could be that the training led to wage increase for a more skilled workforce which raised the denominator (wage bill) without a fully compensatory increase in the sales volume numerator in the short run.

	Revenue/Worker	Sales/Worker	Revenue/Wage	Sales/Wage
adopt_t	0.4176*** (0.0091)	0.2422*** (0.0064)	0.0890*** (0.0094)	-0.0136*** (0.0017)
Fixed-effects	Yes	Yes	Yes	Yes
firm_id	Yes	Yes	Yes	Yes
date	Yes	Yes	Yes	Yes
Observations	26417	26417	26417	26417
R ²	0.83758	0.70290	0.76986	0.60297
Within R ²	0.08478	0.05900	0.00392	0.00284

Clustered (IID) standard-errors in parentheses

*Signif. Codes: *** 0.01, ** 0.05, * 0.1*

TABLE 2: REGRESSION RESULTS

Parallel trends assumption: the credibility of these DiD estimates hinges on the parallel trends assumption. Figures 1 through 4 provide visual evidence supporting this assumption for three out of four (except the revenue per wage, Figure 3) variables. In these figures, the pre-treatment periods (to the left of the dashed line marking the end of Q4 2012) show the treatment and control groups moving in close parallel. There are no statistically significant divergences in trends before the program's implementation which strengthens the case that the post-treatment divergence is indeed caused by the program. As for Figure 3, a parallel trend is visible, too. However, the difference is that after the training the control and treatment converged suggesting that revenue per wage bill became more similar between the two groups.

The "jump" in outcomes for the Figures 1 and 2 (a "dip" for Figure 4) and their treated groups precisely after the end of Q4 2012 is visually obvious. This sharp discontinuity coinciding with the treatment's start, following a period of parallel pre-trends, strongly supports a causal interpretation. The trends continue to diverge (except for Figure 3) in the post-treatment period indicating a persistent effect of the training rather than a temporary shock.

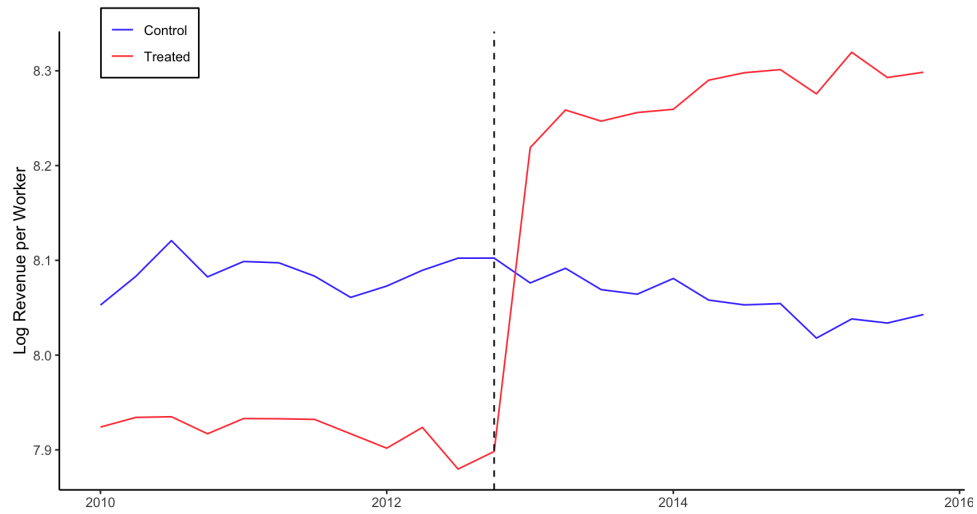


FIGURE 1: PARALLEL TRENDS IN REVENUE PER WORKER

Note: In all graphs onwards, the dashed line represents the end of Q4 2012 (start of treatment).

