A Data-driven Approach to Improve Artisans' Productivity in Distributed Supply Chains

Divya Singhvi
NYU Leonard N. Stern School of Business, divya.singhvi@stern.nyu.edu

Somya Singhvi
USC Marshall School of Business, ssinghvi@marshall.usc.edu

Xinyu Zhang
NYU Leonard N. Stern School of Business, xz1151@stern.nyu.edu

Despite their vital role in the global rural economy, and as a major source of employment for women in the developing world, artisanal supply chains continue to be plagued by low productivity and high poverty levels. Identifying *effective* and *implementable* solutions to improve artisan productivity is a challenging task due to high fragmentation in the upstream parts of the supply chain. This study presents research conducted in close collaboration with one of the leading exporters of handmade rugs in India. Leveraging insights from the field visits, we provide robust empirical evidence that frequent supervisor visits can play a crucial role in improving artisans' productivity. Our results from Instrumental Variables analysis indicate that a one-day decrease in the average number of days between supervisor visits to remote weavers can decrease weaving times by 13.1%-14.1%, which can lead to a 15%-17% increase in monthly income for weavers. We also find that this impact is heterogeneous, with visits to difficult-to-weave rugs, and visits that are more consistently scheduled, leading to maximum productivity gains for the weavers. To capitalize on these insights, we propose a novel predict-then-optimize framework for optimizing supervisor visits in the supply chain. Finally, using real-world data from our collaborator, we demonstrate that the proposed framework can significantly increase weaver productivity even after accounting for various operational and scheduling constraints. This research highlights how supply chain considerations can play a critical role in improving the productivity of the workforce in resource-constrained settings.

Key words: Smallholder artisans, developing countries, poverty alleviation, empirical analysis, multi-method, predict-then-optimize framework, sustainable operations

1. Introduction

Identifying effective strategies for improving productivity and increasing income in distributed supply chains is of significant importance, particularly for developing countries that are dominated by such supply chains. Distributed supply chains are characterized by small-scale producers who operate independently or as part of a larger network, are dispersed across large geographical areas, and rely on traditional techniques and manual labor for production. For instance, millions of smallholder farmers across a vast geographical area dominate the upstream parts of agricultural supply chains (World Bank 2021), and a significant fraction of textile and artisanal production in developing countries is still conducted by artisans from their individual households (Banik 2017). The widespread presence of distributed supply chains in developing countries highlights their potential as a driver of economic growth and sustainable development.

One such distributed supply chain that is prevalent across many developing countries is the artisanal supply chain. The United Nations Educational, Scientific and Cultural Organization/Information Technology Community (UNESCO/ITC) Symposium on Crafts and the International Market adopted the following definition for handicrafts in 1997, "Artisanal products or handicrafts are those produced by artisans, completely by hand or with the help of hand-tools and even mechanical means, as long as the direct manual contribution of the artisan remains the most substantial component of the finished product" (Ted and Marina 2006). The total value of the global artisanal and handicrafts market was $526 Billion (USD) in 2017 and is expected to reach $984 Billion (USD) by 2023 (NEST 2018). It is the second largest employer of the workforce in the developing world after agriculture and acts as a major contributor to the export economy. More importantly, the sector is the second biggest source of employment for rural women. While the artisan sector plays a crucial role in the global rural economy, many workers in this sector struggle with low productivity and poverty (Banik 2017). Given its importance and sustained role in the global rural economy, ex-Secretary of State John Kerry aptly remarked, "If you're looking for innovative ways to help developing countries flourish, artisans are a terrific place to begin."(The Artisan Alliance 2019). Therefore, identifying ways to improve productivity and income in these supply chains is critical for many developing countries.

The existing literature on productivity improvement in developing countries, encompassing fields such as economics (Syverson 2011), operations management (Diwas et al. 2020), and industrial organization (De Loecker and Syverson 2021), predominantly focuses on conventional factories and manufacturing plants. However, the insights from these studies may not be directly transferable to highly distributed supply chains, which feature small-scale, individual household production and considerable heterogeneity. For instance, while proven management strategies like maintaining tidy factory floors, standardizing processes, and implementing quality control practices have been demonstrated to enhance productivity in traditional factories (Bloom et al. 2013a), their application is challenging in distributed supply chains where workers are geographically dispersed and operate from their homes. As a result, there is a pressing need for research that explores productivity improvement strategies specifically designed to address the unique attributes of distributed supply chains.

This work is the outcome of a close collaboration with Jaipur Rugs, one of the largest handloom rug exporters in India. Jaipur Rugs was established in 1978 and employs thousands of rural artisans from as many as 800 villages in Northern India. A majority of the weavers in the company's supply chain, approximately 80%, are rural women from geographically isolated regions in Rajasthan, Gujarat, and Uttar Pradesh. Since most weavers are women in conservative rural areas, the company installs looms in individual households, allowing women to work independently on rug weaving. However, due to high fragmentation in the upstream parts of the supply chain, many artisans suffer from low productivity. Moreover, since artisans are paid based on a piece-rate basis, low productivity translates to low income and leads to high poverty among smallholder artisans. This raises further concern for the artisans' welfare since income matters even more for the subjective well-being (SWB) of people at low-income levels (Diener and Biswas-Diener 2002).

In this paper, we employ a multi-method approach to assess and improve artisans' productivity and income in Jaipur Rugs' supply chain. Through field visits and qualitative interviews, we identify that supervisors, who have regular interactions with artisans, play a pivotal role in this supply chain. As artisans are located in villages spread across a vast geographical area, supervisors are recruited by the company from the local communities and each supervisor is mandated to supervise all artisans in a group of villages. Given that each supervisor is responsible for overseeing hundreds of artisans in the supply chain, it is infeasible for them to visit every artisan every day. Instead, supervisors plan their own schedules and visit a subset of artisans on a daily basis. Each visit from the supervisor includes monitoring artisans' progress, carefully inspecting weaving done for errors, providing feedback on the work done, and assigning the next rug for weaving to the artisan. A key area that can be optimized in the company's supply chain is the scheduling of supervisor visits.

However, it is unclear whether optimizing supervisor visits can benefit artisans in practice. Supervision can play a dual role in rug weaving productivity. On the one hand, it can enhance productivity by ensuring adherence to best practices, maintaining quality standards, providing timely feedback, and motivating workers, as demonstrated in studies by Lurie and Swaminathan (2009), Bloom and Van Reenen (2007) and Bloom et al. (2013b). On the other hand, excessive supervision may hinder productivity by stifling creativity, diverting cognitive resources, limiting autonomy, reducing intrinsic motivation, and causing burnout, as suggested by Gusnard (2005), Amabile et al. (1996), Pierce and Aguinis (2013) and George (2007). Hence, this paper aims to investigate the following key research questions: (i) What is the impact of the frequency of supervisor visits on artisans' productivity? (ii) What factors impact the effectiveness of supervisor visits on artisans' productivity? (iii) How can we optimize supervisor operations to improve artisans' productivity?

1.1. Contributions

Based on the insights from our field visits and empirical analysis using Jaipur Rugs' internal supply chain data, we identify two important findings. First, more frequent supervisor visits can play a key role in improving artisans' productivity. In particular, a one-day decrease in average days between visits can decrease weaving times by 2.8%-14.1%. Economically, the productivity gains from a one-day decrease in average days between supervisor visits translate to a 15%-17% increase in monthly income for the weavers. Second, the relationship between weaving times and supervisor visits is heterogeneous. We find that, (i) visits to weavers with difficult-to-weave rugs and, (ii) visits that are consistently scheduled, are most effective in improving artisans' productivity. We hypothesize that this heterogeneity is due to the fact that supervisor visits help in identifying inadvertent weaving errors that are costly to rectify and these errors are more likely to occur in difficult-to-weave rugs. These insights provide, to the best of our knowledge, the first empirical analysis of the impact of supervision on productivity in distributed supply chains.

Following the findings from our empirical analysis, we devise a *predict-then-optimize* framework to optimize supervisor visits. In particular, our two-step procedure first uses state-of-the-art machine learning methods to predict rugs that are likely to have low productivity. Then, using those predictions, we formulate a scheduling-and-routing optimization problem that optimizes supervisor visits to villages, accounting for various practical travel and working-time-related constraints. We relate our problem to the celebrated prize-collecting TSP problem (Balas 1989, Feillet et al. 2005, Chekuri et al. 2012, Xu et al. 2020) and as a consequence prove the existence of a polynomial time algorithm with a provable performance guarantee. Finally, using data from our collaborator, we perform extensive numerical analysis to demonstrate the value of the proposed methodology. In particular, estimates suggest that optimally targeting supervisor visits can further increase weaver productivity by 3.4%-17.2% even after accounting for various operational and scheduling constraints.

While our empirical analysis is based on data from a specific context within the rug-weaving industry, our study offers valuable insights for other artisanal supply chains in geographically remote areas. The insights generated can directly benefit numerous platforms in developing countries with similar structures, such as Anou (https://www.theanou.com) in Morocco, Fab India (https://www.fabindia.com/) and Mahila Print (https://www.mahilaprint.com) in India, and Soko (https://www.shopsoko.com) in Kenya, that collectively employ millions of artisans. Further, our findings can also be informative for other distributed supply chains in resource-constrained settings that share similar characteristics. The impact of frequent and consistent supervisor visits observed in our study can potentially be extended to such settings, as the challenges faced by the artisans in our study (e.g., identifying and rectifying errors, learning tasks, geographical isolation) are similar to those faced by other smallholder supply chains.

The remainder of the paper is organized as follows. In §2, we discuss relevant literature. In §3, we describe the institutional setting and data in more detail. In §4, we describe the empirical analysis that includes our fixed effects specification as well as two instrumental variables approaches. In

§5, we present the results of the empirical analysis and explore cross-sectional heterogeneity in the impact of supervisor visits. In §6 we present our predict-then-optimize approach to optimize supervisor visits. §7 concludes the paper and identifies some future directions for research.

2. Literature Review

This paper focuses on improving operations in distributed artisanal supply chains of developing countries and makes important contributions to both research and practice. Our work lies at the intersection of three streams of literature.

The first stream is a growing body of operations management research focusing on social impact at the Base of the Economic Pyramid (BoP) (see Sodhi and Tang 2014, Kalkanci et al. 2019, Jónasson et al. 2019, Sunar and Swaminathan 2022, for recent reviews). Researchers have focused on nano-retail operations (Gui et al. 2019, Fransoo and Mora-Quiñones 2021, Escamilla et al. 2021, Fatunde et al. 2021, Acimovic et al. 2022), technology adoption (Guajardo 2019, Uppari et al. 2019, Kundu and Ramdas 2022, Ramdas and Sungu 2022), food and agriculture supply chains (Anupindi and Sivakumar 2007, de Zegher et al. 2018, Ganesh et al. 2019, Levi et al. 2020a,b, Peters et al. 2021, Adebola et al. 2022) and global health (Boutilier and Chan 2020, Gibson et al. 2020, De Boeck et al. 2022, Karamshetty et al. 2022). We add to this literature in two key aspects. First, we focus on a novel setting of artisanal supply chains that employ a major workforce (of mostly women) in many developing countries. Only a handful of studies (Plambeck and Taylor 2016, Chen and Lee 2017, Caro et al. 2021, Tuna and Swinney 2021, Alptekinoğlu and Örsdemir 2022) in operations management focus on textile supply chains. While the focus of these studies is on responsible sourcing, we focus on improving smallholder artisans' productivity in the supply chain. As advocated by Plambeck and Ramdas (2020), this work also helps in empowering rural women who form the majority of the workforce in artisanal supply chains. Second, we demonstrate the importance and value of adopting a field-based, multi-method approach for driving impact in practice. Using empirical analysis and a predict-then-optimize approach, our research shows that optimizing supply chain interventions can play a critical role in improving artisans' productivity and be a win-win for both manufacturers as well as artisans in developing economies.

The second stream of related literature focuses on improving worker productivity in supply chains. Researchers in organizational behavior and operations management have extensively studied key factors that affect worker productivity (see Diwas et al. 2020, for a recent review). Empirical evidence suggests that past experience (Gibbons and Waldman 2004, Lin et al. 2021), task variety (Ramdas et al. 2018), peer effects (Tan and Netessine 2019), team structure (Akşin et al. 2021), feedback mechanisms (Staats et al. 2018, Song et al. 2018, Kotiloglu et al. 2021), feedback specificity (Goodman et al. 2004), leadership (Giardili et al. 2023) and exposure to exports (Atkin et al.

2017) are all important factors that affect workers' productivity.[1] Closest to our setting is Atkin et al. (2017) and Bloom and Van Reenen (2007) who analyze productivity outcomes for rug manufacturers in Egypt, and textile manufacturers in India respectively using RCTs. However, there are major differences in the insights generated from our work. First, while we analyze distributed artisanal supply chains where artisans are located in geographically distant locations, these authors focus on textile manufacturing units (similar to standard factories) where textile workers assemble to produce rugs. Second, our focus is on improving artisan productivity by optimizing internal operations in the supply chain in contrast to learning by exports. We identify the causal impact of supervisor visits on artisan productivity and develop a data-driven optimization-based approach to route supervisor visits built on these insights.

Finally, the third stream of related literature focuses on using machine-learning and data-driven tools in operations management problems (Mišić and Perakis 2020, Baardman et al. 2023). Common applications include problems in revenue management (Ferreira et al. 2016, Cohen et al. 2017), inventory and supply chain management (Mehrotra et al. 2011, Gallien et al. 2015) and healthcare (Chan et al. 2012, Deo et al. 2015). We leverage the popular *predict-then-optimize* framework to optimize supervisor operations. Our underlying optimization problem is a supervisor scheduling-and-routing problem that is similar to the classic prize-collecting traveling salesman problem (PCTSP) (see Balas 1989, Feillet et al. 2005, Archetti et al. 2014, for detailed problem definition). Our optimization problem is most closely related to a variant of the prize-collecting TSP, the *Orienteering Problem* (OP), and its generalization the *Team Orienteering Problem* (TOP), where a single agent (OP) or multiple agents (TOP) travel on the network and construct route(s) that maximize(s) the profits of the trip subject to a fixed budget of costs (Tsiligirides 1984, Golden et al. 1987, Vansteenwegen et al. 2011). Because both the Orienteering Problem and the Team Orienteering Problem are NP-hard (Golden et al. 1987), past studies have focused on developing approximation algorithms with provable performance guarantees (Bansal et al. 2004, Chekuri et al. 2012, Xu et al. 2020, Paul et al. 2022). We show the equivalence of our weekly supervisor scheduling and routing optimization problem with the TOP problem. This equivalence allows us to leverage existing literature (Xu et al. 2020) to develop a polynomial run-time algorithm with a provable guarantee to solve our problem efficiently.

[1] Another stream of literature includes empirical work that analyze ways to conduct quality-control inspections. Current research highlights the importance of well-designed inspection schedules, accounting for factors such as worker fatigue, the timing between inspections, and investigator experience to effectively promote quality and safety improvements across industries (Staats et al. 2017, Ball et al. 2017, Ibanez and Toffel 2019). While the focus of this stream is on inspections, supervisors in our setting focus on multiple tasks, such as motivating workers, monitoring progress, and identifying *inadvertent* weaving errors.

3. Institutional Setting and Data

In order to fully appreciate the operational context, we made a series of field visits (see Appendix §O.1 for more details on field visits) and interacted with different stakeholders in Jaipur Rugs' supply chain. Based on these insights, we first describe the rug-making operations in more detail (also illustrated in Figure 1). Because Jaipur Rugs is vertically integrated, it oversees all tasks (weavers' training, raw material procurement, finishing, weaving, and marketing) in the supply chain. First, raw wool is spinned, hanked, and dyed into different colors. New designs for different rugs are produced by the design team based on market demand and historical trends. A design map along with threads of different colors is supplied to the weavers. Based on a design map, trained weavers then start weaving rugs on handlooms in individual homes. After rugs are fully woven, they are sent to the finishing center where dust and minute fibers are removed. Finally, these rugs are marketed and sold through different channels to local customers as well as exported to other countries. Our research focus is on handloom weaving operations (Step 6 in Figure 1) which is described in more detail next.

Figure 1 **Rug Making Operations**

Note. The supply chain starts with hanking of the raw wool and ends with selling finished rugs through multiple channels including exports, online retail, and showrooms. The focus of this research is rug weaving by artisans (labeled Step 6). Some photos for this figure are reproduced from Ramin (2021).

3.1. Weaving Operations

Weaving handloom rugs is an intricate and intensive task. The handlooms are run without electricity, using the traditional weaving method and the weaving involves constant weaver attention. All types of weaves involve three main actions (shedding, picking, and battening). For details on each of these actions, we refer the interested readers to Taylor (2017). The company compensates the weavers using a piecework scheme that varies based on the knot density. For each knot density, a per square feet rate is used and weavers are paid monthly based on the amount of weaving they finish in that particular month. The company has a significant backlog of customer orders and also maintains an inventory of popular designs. As a result, it never runs out of weaving tasks for weavers. As such, increased productivity for weavers directly translates to increased income in our setting.

Various rug-level features affect weaving times, including the textile used, the density of knots, design, size, and colors. For instance, greater weaving density allows the rugs to become finer and enables more delicate weaving patterns and thus take longer to weave. Next, during family events and harvesting months, weaving speed is similarly impacted because weavers are busy with other activities. A key challenge that also hampers weavers' productivity is inadvertent errors while weaving these rugs. Weaving is a labor-intensive task and it is common for even the most experienced weavers to make errors. For instance, a common mistake discussed during our field visits is slanted weaving. Once it is caught, weavers have to spend many hours rectifying the errors. Further, since weavers are paid for the incremental weaving completed in a month, and not for the reweaving, these errors reduce weavers' productivity. Thus, if not caught on time, these errors can be costly as they directly affect the income of the weavers. While many standard factories implement Poka-yoke (mistake-proofing) methods to prevent production errors (Widjajanto et al. 2020), implementing such methods in our setting is not straightforward. Because of a high level of heterogeneity in both weavers' skills and rug characteristics, identifying such simple techniques for every new rug design is not feasible.[2]

3.2. Supervisor Visits

Unlike standard factories, weaving operations at Jaipur Rugs are highly distributed and spread across a large geographical area (the maximum distance between any two weavers in our dataset is in the order of hundreds of kms). As a result, Jaipur Rugs has invested significant resources to reach and communicate effectively with weavers. The geographical area with looms is divided into multiple different branches. On average, each branch has hundreds of weavers and the branch manager is responsible for assigning and monitoring the weaving progress for his respective looms. For each branch, Jaipur Rugs maintains a list of rugs that need to be assigned to weavers. Every time a weaver is done weaving the current rug, the branch manager assigns her the next rug from this list of rugs if/when she becomes available. Once the new rug assignment is made, the weaver

[2] Indeed, Jaipur Rugs has implemented (before 2016) some operational interventions, with limited success, that are preventive in nature (e.g., drawing straight lines with special blue ink to avoid slanted weaving). While our interactions suggest that errors have decreased after such interventions, a significant number of errors continue to be identified by supervisors during their visits.

is provided with the design map of the rug as well as threads of different colors that are to be used in the rug.

To assist the branch managers in monitoring progress and minimizing weaving errors, each branch also has multiple quality supervisors who are tasked with physically visiting weavers at regular intervals. Since weavers are located in remote locations, supervisors act as a major node of communication between the company and the weavers. Supervisors are hired from local communities, paid a fixed monthly income, and treated as regular employees of the company. Every supervisor is pre-assigned a set of looms and is tasked with monitoring the progress of weaving in those looms. Since the company doesn't require supervisors to follow exact visit schedules, supervisors currently plan their own schedules based on their individual preferences. Each loom visit lasts about 15-20 minutes and involves a combination of tasks. First, the supervisors maintain a log and update the amount of work done by the weaver since his last visit. Next, they carefully inspect the rug to ensure that there are no weaving errors. Finally, if the weaver finishes the current rug, the supervisor ensures that the raw materials and design map for the next rug are sent to the weaver.

Our objective is to find the causal impact of supervisor visits on weavers' productivity. Supervisors in our setting are mandated to perform multiple tasks when they visit a weaver. Consequently, it is uncertain whether frequent visits by supervisors will result in an improvement in the productivity of weavers. If supervisors concentrate their efforts on motivating weavers and identifying any weaving errors during their visits, it is reasonable to assume that productivity will be positively impacted. This is because errors that are identified early on in the process require less reweaving, whereas errors detected later necessitate a greater amount of reweaving. Therefore, if supervisors assist in identifying errors, frequent visits can enhance productivity by reducing the amount of reweaving required for each rug. However, visits may have no effect on productivity if supervisors primarily focus on monitoring progress and updating logs, or if such errors are already minimal in the supply chain. Finally, frequent supervisor visits may actually lower weavers' productivity if weavers start to focus excessively on avoiding weaving errors or if the supervisor visits make them anxious (Gusnard 2005, George 2007, Pierce and Aguinis 2013). To summarize, it is not obvious a-priori whether more frequent supervisor visits will increase weavers' productivity in the supply chain.