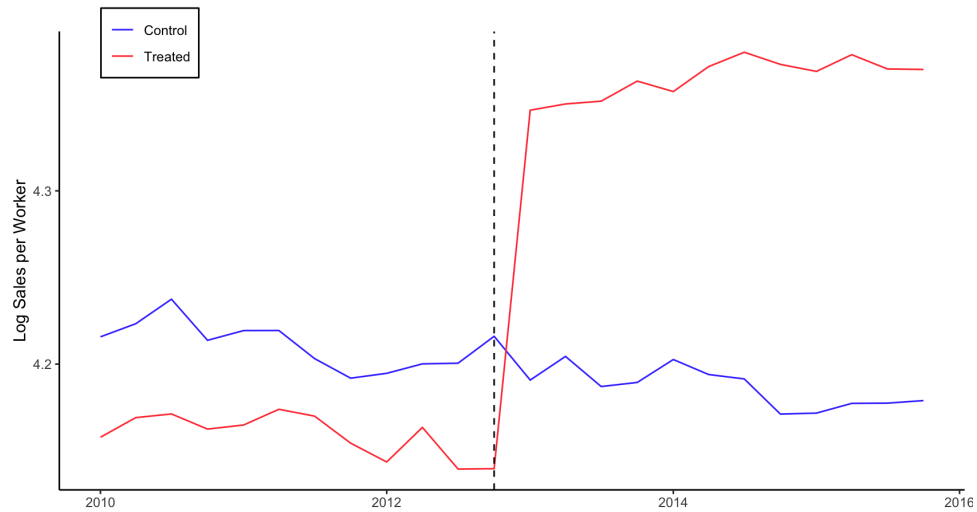


FIGURE 2: PARALLEL TRENDS IN SALES PER WORKER



FIGURE 3: PARALLEL TRENDS IN REVENUE PER WAGE BILL



FIGURE 4: PARALLEL TRENDS IN SALES PER WAGE BILL

Robustness checks: Further analysis succeeds in bolstering the robustness of these findings. Table 3 presents a balance test of pre-treatment characteristics. While the treated and control groups are not perfectly balanced (especially in terms of standard deviations of their mean variables and firm employment numbers), the DiD design effectively controls for these time-invariant differences. The critical requirement of parallel pre-treatment trends and not levels is still met. This is confirmed by the previous graphs.

Variable	Control Mean	Control SD	Treatment Mean	Treatment SD
Log Revenue per Worker	8.09	0.53	7.92	0.63
Employment	72.35	21.63	63.41	25.13
Revenue	291602	185364	253877	258606
Treated Firm	0		1	

Note: Means and standard deviations (SD) are computed for the pre-treatment period (before Q1 2013). Treated Firm is a binary indicator (0 = Control, 1 = Treatment).

TABLE 3: BALANCE FOR PRE-TREATMENT CHARACTERISTICS

The study also reveals heterogeneous, but not dramatically different, effects across sectors as detailed in Table 4. The treatment effect, measured through log revenue per worker, is consistent and positive in all three sectors (hospitality, retail, and fast food) ranging between 38 and 43.8 percent. This is close to the average effect derived for the general case (Figure 1). The stability of this large, positive effect across three distinct industries suggests the training's value is not confined to a specific business model, but generalizes across service-oriented sectors. Thus, the cross-sector validation enhances the external validity of the overall result. Furthermore, the parallel trends assumption holds within each sector as they closely resemble the trends of Figure 1. This is visible while looking at Figures 5 through 7 for respective sectors. They show no pre-trend differences and a clear post-treatment divergence for each industry.