3.3. Data and Variables

We obtain the following proprietary datasets from our collaborator. First, for all branches in the state of Rajasthan, we have a rug-level dataset that consists of extensive covariate information for all rugs that were assigned to weavers between 2017-2021. We observe the date on which the rug was assigned, the loom to which the rug was assigned, the size (in feet) of the loom, and

Table 1 Summary of Variables

	Explanation	Mean	Min	Max	SD
Weaving Time (days per rug)	Number of days taken to complete a rug	68.9	21	151	25.8
Avg days between visits (days)	Average number of days between supervisor visits per rug	9.0	2.8	19.7	4.5
Product Cubage (sq. feet)	Cubage of the rug	85.3	23.7	189.8	35.8
Total Colors	Total colors in the rug	7.4	1	25	4.1
Weaving Density(knots per sq. inch)	Density of knots weaved in a rug	63.7	36	121	21.78
Avg Temperature (K)	Average temperature across days during weaving per rug	299.1	285.1	310.1	5.8
Distance (kms)	Driving distance between weaver and supervisor pair	18.5	0.09	53.8	13.5
Looms per supervisor	Number of looms assigned to each supervisor	46	21	75	20
Looms per village	Number of looms in each village	6	1	51	8

Note. The total number of supervisors in the dataset is 20, the total number of villages in the dataset is 50, and the total number of rugs in the dataset is 8,061.

the weaving time (in days) to complete the rug. In addition, we also observe the following rug characteristics: design code of the rug[3], size of the rug, the textile used in the rug, length and width of the rug, number of colors in the rug, weaving density in the rug and unique ID of the supervisor who monitored the rug. Second, we obtain a supervisor visit dataset that contains the exact dates on which the assigned supervisor visited the rug during the course of its completion. Using this data, we calculate the average number of days between consecutive visits by the supervisor for each rug. We also obtain the geo-location of all installed looms and supervisors' home locations. We leverage open-sourced Google Map API to query the driving distances between the supervisor's home and his assigned loom locations. This driving distance measure is used to do our IV analysis which is discussed in more detail in §4.3. Finally, we collect satellite-based daily temperature data from National Oceanic and Atmospheric Administration (NOAA) for all loom locations since temperature may also affect productivity (Burke et al. 2015) and visit schedules of the supervisors. To identify the impact of supervisor visits on weaving times, we aggregate the data at the rug level. We choose this level of aggregation because the supervisor visits each rug multiple times in our dataset and it is difficult to control for spillover effects on productivity from consecutive visits for the same rug. Nevertheless, we confirm that our results remain consistent when we use individual visit level datasets (see §5.3 for more details). A summary of key variables used in our analysis is provided in Table 1.

[3] Weavers follow a design map when weaving rugs, and each design map has a unique design code of the rug.

4. Empirical Analysis

We start this section by providing some model-free evidence from the collaborator's data. Let I_{rl} be the average number of days between visits for rug r assigned on loom l. Rugs with the above-median average number of days between visits take 71 weaving days on average for completion while those with the below-median average number of days between visits take 67 weaving days on average for completion and this difference is statistically significant (t-test, $p = 0.0001$). Figure O.1 in Appendix O.2 illustrates this insight. This result provides some model-free evidence that more frequent supervisor visits (equivalently, lower average number of days between visits) may reduce weaving times and thereby increase weavers' productivity.

4.1. Fixed Effects (FE) Specification

We now turn to a regression framework to estimate the impact of supervisor visits on weaving times. Our base specification is a fixed-effects model with an extensive set of controls to account for factors that affect weaving times. We estimate the following model as our main FE specification:

$$\log(W_{rlt}) = \beta_1 I_{rl} + \beta \mathbf{X}_{rlt} + \phi_l + \delta_s + \psi_h + \omega_y + \epsilon_{rlt}. \tag{1}$$

$\log(W_{rlt})$ is the logarithm of the total weaving days taken to complete rug r on loom l assigned in time t. I_{rl} is the average number of days between supervisor visits to rug r on loom l. $\mathbf{X}_{rlt}$ includes the following time-invariant and time-variant rug and loom-level attributes. The first set of controls is rug-level features that affect weaving times. These include time-invariant variables such as the total number of colors in the rug, the weaving density of the rug, the total cubage of the rug, and the daily temperature at loom location l averaged across all days on which rug r was actively being weaved. Finally, we control for the weaver experience in the following manner. We compute a new variable, E_{rlt}, that is the log transformation of the total cubage weaved by the loom l until the date of the assignment t of the focal rug r, and encode all rugs for which E_{rlt} is below (above) the median of the distribution of E_{rlt} as "low-experience" ("high-experience"). We add this experience dummy in our specification. In §5.3, we consider alternate experience definitions and confirm consistent results.

In addition to these controls, we also add loom, supervisor, harvesting season, and year fixed effects (ϕ_l, δ_s, ψ_h, ω_y). Loom and supervisor fixed effects control for idiosyncratic differences between weaving speeds and monitoring quality of weavers and supervisors respectively. Harvesting season and year fixed effects adjust for seasonality and trends.[4] Finally, ϵ_{rlt} is the idiosyncratic shock to total weaving days for rug r on loom l at time t. Standard errors are clustered at the loom level to account for correlation across observations from the same loom. Our key variable of interest in this specification is I_{rl}.

[4] Our results remain consistent if we control for month fixed effects instead of harvesting season fixed effects. See Appendix §O.3 for more details.

4.2. Identification Challenges

Although this fixed effects specification controls for observed and unobserved heterogeneity at the rug, loom, and supervisor levels, there are two econometric challenges. The first concern is reverse causality. While we are interested in assessing the impact of visit frequency on weaving times, it is plausible that if weaving times are longer than expected, supervisors increase their frequency of visits. Given that the weaving times are long and there are multiple visits for the same rug, we expect that our key variable of interest, "average days between visits" should not be significantly affected by this concern. However, if this concern is indeed significant, rugs with longer weaving times would in fact have more frequent visits and smaller average days between visits by the supervisor. Since this would lead to an underestimation of the true effect size, the estimated β_1 in Equation 1 is a conservative estimate of the true effect size. Note that the reverse causality concerns are most severe for rugs *after* they are delayed because supervisors may increase efforts on rugs that are already delayed to speed up the work. In addition to the IV analysis, we perform an additional robustness check (see §5.3) by re-estimating Equation 1 using data from all rugs *before* they were delayed and confirm consistent results.

The second key concern is that of unobserved confounders. Our model can still lead to biased estimates if there are unobserved variables that are correlated with both weaving time and average days between visits. One omitted variable is the weaver's expected household expenditure (e.g., educational and medical expenses) during different times of the year. An increase in expenditure may increase her intrinsic motivation to work and thus reduce weaving times. However, given the close social bonds in rural communities, if the supervisor anticipates increased motivation, he may reduce the frequency of visits to the weaver. Thus, household expenditure could be positively correlated to the average number of days between supervisor visits but negatively correlated to weaving times. The omission of household expenditure would thus lead to an underestimation of the true impact of supervisor visits on weaving times.

4.3. Identification using Instrumental Variables

To tackle the above challenges, we use the instrumental variable (IV) approach, which has been widely used in the empirical literature (Angrist and Pischke 2009). In order to find a good IV, two conditions must be met. First, the IV should be correlated with the endogenous variable (Relevance Condition). Second, the IV should be uncorrelated with the error term (Exclusion Restriction). Under these conditions, the IV is correlated with the dependent variable only through the endogenous variable. We use two types of IVs in our estimation.

The first IV is the distance between the supervisor's home location, and the location of the weaver working on rug r (d_{rls}). Each supervisor is hired from rural communities to manage a large geographical area leading to natural variation in his distance from weavers. Further, a larger distance to rug r should lead to less frequent visits by the supervisors due to the additional travel required to visit these rugs. Finally, distance should not directly affect the total weaving times, thus satisfying the exclusion condition. In line with our mechanism, we observe that the coefficient of d_{rls} is positive and statistically significant in the first-stage regression (see Table O.1). In addition, the Cragg-Donald F statistic is greater than 10 which verifies that the IV is not weak. Because the distance IV, d_{rls}, does not vary between the same loom-supervisor pair, we cannot control for loom fixed effects in this specification. Instead, we add additional controls at the loom level. In particular, we control for loom size and also add fixed effects for the district in which the loom is installed to control for district-level heterogeneity.

We further supplement this IV with a Hausman-type IV following the literature (Cameron and Trivedi 2005, Caro et al. 2021). In particular, we use the average number of days between visits by supervisor s to the previous rug $r-1$ on the same loom l, I_{r-1l} as an IV for I_{rl}. This IV should satisfy the relevance condition since supervisors prefer consistency in their schedules: the frequency of visits by the supervisor to the previous rug assigned on the same loom should be correlated to the frequency of visits for this focal rug. Further, this IV should satisfy the exclusion restriction under the assumption that visit frequency during previous rugs does not directly impact the weaving times of the focal rug after controlling for the weaver experience. Under this assumption, I_{r-1l} should be independent of the focal rug r's error term. In line with this hypothesis, we again find that the coefficient of I_{r-1l} is positive and statistically significant. Further, the Cragg-Donald F statistic verifies that the IV is not weak. These first-stage results for both IVs are reported in Tables O.1 of the online appendix.

5. Results

In this section, we first discuss the estimation results from the fixed-effects model and instrumental variables model (§5.1). Next, we examine cross-sectional heterogeneity in effect across rugs with different features (§5.2). Finally, we present results from multiple robustness tests that confirm the validity of our findings (§5.3).

5.1. Estimation Results

Table 2 contains the results of our analysis on the impact of the average number of days between visits on weaving times. Column (1) shows results from the FE model (Equation 1). We find that a one-day increase in average days between visits increases total weaving days by 2.8% ($p = 0.002$). Columns (2) and (3) show results with the distance and Hausman-type IVs respectively. The results are directionally consistent with the FE model. In particular, we find that a one-day increase in average days between visits increases total weaving times by 13% and 14% respectively. From all three estimates, we can conclude that supervisor visits have a significant positive impact on weaving times.

One noteworthy aspect that warrants additional discussion is the significantly magnified impact of supervisor visits on weaving times estimated from IV analysis. The difference suggests that the impact of supervisor visits is heterogeneous (Kundu and Ramdas 2022). This is because IV analyses estimate local average treatment effects (LATE), while the FE model estimates average treatment effect (ATE) (Angrist and Pischke 2009). The LATE is for "compliers", the subset of the population that is affected by the instrument in question. Nevertheless, consistent and statistically significant effects in all three models confirm the positive effects of supervisor visits on weaving times. The coefficient estimates for the rest of the control variables are consistent with our intuition. Rugs with larger product cubage, larger knot density, more colors in the design, and lower temperatures are all associated with longer weaving times.

Recall from Table 1 that rugs take 69 days on average to complete and supervisors visit every loom in 9 days on average in our dataset. Interpreting the results with the IV, this result suggests that all else equal, if supervisors instead visit every loom in 8 days on average, the total weaving time on average for rugs would decrease to 59 days. Assuming that the weavers can then work on additional rugs given that the company has enough demand and backlog for rugs, this directly translates to a substantial 15%-17% increase in monthly income for the weavers. Our field interactions suggest that supervisors do not influence weavers' working hours since most weavers already work full-time on the loom. This suggests that the estimated increase in monthly income is due to increased efficiency in weaving and would not compromise the artisans' welfare. Finally, increased weaver productivity can be a win-win proposition as customer satisfaction is expected to improve substantially if delays in delivering finished rugs to Jaipur Rugs' customers are reduced.

5.2. Heterogeneous Effects

The above analysis establishes that visits by quality supervisors lead to a significant reduction in weaving times. In what follows, we provide evidence that the impact of visits on weaving times also has cross-sectional heterogeneity.

5.2.1. Impact on Difficult-to-Weave Rugs We hypothesize that the impact of visits from quality supervisors will be greater for rugs that are more difficult to weave. A key objective of supervisor visits is to identify inadvertent errors during weaving. Once a mistake is identified, weavers often have to unthread the rug and reweave it from the point of the error. If visits are less frequent, the amount of work that needs to be redone (conditional on an error occurring) is more, and total weaving time also increases. While it would be ideal to analyze data on inadvertent errors identified during individual visits by supervisors, that data is unfortunately not collected by

Table 2 **Estimated Impact of Avg. Days Between Supervisor Visits on Log-transformed Weaving Times**

Variable	(1) Fixed Effects	(2) IV1 (Driving Distance)	(3) IV2 (Previous Visit)
Avg. days between visits	0.028***	0.131**	0.141***
	(0.002)	(0.043)	(0.022)
Product Cubage	0.005***	0.038***	0.004***
	(0.000)	(0.000)	(0.000)
Knot Density	0.006***	0.005***	0.005***
	(0.000)	(0.000)	(0.001)
Number of Colors	0.006***	0.005***	0.006***
	(0.001)	(0.002)	(0.002)
Avg. temperature	-0.007***	-0.013**	-0.017***
	(0.001)	(0.003)	(0.002)
Size of Loom	–	-0.034***	-0.037***
		(0.005)	(0.006)
Other Controls	Y	Y	Y
Observations	8,006	8,006	7,150

Notes. "–" means the variable is not present in the model. Standard errors (in parentheses) are clustered at the loom level. Other controls include supervisor, time, and experience fixed effects in all three specifications. In addition, we control for loom fixed effects in Column (1) and district fixed effects in Column (2) and Column (3). ***: $p < 0.01$; **: $p < 0.05$; *: $p < 0.1$.

our collaborator. Nevertheless, we collected survey data from the weavers on how supervisor visits affect their work. Out of the 1950 completed responses we received, 99.43% of the weavers find supervisor visits "help in identifying errors (slanted weaving, wrong threads, etc.)" (with 83.79% strongly agree and 15.64% somewhat agree). Our interactions with multiple branch managers and supervisors also suggest that the chances of such errors are more in rugs that are more difficult to weave. As such, rugs that are more difficult to weave should benefit more from supervisor visits and see a stronger impact on weaving times. Recall from §3.1 that rugs differ significantly with regard to weaving difficulty. The first feature that we use as a proxy for weaving difficulty is the number of colors in the design map. Our field visits and interactions with the weavers suggest that rugs with multiple colors are significantly more difficult to weave. This is because weavers have to switch between threads of different colors based on the design map during weaving. This switching between colors is difficult and it is easy to mistake between shades and use the wrong thread if weavers are not attentive. Thus, as the number of colors in the rug increases, weaving difficulty also increases. This difference is also reflected in our data. Rugs in the lowest quartile based on the number of colors take fewer days to complete than all other rugs and the difference is statistically significant (Wilcoxon rank sum test, $p = 0.0001$). The second feature that we use as a proxy for weaving difficulty is the weaving density or the number of knots per square inch. As the weaving density increases, rugs become finer, and more delicate patterns can be weaved on them. However, this increased density also increases weaving difficulty since the weaver needs to weave a higher number of knots in the same area. While rugs with smallest weaving density of 36 knots per square inch in our dataset take 57 days, those with the largest weaving density of 121 knots per square inch take 72 days and the difference is statistically significant (Wilcoxon rank sum test, $p = 0.0001$).

Following the literature (Parker et al. 2016, Zhang et al. 2018), we estimate the following model in order to test for differential impact from supervisor visits:

$$\log(W_{rlt}) = \beta_0 I_{rl} + \beta_1 I_{rl} \times Z_{rH} + \beta \mathbf{X}_{rlt} + \phi_l + \delta_s + \psi_h + \omega_y + \epsilon_{rlt}. \tag{2}$$

Using weaving density (number of colors) as a proxy for weaving difficulty, we define Z_{rH} as an indicator variable that is equal to 1 if weaving density (number of colors) in the rug is greater than or equal to 64 knots per sq. inch (in the top three quartiles) and 0 otherwise in Equation 2.[5] All other variables remain the same as before. The baseline group in the specification represents rugs with lower weaving difficulty. Thus, coefficient β_1 captures the additional impact of supervisor visits on rugs with more weaving difficulty. Since we already control for both the weaving density as well as the total number of colors in the rug, we do not control for Z_{rH} directly in this specification.

The regression results for the two proxies are presented in Columns (1) and (2) of Table 3 and demonstrate that supervisor visits have a stronger positive impact on rugs with greater weaving difficulty. We observe that the coefficient for the interaction term in row 2 (β_1) is positive and statistically significant in both the models. Interpreting estimates from the IV analysis in Column (1), a one-day increase in average days between the visits increases total weaving times by an additional 0.6% for rugs with higher weaving density (or a 27% relative increase over rugs with lower weaving density). Note that the heterogeneous impact for rugs with additional colors is also significant but much smaller in magnitude (7% relative increase over rugs with fewer colors). This result suggests that errors associated with higher weaving density are more common and may benefit more from increased supervision than errors related to color mismatches. We confirm that these results remain consistent with IV analysis. Overall, the results highlight that the impact of supervisor visits is heterogeneous, and rugs with greater weaving difficulty benefit more from supervisor visits.

5.2.2. Impact from Consistency in Visits One aspect that may additionally influence the effectiveness of supervisor visits is the consistency in their visit schedule to the focal rug. To measure consistency in supervisor visits, we use the variance of the number of days between each supervisor visit for a rug. A low (high) variance in the number of days between the visits implies that the supervisors maintained a consistent (inconsistent) schedule while the weaver was working

[5] The results remain consistent if we instead use median of the distribution of number of colors to classify rugs with higher weaving difficulty.

on the rug. We hypothesize that in addition to frequent visits (with a lower average number of days between visits), visits that are consistent (with a lower variance in the number of days between visits) are most effective in improving weavers' productivity. Since most of the weavers are from rural households and have known supervisors for years, supervisors do not announce their visits in advance. It is worth noting that there is a significant body of literature (Van Loocke and Put 2011) that suggests that surprise (or equivalently, inconsistent) audit schedules may have a greater impact on improving performance and quality standards. However, in contrast to the extant literature which focuses on strategic violations that need to be detected, supervisors in our case primarily focus on inadvertent errors as there is no incentive for weavers to deliberately make these errors. As supervisors consistently visit weavers over time, weavers may come to expect regular intervals of supervision for the focal rug. This expectation may motivate weavers to complete the rug in a timely manner as they anticipate the supervisor to consistently monitor their progress. Finally, if the supervisor visits the weaver frequently *and* consistently, he may be able to detect weaving errors sooner, which can further help to reduce the reweaving costs. We estimate the following model in order to test for the differential impact from consistent supervisor visits:

$$\log(W_{rlt}) = \beta_0 I_{rl} + \beta_2 Z_{rH} + \beta_1 I_{rl} \times Z_{rH} + \beta \mathbf{X}_{rlt} + \phi_l + \delta_s + \psi_h + \omega_y + \epsilon_{rlt}. \tag{3}$$

Using variance in the number of days between visits as a measure for consistency, we define Z_{rH} in Equation 3 as an indicator variable that is equal to 1 if the variance in visits for a rug is greater than the median of variance in visits over all observations in our data and 0 otherwise. All other variables remain the same as before. Thus, coefficient β_2 captures the difference in weaving times between rugs with above-median and below-median variances in visit schedules, coefficient β_0 captures the impact from increasing average days between visits for rugs with below-median variance, and coefficient β_1 captures the incremental effect from increasing the average days between visits for rugs with above-median variance in the number of days between visits.

The regression results from Equation 3 are given in Column (3), Table 3 and demonstrate that consistent visit schedules can further benefit the weavers by improving their productivity and reducing weaving times. We observe that coefficient β_2 in row 3 is positive and statistically significant, suggesting that the average weaving time in rugs with inconsistent visit schedules is higher and the difference is statistically significant. Further, the coefficient for the interaction term in row 2 (β_1) is positive and statistically significant. Interpreting estimates from the analysis, a one-day increase in average days between visits is associated with an additional 0.6% increase in the total weaving time when supervisor schedules are inconsistent (or a 19% increase relative to the rugs with consistent visits). We confirm that this result remains consistent with the IV analysis. This result suggests that supervisor visits that are regular and consistent are most beneficial in improving the weavers' productivity.

Table 3 Estimated Differential Impact of Avg. Days Between Supervisor Visits on Log-transformed Weaving Times

Variable	(1) Knot Density	(2) Number of Colors	(3) Consistency in Visits
Avg. days between visits	0.022***	0.027***	0.031***
	(0.003)	(0.002)	(0.002)
Avg. days between visits $\times Z_{rH}$	0.006***	0.002*	0.006**
	(0.002)	(0.001)	(0.002)
Z_{rH}	–	–	0.069***
			(0.001)
Product Cubage	0.005***	0.005***	0.005***
	(0.000)	(0.000)	(0.000)
Knot Density	0.006***	0.006***	0.006***
	(0.000)	(0.001)	(0.001)
Number of Colors	0.007***	0.005***	0.007***
	(0.001)	(0.001)	(0.001)
Avg. temperature	-0.007***	-0.007***	-0.006***
	(0.001)	(0.001)	(0.001)
Other Controls	Y	Y	Y
Observations	8,006	8,006	8,006

Notes. "–" means the variable is not present in the model. Standard errors (in parentheses) are clustered at the loom level. Other controls include loom, supervisor, time, and experience fixed effects. ***: $p < 0.01$; **: $p < 0.05$; *: $p < 0.1$.

5.3. Robustness Tests

We perform a series of robustness checks to further strengthen our results. Our results remain consistent under, (i) alternative dependent variable; (ii) alternative experience definitions; (iii) alternative data preparation; (iv) additional controls; (v) village-level aggregation to account for spillover concerns; (vi) alternative clustering; (vii) data filtering to tackle COVID-19 interference; (viii) using weavers' payment data as an alternate data source; (ix) redoing the analysis at an individual visit level. Appendix O.3 presents the details of these analyses.

6. Framework for Optimizing Supervisor Visits

Following results from §5, one strategy to increase productivity could be to simply increase supervisor visits for all the weavers. However, uniformly increasing visits for all looms is operationally infeasible without hiring more supervisors.[6] Instead, an alternate strategy could be to selectively increase visits for rugs assigned to weavers with low productivity. To implement such a targeted supervision strategy, two steps are necessary: (i) *prediction* of whether a rug will have lower-than-expected productivity; (ii) *optimization* of supervisor schedules to ensure additional visits to rugs

[6] In exploring potential interventions to increase supervision, we considered the possibility of leveraging digital tools to supervise weavers remotely. However, we identified two significant challenges that may limit the effectiveness of this approach. First, the majority of the weavers in Jaipur Rugs' supply chain use basic phones that lack video capabilities, making it difficult to conduct digital visits. Second, our field visits revealed that most households share a single mobile phone, which is usually controlled by male members, whereas the majority of the weavers in the network are women. These challenges could impede the adoption and effectiveness of digital tools for remote supervision. Nevertheless, this direction remains an important avenue for future research.

predicted to have lower-than-expected productivity. In what follows, we discuss each of these steps in detail.