Sector	Treatment Effect	N
Hospitality	0.437*** (0.0183)	7323
Retail	0.380*** (0.0162)	11099
Fast Food	0.438*** (0.0217)	5775
Fixed-effects	Yes	
firm_id	Yes	
date	Yes	

Clustered (IID) standard-errors in parentheses

*Signif. Codes: *** 0.01, **: 0.05, *: 0.1*

TABLE 4: SECTOR-SPECIFIC TREATMENT EFFECTS

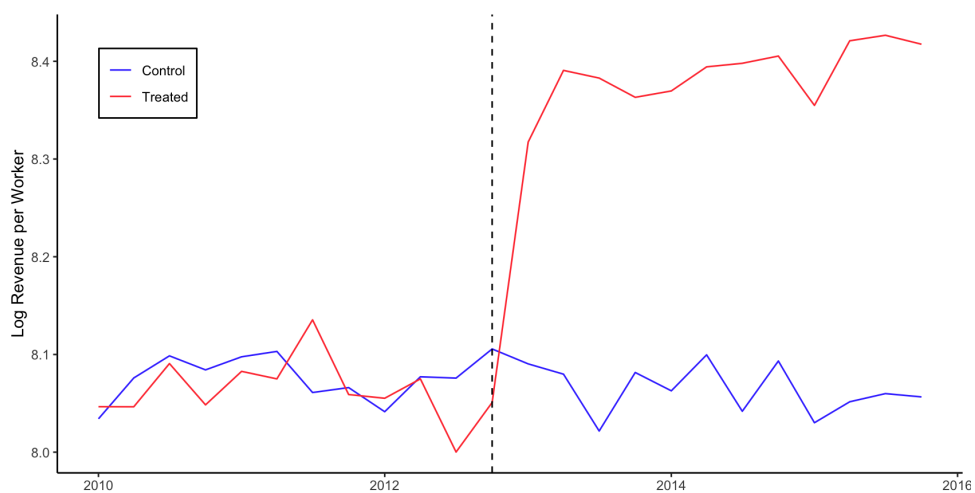


FIGURE 5: PARALLEL TRENDS IN HOSPITALITY

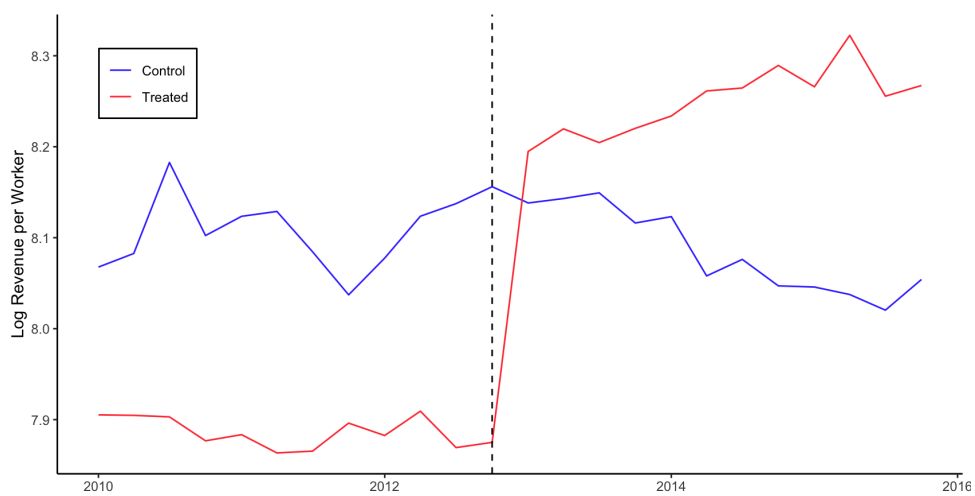


FIGURE 6: PARALLEL TRENDS IN RETAIL

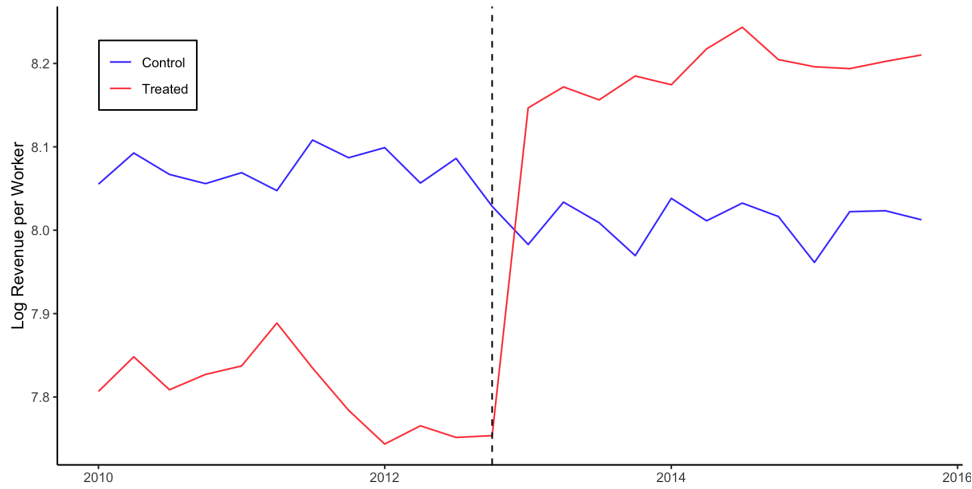


FIGURE 7: PARALLEL TRENDS IN FAST FOOD

V. Concluding Remarks

In conclusion, the study provides robust and causal evidence that a managerial training program significantly enhanced productivity in small and medium-sized enterprises. Using a difference-in-differences design, the analysis finds that the program led to a substantial 41.8% increase in revenue per worker and a 24.2% increase in sales per worker. For revenue per worker, these positive effects are consistent across hospitality, retail, and fast food sectors which demonstrates the broad applicability of the intervention. The results withstand checks for parallel pre-trends and sample balance that further lend credibility to the causal claim that training and not other factors drove the improvement.

This effectively entails policy implications. The recommendation seems to be clear. Scaling up this specific training program could prove highly effective for boosting productivity within the small business sector. The large and positive returns suggest that social benefits of such an initiative could outweigh the costs. The cross-sector validity indicates the curriculum's principles may be generalizable which makes it a strong candidate for wider implementation. Nevertheless, there are some missing factors that could prove fundamental for optimal policy design. Firstly, the negative effect on sales per wage bill suggests the training may have increased labor costs possibly through higher wages