6.1. Predicting Rugs with Lower Productivity

In order to target supervisor visits, we need a forecast of whether rug r will have lower-than-expected productivity or not. As previously discussed, we have access to very detailed data on past rugs, along with information on baseline weaving days, as well as the realized weaving days for every rug. We leverage this data to formulate a supervised learning task and then use state-of-the-art ML methods to predict which rugs will have lower-than-expected productivity at any instance. Rugs with lower-than-expected productivity are simply defined as those rugs whose weaving days are higher than the baseline. More formally, let W_r denote the total weaving days of rug r and H_r denote the baseline weaving days for rug r. Note that H_r depends on rug-level features (such as weaving density, and product cubage) and is pre-decided by Jaipur Rugs. A rug is classified to have lower-than-expected productivity if $W_r > H_r$. We let Y_r be 1 if the rug has lower-than-expected productivity and 0 otherwise. [7]

Then the objective of the predictive task is to estimate a mapping function $f(X_r): \mathbb{R}^d \to \mathbb{R}$ that predicts the probability of lower-than-expected productivity for rug r with features $X_r \in \mathbb{R}^d$. To estimate such a function f, one can solve the following error-minimization problem

$$\min_{f \in \mathcal{F}} \sum_{r \in [N]} (Y_r - f(X_r))^2 ,$$

where $\mathcal{F}$ is the class of predictive functions (for e.g., generalized linear model, decision trees, random forests, etc.) and N is the total number of rugs in the data set. Naturally, the lower the error, the better does the model fit for the available data. Nevertheless, to ensure good out-of-sample performance, controlling over-fitting is necessary. In Appendix O.5, we briefly discuss, (i) the feature engineering for this prediction task, (ii) the predictive accuracy of different machine learning models, (iii) the train-test-validate framework (to control for over-fitting).[8]

We compare different algorithms based on four popularly used metrics for classification: (i) accuracy (ii) balanced accuracy (iii) ROC AUC and (iv) F1 score. Appendix O.4 describes each of these metrics in more detail. Table O.5 in Appendix O.5 presents results from different benchmark methods on these metrics. We find that the XGBoost classifier outperforms all other methods on all metrics with an out-of-sample balanced accuracy of 0.71 showing robust predictive performance. Hence, we use this method to make out-of-sample predictions for rugs.

[7] Using this classification rule, we find that 57.31 % of the rugs have lower-than-expected productivity in the historical data.

[8] Note that we use outcome predictions to *target* supervisor visits instead of directly using the empirical estimates identified in §4. This is motivated by recent research that has shown that in several cases simple outcome prediction model is preferable to using HTE estimation when optimizing treatment assignment (Fernández-Loría and Provost 2022).

6.2. Optimizing Supervisor Visits

Next, we consider an optimization framework to optimize the routes of existing supervisors and generate schedules that target rugs predicted to have lower-than-expected productivity.[9]

6.2.1. Model and Notation Consider a rug r located in village v_j that is predicted to have lower-than-expected productivity based on the machine learning model.

To accomplish these additional supervisor visits, we let Ω_j^T be the pre-decided target number of supervisor visits for a village v_j over the next T days (for e.g., 1 week), given by

$$\Omega_j^T := \begin{cases} \bar{c} & \text{if there exists a rug in village } v_j \text{ with lower-than-expected productivity,} \\ \underline{c} & \text{otherwise.} \end{cases} \tag{4}$$

We do not optimize the number of target visits $(\underline{c}, \bar{c})$ but assume it to be given in this formulation because of operational considerations. Based on our discussions with the management team, suggesting highly variable visits for each supervisor-loom pair may raise fairness concerns and make proposed solutions difficult to implement. As a result, the chosen target function in Eq. 4, (i) uniformly increases visits to $\bar{c}$ for all rugs that are predicted to have lower-than-expected productivity, (ii) assigns target visits at village level and assumes that all looms in a village will be visited by the supervisor every time he makes a visit to a village. Nevertheless, since a continuous objective that optimizes the number of visits directly based on the productivity predictions could be of independent interest, we explore this direction further in Appendix O.7. As discussed, such an objective leads to highly variable visits, both across different villages, and over time.

Since none of the supervisors share any villages, the optimization problem can be solved separately for each supervisor. We aim to generate a supervisor route that minimizes the difference between target visits ($\boldsymbol{\Omega}$) and actual visits by the supervisor. More specifically, let y_j^t be a binary decision variable that is 1 if village v_j is visited on day t, and 0 otherwise. Then, we seek to optimize the following objective: $\min_y \sum_{v_j} |\Omega_j^T - \sum_t y_j^t|$. To introduce other operational constraints that affect supervisor schedules, we define additional notation. Let ν denote the average travel speed (in kilometers per hour) of the quality supervisor, and τ denote the time (in hours) the supervisor spends inspecting each loom. Let n_j be the number of rugs being weaved in village v_j, and $T_{\max}$ be an upper bound on the maximum number of hours that the supervisor can work on any given day. In our case, $T_{\max}$ ranges from 5-7 hours. Let V denote a set (or a subset) of villages assigned to a supervisor, along with the supervisor's home village. Let $G(V)$ represent a fully-connected

[9] Another intervention that can be effective in increasing supervisor visits is to hire more supervisors. We provide a detailed analysis of this intervention in Appendix §O.6. Our focus on optimizing the routes of current supervisors is motivated by our interactions with the management team at Jaipur Rugs, who informed us that recruiting competent supervisors can be challenging due to the specialized skills and experience required. In fact, supervisors in the network have an average experience of more than 5 years, which underscores the difficulty of finding suitable candidates.

undirected graph over the set of villages V. An edge connecting villages (v_i, v_j) on this graph has a weight equal to the shortest driving distance between the two villages (D_{ij}). Let x_{ij}^t be a binary decision variable that is 1 if arc (v_i, v_j) is traversed on day t, and 0 otherwise.

We are now in a position to present the Optimal Supervisor Visit Problem (OSVP) given by,

$$
\begin{aligned}
\mathbf{OSVP}(\boldsymbol{\Omega}) = \min_{\mathbf{x},\mathbf{y},\mathbf{w}} \quad & \sum_{v_j} w_j && \text{(5a)}\\
\text{s.t.} \quad & \sum_t y_j^t \geq \underline{c}, \ \forall\, v_j && \text{(5b)}\\
& \sum_t y_j^t \leq \bar{c}, \ \forall\, v_j && \text{(5c)}\\
& \sum_{(v_i,v_j)} \left(\frac{D_{ij}}{\nu} + \tau \cdot n_j\right) \cdot x_{ij}^t \leq T_{\max}, \ \forall\, t && \text{(5d)}\\
& \sum_{\substack{(v_i,v_j)\\ \text{for } v_j \neq v_i}} x_{ij}^t = y_i^t, \ \forall\, v_i, t && \text{(5e)}\\
& \sum_{\substack{(v_j,v_i)\\ \text{for } v_j \neq v_i}} x_{ji}^t = y_i^t, \ \forall\, v_i, t && \text{(5f)}\\
& \sum_{\substack{(v_i,v_j)\\ \text{for } v_i \in S,\\ v_j \notin S}} x_{ij}^t \geq y_h^t, \ \forall\, S \subseteq V \backslash \{v_0\}, v_h \in S, t && \text{(5g)}\\
& \Omega_j^T - \sum_t y_j^t \leq w_j, \ \forall\, v_j && \text{(5h)}\\
& -\Omega_j^T + \sum_t y_j^t \leq w_j, \ \forall\, v_j && \text{(5i)}\\
& y_j^t, x_{ij}^t \in \{0,1\}, \ \forall\, v_i, v_j, t && \text{(5j)}\\
& w_j \geq 0, \ \forall\, v_j && \text{(5k)}
\end{aligned}
$$

Notice that $\mathbf{OSVP}(\boldsymbol{\Omega})$ optimizes both scheduling (y_j^t) and routing (x_{ij}^t) decisions of the quality supervisor. Constraints (5b) and (5c) guarantee that for each village, there are at least $\underline{c}$ supervisor visits and at most $\bar{c}$ supervisor visits across the planning time horizon T. Constraint (5d) is a maximum working time constraint on any day t. By limiting the maximum working time, we implicitly set an upper bound on the distance traveled by a supervisor, which is an important constraint to account for in practice (Cappanera and Scutellà 2015). Constraints (5e) and (5f) ensure that one arc enters and one arc leaves each visited village. Subtours are eliminated through constraint (5g). It indicates that if village $v_h \in S$ is visited, an arc necessarily leaves the set S, thus breaking the subtours within S (Feillet et al. 2005). Finally, constraints (5h) and (5i) link together the scheduling variables with the objective function.

The **OSVP** problem proposes a static optimization framework to find an optimized route for the next T periods (for example, the next week). This problem can be resolved every T periods (with updated predictions) to continuously optimize and update routes for the supervisors over time.[10] The **OSVP** problem is infeasible when the minimum number of visits to villages ($\underline{c}$) is very high relative to the number of time periods (T). Similarly, if the total time period (T) is very large relative to the maximum number of visits ($\bar{c}$), then the problem has a trivial optimal solution (visit one village every day). Nevertheless, in the intermediate ranges which is of interest in our problem setting, solving $\mathbf{OSVP}(\mathbf{\Omega})$ directly is computationally expensive. In particular, the formulation contains an exponential number of constraints (in $|V|T$, the product of the number of villages and time horizon), and $(|V|^2+|V|)T$ decision variables. Hence, in what follows, we draw the connection between $\mathbf{OSVP}(\mathbf{\Omega})$ and the well-studied prize orienteering problem (Archetti et al. 2014), and leverage a recent result on team orienteering problem to solve $\mathbf{OSVP}(\mathbf{\Omega})$ in polynomial time with a provable performance guarantee.

6.2.2. Team OP Based Approach for Solving $\mathbf{OSVP}(\mathbf{\Omega})$ We begin with a brief description of a variant of the classical Team Orienteering Problem (TOP), which was first introduced by Butt and Cavalier 1994 and Chao et al. 1996. The problem considers K vehicles, $K \geq 1$ navigating on a graph $G(V)$ with profits pre-assigned at each village v_i, where the goal is to generate K routes (one route for each vehicle) that maximizes the total profits collected by all vehicles under global constraints of each vehicle route (for example, vehicle travel time). Note that each vehicle has the binary decision of whether to visit a village and each village can be visited by multiple vehicles. More specifically, we let $\Delta_i(n_i)$ be the profit of visiting village v_i, n_i number of times among the K vehicles, where $\Delta_i(\cdot)$ is a non-decreasing submodular profit function for village v_i that generates diminishing marginal returns with additional visits. Let $P_1, P_2 \dots P_K$ be the K paths for K vehicles, each starting and ending at a specified home location v_0, and we denote $\xi(P_k)$ as the cost for taking path P_k. Then the TOP problem with vehicle travel budget constraint B_k, $1 \leq k \leq K$ is given by

$$\mathbf{TOP}(\mathbf{\Delta}, G(V)) = \max \left\{ \sum_{v_i \in \bigcup_{k=1}^{K} P_k} \Delta_i(n_i) |\ \xi(P_k) \leq B_k,\ \forall 1 \leq k \leq K \right\} \tag{6}$$

Notice that both $\mathbf{OSVP}(\mathbf{\Omega})$ and the Team Orienteering Problem $\mathbf{TOP}(\mathbf{\Delta}, G(V))$ consider routing under travel budget, yet they differ from each other mainly in two aspects.

1. *Time horizon:* TOP considers a single time period problem, where the decisions are made for a single time step over K vehicle routes. In contrast, $\mathbf{OSVP}(\mathbf{\Omega})$ considers the problem of planning and scheduling over multiple time periods for a single quality supervisor, and total visits across all time periods are constrained based on $\mathbf{\Omega}$.

[10] While the current objective minimizes the difference between target and assigned visits, an alternate objective could also be to ensure that the proposed routes remain as close as possible to the existing routes that the supervisors use to visit different villages. Such an objective would simplify implementation in practice since supervisors would need to make minimal changes to their existing schedules. We can easily update the formulation to account for such a consideration. In Appendix O.8, we discuss and present this alternative formulation.

2. *Profits of village visits*: While $\mathbf{TOP}(\mathbf{\Delta}, G(V))$ problem maximizes profits defined by $\mathbf{\Delta}(\cdot)$ across K vehicles, the $\mathbf{OSVP}(\mathbf{\Omega})$ problem does not model explicit village-level profits. In particular, $\mathbf{OSVP}(\mathbf{\Omega})$ ensures that each village is visited at most once per day and the total number of visits is upper and lower bounded by $\bar{c}$ and $\underline{c}$ respectively.

Given these differences, how to use Team Orienteering Problem to solve $\mathbf{OSVP}(\mathbf{\Omega})$ is not obvious a-priori. In what follows, we identify a specific village-level profit function $\tilde{\mathbf{\Delta}}(\cdot)$ that establishes the equivalence between $\mathbf{OSVP}(\mathbf{\Omega})$ and $\mathbf{TOP}(\mathbf{\Delta}, G(V))$ problem under mild conditions.

Intuition: Recall that the objective of the $\mathbf{OSVP}(\mathbf{\Omega})$ problem is to match the number of visits to each village according to $\mathbf{\Omega}$. Hence, one intuitive strategy to define village-level profits is to simply let the profit collected from each visit to a village v_i as the difference between total visits to the village so far versus the number of visits prescribed by Ω_i^T. Note that the profit described above is a decreasing step function: as the number of visits to village v_i increases before surpassing Ω_i^T, the difference from its target visit decreases, leading to diminishing profitability for further visits to that village. Following this intuition, we formalize the idea of profit gain from each village visit and define a village-level non-decreasing submodular profit function.

Definition 1. For any village $v_i \in V$, the profit of each village visit $\delta_i(\cdot)$ is a function of unfulfilled target visits Ω_i^T. More specifically, $\delta_i(c_i) = (\Omega_i^T - c_i)^+$, where c_i is the number of times village v_i has been visited prior to the current visit, $0 \le c_i < \Omega_i^T$.

The total profit from visiting village v_i, n_i times, denoted as $\tilde{\Delta}_i(n_i)$ is thus

$$\tilde{\Delta}_i(n_i) = \sum_{c_i=0}^{n_i-1} \delta_i(c_i) = \sum_{c_i=0}^{n_i-1} (\Omega_i^T - c_i)^+, \tag{7}$$

where we have $\tilde{\Delta}_i(0) = 0$ and $n_i \ge 1$.

Note that for each additional visit, the marginal profit decreases by one unit, which characterizes the submodularity property of $\tilde{\mathbf{\Delta}}(\cdot)$, and the profit function $\tilde{\mathbf{\Delta}}(\cdot)$ is non-decreasing. Figure O.2 in Appendix O.2 demonstrates the definition of δ_i and $\tilde{\Delta}_i(\cdot)$. In Proposition 1 we show the existence of a computationally tractable algorithm for solving the supervisor scheduling and routing problem with a provable performance guarantee.

Proposition 1 *Assume the* $\mathbf{OSVP}(\mathbf{\Omega})$ *problem has feasible solutions and* $\underline{c}=0$*, then there exists a polynomial time algorithm to solve* $\mathbf{OSVP}(\mathbf{\Omega})$ *with a competitive ratio of* $1-(\frac{1}{e})^{\frac{1}{2+\epsilon}}$*, where* ϵ *is any fixed constant with* $0 < \epsilon \le 1$*.*

While the proof of the Proposition is relegated to Appendix O.10, we discuss the main intuition here. We prove Proposition 1 by showing the equivalence of $\mathbf{OSVP}(\mathbf{\Omega})$ and $\mathbf{TOP}(\tilde{\mathbf{\Delta}}, G(V))$. In particular, we prove that the village-level profit function $\tilde{\Delta}(\cdot)$ is non-decreasing and submodular and there exists an equivalence between $\mathbf{OSVP}(\mathbf{\Omega})$ and $\mathbf{TOP}(\tilde{\mathbf{\Delta}}, G(V))$. Leveraging results from Xu et al. (2020), there is a $(1-(\frac{1}{e})^{\alpha})$-approximation algorithm for $\mathbf{OSVP(\Omega)}$, assuming the existence of a α-approximation solution for the classic Orienteering Problem (Golden et al. 1987), $0<\alpha<1$. Applying the state-of-art $\frac{1}{2+\epsilon}$-approximation algorithm for the Orienteering Problem due to Chekuri et al. (2012), we prove that the supervisor visit optimization problem $\mathbf{OSVP(\Omega)}$ can be solved efficiently with a competitive ratio of $1-(\frac{1}{e})^{\frac{1}{2+\epsilon}}$.

6.3. Case Study: Impact of Targeted Supervisor Visits

In this section, we discuss our end-to-end predict-then-optimize approach and its impact on weaver productivity by applying it to the real-world data from our collaborator's supply chain.

Problem setting and baseline: Recall from §6.1 that we have access to detailed rug-level data which is used for the prediction task. As before, we split the data into training and testing sets where we keep the last three months of data on rugs that are actively being weaved in the test set. For all *active* rugs in the test set, we perform the following four steps in every one-week period to generate weekly schedules: (i) Use the ML model to predict whether each *active* rug in the test set will have lower-than-expected productivity or not. Subsequently, identify villages that have any rug which is predicted to have lower-than-expected productivity. (ii) Let the corresponding target visits by the supervisor for these villages, $\bar{c}$, be 2 over the next week. Otherwise, for all other villages let the target visits over next week, $\underline{c}$, be 1. (iii) Design schedules and routes for the next week for each supervisor using the optimization formulation described in §6.2. (iv) Update rug-level features using information generated by supervisor visits and start from Step (i) again.

Let LR_a (HR_a) be the set of rugs in the test dataset that had lower-than-expected (higher-than-expected) productivity. Similarly, let LR_p (HR_p) be the set of rugs in the test dataset that are *predicted* to have lower-than-expected (higher-than-expected) productivity. We compare the performance of the baseline policy of supervisor visits to two different policies that we discuss next.

1. *Policy MPTO (Proposed visits from ML prediction and optimization framework)* This policy outputs supervisor routes generated by the predict-then-optimize framework and we track the total number of visits it recommends for rugs in LR_a and HR_a separately as a metric of efficiency. Note that the predict-then-optimize framework is effective if a relatively higher number of visits is scheduled for rugs in the set LR_a. For the problem scale of interest to our collaborator, we are able to solve the exact **OSVP** problem on Lenovo SD650 with an Intel 2.9 GHz eight-core processor and 96 GB of RAM in 48.75 seconds on average. For large-scale problems where an exact IP solution is not feasible, one can leverage the polynomial-time algorithm discussed in §6.

2. *Policy NPTO (Proposed visits from naive prediction and optimization framework)* Although the previous metric tracks the predict-then-optimize framework, it does not reveal whether the ML-based prediction models are essential for improving performance over the baseline. To better

understand the value of ML-based predictions, we also compare the results to a naive prediction model. Specifically, we calculate the cumulative work done by the weaver during the previous two supervisor visits for each rug. If the amount of work done is lower (higher) than what it should be based on the standard weaving rate for a given rug r[11], we classify it to the set LR_p (HR_p), respectively. If the optimization framework with this naive prediction model performs equally well as the previous policy, it would suggest that the ML-based prediction model is not critical to our framework.

Estimated Improvement: In Table 4, we provide a simple comparison between the current route followed by Supervisor X (Policy Baseline, on the left) and our proposed visit route (generated by Policy MPTO, on the right). Note that villages with rugs in LR_a are marked in bold in Table 4 and ideally should be visited twice. Comparing the current and optimized routes, we observe that some daily tours are very similar, such as Day 3 of current routes and Day 6 of optimized routes (blue cells), and Day 4 of current routes and optimized routes (yellow cells). However, the optimized routes are more efficient in two ways: (i) optimizing the sequence of village visits (blue cells) and (ii) grouping together nearby villages to save time and effort (yellow and red cells). Although the proposed routes increase Supervisor X's workload by 13.24% (from 68 to 77 looms per week), they increase visits to villages with lower-than-expected productivity rugs from 57.14% to 85.71%. In particular, while current routes visit four out of seven villages with lower-than-expected productivity rugs twice, optimized routes visit six out of these seven villages twice. These results provide insights into the productivity improvement from an individual supervisor's perspective by comparing the changes from current to proposed routes.

Next, we analyze the improvement from our framework at an aggregate level. Figure 2 illustrates the overall effectiveness of the two policies vis-a-vis the current baseline policy. On the left (right) we plot the average weekly visits from the three policies across villages that had any rug with lower-than-expected (higher-than-expected) productivity in the test dataset. With regards to the effectiveness of the ML prediction and optimization framework (Policy MPTO), we find that average weekly visits to lower-than-expected productivity rugs (1.60 visits per week) is 33.65% more than the visits proposed on higher-than-expected productivity rugs (1.20 visits per week), thereby showing that the framework is able to appropriately target visits on lower-productivity rugs. Further, in comparison to the baseline policy, Policy MPTO increases weekly visits to lower-than-expected productivity rugs by 24.69% (from 1.29 to 1.60) while it decreases weekly visits to higher-than-expected productivity rugs by 8.40% (from 1.31 to 1.20). These results suggest that Policy MPTO is able to more efficiently allocate supervisor visits to looms that have lower-than-expected productivity. Comparing average weekly supervisor visits from naive prediction (Policy

[11] Standard weaving rate of rug r is defined as the product cubage of rug r divided by baseline weaving days H_r.