for upskilled workers. A cost-benefit analysis incorporating these increased wage costs is needed to calculate the program's precise net fiscal return. Second, some literature on management practices (Bloom et al., 2019) indicate that the specific content of the training and the quality of its delivery are critical mediators of its success. Investigating which managerial practices (such as inventory management, marketing, or HR practices) were responsible for the productivity gains could very well strengthen the argument for a nationwide rollout. It would allow policymakers to refine the curriculum for even greater impact which could then be followed by testing the effects for larger firms and the impacts of training on their outcomes.

References

- [1] Bloom, N., Brynjolfsson, E., Foster, L., Jarmin, R., Patnaik, M., Saporta-Eksten, I., & Van Reenen, J. (2019). What drives differences in management practices? *American Economic Review*, 109(5), 1648-83.
- [2] Bloom, N., Eifert, B., Mahajan, A., McKenzie, D., & Roberts, J. (2013). Does management matter? Evidence from India. *Quarterly Journal of Economics*, 128(1), 1-51.
- [3] Brooks, W., Donovan, K., & Johnson, T. R. (2018). Mentors or teachers? Microenterprise training in Kenya. *American Economic Journal: Applied Economics*, 10(4), 196-221.
- [4] Calderon, G., Cunha, J. M., & De Giorgi, G. (2020). Business literacy and development: Evidence from a randomized controlled trial in rural Mexico. *Economic Development and Cultural Change*, 68(2), 507-540.
- [5] Cunningham, S. (2021). Causal inference: the mixtape. 1st edition. New Haven: *Yale University Press*.
- [6] Higuchi, Y., Nam, V. H., & Sonobe, T. (2015). Sustained impacts of Kaizen training. *Journal of Economic Behavior & Organization*, 120, 189-206.
- [7] McKenzie, D., & Woodruff, C. (2014). What are we learning from business training and entrepreneurship evaluations around the developing world? *The World Bank Research Observer*, 29(1), 48-82.