Table 4 Comparison of current weekly routes and the proposed routes for Supervisor X

	Current routes	Total looms	Distance (kms)	Optimized routes	Total looms	Distance (kms)
Day 1	village A, **village B**, **village C**	11	65.92	**village I**, **village E**	14	45.12
Day 2	**village D**, **village E**	11	53.78	**village C**, **village K**	13	62.78
Day 3	village F, **village G**, village H, **village I**	12	55.56	**village D**, **village E**, **village G**	12	54.52
Day 4	village J, **village K**, village L, village M	12	70.31	village M, **village K**, village J, village L, **village B**	13	66.09
Day 5	village A, **village B**, **village C**	11	65.92	village A, **village C**, **village D**	13	75.98
Day 6	**village D**, **village E**	11	53.78	**village G**, village F, village H, **village I**	12	49.46
Total	18 villages	68	365.27	19 villages	77	353.95

Note. Jaipur Rugs supervisors work for 6 days per week. The current routes (on the left) present a weekly schedule followed by a randomly chosen supervisor from the dataset, whereas the proposed routes (on the right) present the optimized schedule generated by Policy MPTO. Villages with rugs in LR_a are marked in bold. Cells of the same color follow similar daily routes but have improved efficiency in the proposed routes.

NPTO) with the baseline policy, we again find that even with naive prediction, the optimization framework is able to target rugs with lower-than-expected productivity more effectively. However, Policy MPTO is better able to utilize supervisor resources to target rugs with lower-than-expected productivity (1.60 visits versus 1.54 visits). Recall from §4.3 that reducing the average number of days between visits by a day can lead to a 2.8-14.1% decrease in weaving times. Using these estimates, we can calculate the effect on weaving times from increased visits on lower-than-expected productivity rugs under the MPTO Policy. Our estimates indicate that increased supervisor visits can decrease weaving times by 2.35-11.83 days per rug, which translates to a 3.4-17.2% increase in weaver productivity (recall 68.9 days for average rug weaving time). It subsequently reduces the number of rugs with lower-than-expected productivity by 1.70%-12.17% while increasing the supervisor visit consistency by 71.9% (variance of days between supervisor visits reduced from 9.10 to 2.56) (see Appendix §O.9 for more details on the improvement).

7. Conclusions

The artisanal supply chain plays a major role in the rural economy and is a major employer for women across the developing world. Nevertheless, the industry remains highly fragmented and a key challenge is the low productivity of artisans in the supply chain due to limited supervision. By collaborating closely with Jaipur Rugs, a major employer of smallholder weavers in India, we provide empirical evidence that frequent supervision can play an important role in improving artisans' productivity. Further, this impact is heterogeneous, and weavers working on difficult-to-weave rugs benefit the most from supervisor visits. Based on the empirical insights, we develop a novel predict-then-optimize framework for optimizing supervisor visits and perform numerical

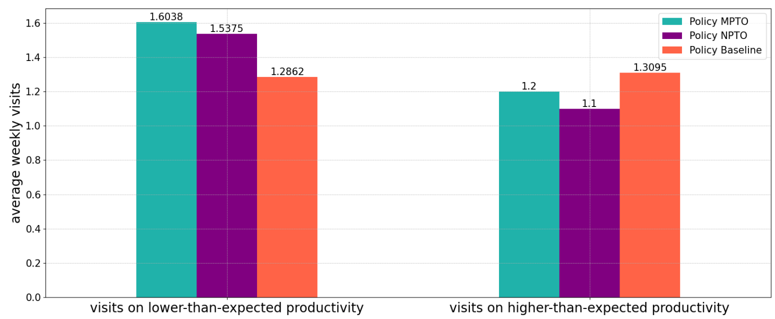


Figure 2 **Average weekly village visits for lower-than-expected productivity and higher-than-expected productivity rugs, Policy MPTO is in green, Policy NPTO is in magenta and Baseline Policy is in orange in the figure.**

experiments to show that this approach can considerably improve the productivity of rural artisans in the supply chain. We believe that the methods introduced in this paper can provide useful insights and tools for other researchers and practitioners to optimize supervision in other distributed supply chains in resource-constrained settings. This work also opens multiple avenues for future research. In particular, while we focus on optimizing supervisor visits for improving productivity, other important directions which can also benefit weavers include optimizing incentive contracts for smallholder weavers, as well as optimizing weaver-rug matching assignments taking into account their individual preferences and resource constraints. Finally, extending the predict-then-optimize framework to decision-aware learning (Kotary et al. 2021, Chung et al. 2022, Elmachtoub and Grigas 2022) and evaluating its impact could be another interesting direction for future research.

8. Acknowledgements

We are grateful to Nand Kishor Chaudhary and Nitesh Chaudhary for enabling this collaboration; to Amit Dagar and Surendra Dhakad for their tremendous support in the field; to all the weavers and other staff members at Jaipur Rugs for being so welcoming during our field trips, and to Benjamin Liu for his excellent research assistance. We gratefully acknowledge financial support from the Institute for Outlier Research in Business at the USC Marshall School of Business, the Lloyd Greif Center for Entrepreneurial Studies at the USC Marshall School of Business, and the NYU Stern Center for Sustainable Business Research Grant Program.

References

Acimovic, Jason, Chris Parker, David F. Drake, Karthik Balasubramanian. 2022. Show or tell? improving inventory support for agent-based businesses at the base of the pyramid. *Manufacturing & Service Operations Management* **24**(1) 664–681.

Adebola, Olufunke, Priyank Arora, Can Zhang. 2022. Farm equipment sharing in emerging economies. *Available at SSRN 4190725* .

Akşin, Zeynep, Sarang Deo, Jónas Oddur Jónasson, Kamalini Ramdas. 2021. Learning from many: Partner exposure and team familiarity in fluid teams. *Management Science* **67**(2) 854–874.

Alptekinoğlu, Aydın, Adem Örsdemir. 2022. Is adopting mass customization a path to environmentally sustainable fashion? *Manufacturing & Service Operations Management* .

Amabile, Teresa M, Regina Conti, Heather Coon, Jeffrey Lazenby, Michael Herron. 1996. Assessing the work environment for creativity. *The Academy of Management Journal* **39**(5) 1154–1184.

Angrist, Joshua D, Jörn-Steffen Pischke. 2009. *Mostly harmless econometrics: An empiricist's companion*. Princeton university press.

Anupindi, Ravi, S Sivakumar. 2007. Supply chain reengineering in agri-business: A case study of itc's e-choupal. *Building supply chain excellence in emerging economies*. Springer, 265–307.

Archetti, Claudia, M Grazia Speranza, Daniele Vigo. 2014. Chapter 10: Vehicle routing problems with profits. *Vehicle routing: Problems, methods, and applications, second edition*. SIAM, 273–297.

Atkin, David, Amit K Khandelwal, Adam Osman. 2017. Exporting and firm performance: Evidence from a randomized experiment. *The Quarterly Journal of Economics* **132**(2) 551–615.

Baardman, Lennart, Rares Cristian, Georgia Perakis, Divya Singhvi, Omar Skali Lami, Leann Thayaparan. 2023. The role of optimization in some recent advances in data-driven decision-making. *Mathematical Programming* **200**(1) 1–35.

Balas, Egon. 1989. The prize collecting traveling salesman problem. *Networks* **19**(6) 621–636.

Ball, George, Enno Siemsen, Rachna Shah. 2017. Do plant inspections predict future quality? the role of investigator experience. *Manufacturing & Service Operations Management* **19**(4) 534–550.

Banik, Subhamoy. 2017. A study on financial analysis of rural artisans in india: issues and challenges. *International Journal of Creative Research Thoughts (IJCRT)* **5**(4).

Bansal, Nikhil, Avrim Blum, Shuchi Chawla, Adam Meyerson. 2004. Approximation algorithms for deadline-tsp and vehicle routing with time-windows. *Proceedings of the thirty-sixth annual ACM symposium on Theory of computing*. 166–174.

Bloom, Nicholas, Benn Eifert, Aprajit Mahajan, David McKenzie, John Roberts. 2013a. Does management matter? evidence from india. *The Quarterly journal of economics* **128**(1) 1–51.

Bloom, Nicholas, Benn Eifert, Aprajit Mahajan, David McKenzie, John Roberts. 2013b. Does management matter? evidence from india. *The Quarterly Journal of Economics* **128**(1) 1–51.

Bloom, Nicholas, John Van Reenen. 2007. Measuring and explaining management practices across firms and countries. *The Quarterly Journal of Economics* **122**(4) 1351–1408.

Boutilier, Justin J, Timothy CY Chan. 2020. Ambulance emergency response optimization in developing countries. *Operations Research* **68**(5) 1315–1334.

Burke, Marshall, Solomon M Hsiang, Edward Miguel. 2015. Global non-linear effect of temperature on economic production. *Nature* **527**(7577) 235–239.

Butt, Steven E, Tom M Cavalier. 1994. A heuristic for the multiple tour maximum collection problem. *Computers & Operations Research* **21**(1) 101–111.

Cameron, A Colin, Pravin K Trivedi. 2005. *Microeconometrics: methods and applications*. Cambridge university press.

Cappanera, Paola, Maria Grazia Scutellà. 2015. Joint assignment, scheduling, and routing models to home care optimization: A pattern-based approach. *Transportation Science* **49**(4) 830–852.

Caro, Felipe, Leonard Lane, Anna Sáez de Tejada Cuenca. 2021. Can brands claim ignorance? unauthorized subcontracting in apparel supply chains. *Management Science* **67**(4) 2010–2028.

Chan, Carri W, Vivek F Farias, Nicholas Bambos, Gabriel J Escobar. 2012. Optimizing intensive care unit discharge decisions with patient readmissions. *Operations research* **60**(6) 1323–1341.

Chao, I-Ming, Bruce L Golden, Edward A Wasil. 1996. The team orienteering problem. *European journal of operational research* **88**(3) 464–474.

Chekuri, Chandra, Nitish Korula, Martin Pál. 2012. Improved algorithms for orienteering and related problems. *ACM Transactions on Algorithms (TALG)* **8**(3) 1–27.

Chen, Li, Hau L Lee. 2017. Sourcing under supplier responsibility risk: The effects of certification, audit, and contingency payment. *Management Science* **63**(9) 2795–2812.

Chen, Tianqi, Carlos Guestrin. 2016. XGBoost: A scalable tree boosting system. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. KDD '16, ACM, New York, NY, USA, 785–794. doi:10.1145/2939672.2939785. URL `http://doi.acm.org/10.1145/2939672.2939785`.

Chung, Tsai-Hsuan, Vahid Rostami, Hamsa Bastani, Osbert Bastani. 2022. Decision-aware learning for optimizing health supply chains. *arXiv preprint arXiv:2211.08507* .

Cohen, Maxime C, Ngai-Hang Zachary Leung, Kiran Panchamgam, Georgia Perakis, Anthony Smith. 2017. The impact of linear optimization on promotion planning. *Operations Research* **65**(2) 446–468.

De Boeck, Kim, Catherine Decouttere, Jónas Oddur Jónasson, Nico Vandaele. 2022. Vaccine supply chains in resource-limited settings: Mitigating the impact of rainy season disruptions. *European Journal of Operational Research* **301**(1) 300–317.

De Loecker, Jan, Chad Syverson. 2021. An industrial organization perspective on productivity. *Handbook of industrial organization*, vol. 4. Elsevier, 141–223.

de Zegher, Joann F, Dan A Iancu, Erica L Plambeck. 2018. Sustaining smallholders and rainforests by eliminating payment delay in a commodity supply chain-it takes a village .

Deo, Sarang, Kumar Rajaram, Sandeep Rath, Uday S Karmarkar, Matthew B Goetz. 2015. Planning for hiv screening, testing, and care at the veterans health administration. *Operations research* **63**(2) 287–304.

Diener, Ed, Robert Biswas-Diener. 2002. Will money increase subjective well-being? *Social indicators research* **57** 119–169.

Diwas, KC, et al. 2020. Worker productivity in operations management. *Foundations and Trends® in Technology, Information and Operations Management* **13**(3) 151–249.

Elmachtoub, Adam N, Paul Grigas. 2022. Smart "predict, then optimize". *Management Science* **68**(1) 9–26.

Escamilla, Rafael, Jan C Fransoo, Christopher S Tang. 2021. Improving agility, adaptability, alignment, accessibility, and affordability in nanostore supply chains. *Production and Operations Management* **30**(3) 676–688.

Fatunde, Olumurejiwa, Andre Calmon, Joann F de Zegher, Gonzalo Romero. 2021. The value of long-term relationships when selling to informal retailers-evidence from india. *Rotman School of Management Working Paper* (3792639).

Feillet, Dominique, Pierre Dejax, Michel Gendreau. 2005. Traveling salesman problems with profits. *Transportation science* **39**(2) 188–205.

Fernández-Loría, Carlos, Foster Provost. 2022. Causal classification: Treatment effect estimation vs. outcome prediction. *Journal of Machine Learning Research* **23**(59) 1–35.

Ferreira, Kris Johnson, Bin Hong Alex Lee, David Simchi-Levi. 2016. Analytics for an online retailer: Demand forecasting and price optimization. *Manufacturing & service operations management* **18**(1) 69–88.

Fransoo, Jan C, Camilo Andrés Mora-Quiñones. 2021. Decarbonizing urban logistics: Perspectives for low- and middle-income countries. *Available at SSRN 3943360* .

Gallien, Jérémie, Adam J Mersereau, Andres Garro, Alberte Dapena Mora, Martín Nóvoa Vidal. 2015. Initial shipment decisions for new products at zara. *Operations Research* **63**(2) 269–286.

Ganesh, Maya, Sarang Deo, Sripad K Devalkar. 2019. Leveraging digital technology to improve monitoring and planning in public sector supply chains: Evidence from india's food security program. *Indian School of Business* .

George, Jennifer M. 2007. Creativity in organizations. *The Academy of Management Annals* **1**(1) 439–477.

Giardili, Soledad, Kamalini Ramdas, Jonathan W Williams. 2023. Leadership and productivity: a study of us automobile assembly plants. *Management Science* **69**(3) 1500–1517.

Gibbons, Robert, Michael Waldman. 2004. Task-specific human capital. *American Economic Review* **94**(2) 203–207.

Gibson, Emma, Sarang Deo, Jónas Oddur Jónasson, Mphatso Kachule, Kara Palamountain. 2020. Redesigning sample transportation in malawi through improved data sharing and daily route optimization. *Available at SSRN 3712556* .

Golden, Bruce L, Larry Levy, Rakesh Vohra. 1987. The orienteering problem. *Naval Research Logistics (NRL)* **34**(3) 307–318.

Goodman, Jodi S, Robert E Wood, Margaretha Hendrickx. 2004. Feedback specificity, exploration, and learning. *Journal of Applied Psychology* **89**(2) 248.

Guajardo, Jose A. 2019. How do usage and payment behavior interact in rent-to-own business models? evidence from developing economies. *Production and Operations Management* **28**(11) 2808–2822.

Gui, Luyi, Christopher S Tang, Shuya Yin. 2019. Improving microretailer and consumer welfare in developing economies: Replenishment strategies and market entries. *Manufacturing & Service Operations Management* **21**(1) 231–250.

Gusnard, Debra A. 2005. Being a self: considerations from functional imaging. *Consciousness and cognition* **14**(4) 679–697.

Ho, Tin Kam. 1995. Random decision forests. *Proceedings of 3rd international conference on document analysis and recognition*, vol. 1. IEEE, 278–282.

Ibanez, Maria R, Michael W Toffel. 2019. How scheduling can bias quality assessment: Evidence from food-safety inspections. *Management Science* **65**(6) 2626–2647.

Jónasson, Jónas Oddur, Kamalini Ramdas, Alp Sungu. 2019. Social impact operations at the global base of the pyramid. *Production and Operations Management* .

Kalkanci, Başak, Morvarid Rahmani, L Beril Toktay. 2019. The role of inclusive innovation in promoting social sustainability. *Production and Operations Management* **28**(12) 2960–2982.

Karamshetty, Varun, Harwin De Vries, Luk N Van Wassenhove, Sarah Dewilde, Warnyta Minnaard, Dennis Ongarora, Kennedy Abuga, Prashant Yadav. 2022. Inventory management practices in private healthcare facilities in nairobi county. *Production and Operations Management* **31**(2) 828–846.

Ke, Guolin, Qi Meng, Thomas Finley, Taifeng Wang, Wei Chen, Weidong Ma, Qiwei Ye, Tie-Yan Liu. 2017. Lightgbm: A highly efficient gradient boosting decision tree. *Advances in neural information processing systems* **30** 3146–3154.

Kotary, James, Ferdinando Fioretto, Pascal Van Hentenryck, Bryan Wilder. 2021. End-to-end constrained optimization learning: A survey. *arXiv preprint arXiv:2103.16378* .

Kotiloglu, Serhan, Yan Chen, Thomas Lechler. 2021. Organizational responses to performance feedback: A meta-analytic review. *Strategic Organization* **19**(2) 285–311.

Kundu, Amrita, Kamalini Ramdas. 2022. Timely after-sales service and technology adoption: Evidence from the off-grid solar market in uganda. *Manufacturing & Service Operations Management* .

Levi, Retsef, Manoj Rajan, Somya Singhvi, Yanchong Zheng. 2020a. The impact of unifying agricultural wholesale markets on prices and farmers' profitability. *Proceedings of the National Academy of Sciences* **117**(5) 2366–2371.

Levi, Retsef, Manoj Rajan, Somya Singhvi, Yanchong Zheng. 2020b. Improving farmers' income on online agri-platforms: Theory and field implementation of a two-stage auction. *Available at SSRN 3486623* .

Lin, Wilson, Susan Feng Lu, Tianshu Sun. 2021. Worker experience and donor heterogeneity: The impact of charitable workers on donors' blood donation decisions .

Lurie, Nicholas H, Jayashankar M Swaminathan. 2009. Is timely information always better? the effect of feedback frequency on decision making. *Organizational Behavior and Human decisión processes* **108**(2) 315–329.

Mehrotra, Mili, Milind Dawande, Srinagesh Gavirneni, Mehmet Demirci, Sridhar Tayur. 2011. Or practice—production planning with patterns: A problem from processed food manufacturing. *Operations research* **59**(2) 267–282.

Mišić, Velibor V, Georgia Perakis. 2020. Data analytics in operations management: A review. *Manufacturing & Service Operations Management* **22**(1) 158–169.

NEST. 2018. The state of the handworker economy 2018. *NEST* .

Parker, Chris, Kamalini Ramdas, Nicos Savva. 2016. Is IT enough? Evidence from a natural experiment in India's agriculture markets. *Management Science* **62**(9) 2481–2503.

Paul, Alice, Daniel Freund, Aaron Ferber, David B Shmoys, David P Williamson. 2022. Erratum to "budgeted prize-collecting traveling salesman and minimum spanning tree problems". *Mathematics of Operations Research* .

Pedregosa, F., G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, E. Duchesnay. 2011. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research* **12** 2825–2830.

Peters, Koen, Sérgio Silva, Rui Gonçalves, Mirjana Kavelj, Hein Fleuren, Dick den Hertog, Ozlem Ergun, Mallory Freeman. 2021. The nutritious supply chain: optimizing humanitarian food assistance. *INFORMS Journal on Optimization* **3**(2) 200–226.

Pierce, Jason R, Herman Aguinis. 2013. The too-much-of-a-good-thing effect in management. *Journal of Management* **39**(2) 313–338.

Plambeck, Erica, Kamalini Ramdas. 2020. Alleviating poverty by empowering women through business model innovation: Manufacturing & service operations management insights and opportunities. *Manufacturing & Service Operations Management* **22**(1) 123–134.

Plambeck, Erica L, Terry A Taylor. 2016. Supplier evasion of a buyer's audit: Implications for motivating supplier social and environmental responsibility. *Manufacturing & Service Operations Management* **18**(2) 184–197.

Ramdas, Kamalini, Khaled Saleh, Steven Stern, Haiyan Liu. 2018. Variety and experience: Learning and forgetting in the use of surgical devices. *Management Science* **64**(6) 2590–2608.

Ramdas, Kamalini, Alp Sungu. 2022. Capping mobile data access creates value for bottom-of-the-pyramid consumers–experimental evidence from a mumbai settlement. *Available at SSRN 4012172* .

Ramin, Cathryn Jakobson. 2021. India's rug saint. *Cratsmanship Quarterly* .

Sodhi, ManMohan S, Christopher S Tang. 2014. Supply-chain research opportunities with the poor as suppliers or distributors in developing countries. *Production and operations management* **23**(9) 1483–1494.

Song, Hummy, Anita L Tucker, Karen L Murrell, David R Vinson. 2018. Closing the productivity gap: Improving worker productivity through public relative performance feedback and validation of best practices. *Management Science* **64**(6) 2628–2649.

Staats, Bradley R, Hengchen Dai, David Hofmann, Katherine L Milkman. 2017. Motivating process compliance through individual electronic monitoring: An empirical examination of hand hygiene in healthcare. *Management Science* **63**(5) 1563–1585.

Staats, Bradley R, Diwas S Kc, Francesca Gino. 2018. Maintaining beliefs in the face of negative news: The moderating role of experience. *Management Science* **64**(2) 804–824.

Sunar, Nur, Jayashankar M Swaminathan. 2022. Socially relevant and inclusive operations management. *Production and Operations Management* .

Syverson, Chad. 2011. What determines productivity? *Journal of Economic literature* **49**(2) 326–365.

Tan, Tom Fangyun, Serguei Netessine. 2019. When you work with a superman, will you also fly? an empirical study of the impact of coworkers on performance. *Management Science* **65**(8) 3495–3517.

Taylor, Lindsey Waterton. 2017. Woven fabrics. *Textile and Clothing Design Technology*. CRC Press, 143–194.

Ted, Barber, Krivoshlykova Marina. 2006. Global market assesment for handicrafts. *USAID* .

The Artisan Alliance. 2019. 2019-2020 impact report from the artisan alliance. *The Aspen Institute* .

Tsiligirides, Theodore. 1984. Heuristic methods applied to orienteering. *Journal of the Operational Research Society* **35**(9) 797–809.

Tuna, Ali Kaan, Robert Swinney. 2021. Are fast supply chains sustainable? *Available at SSRN 3995549* .

Uppari, Bhavani Shanker, Ioana Popescu, Serguei Netessine. 2019. Selling off-grid light to liquidity-constrained consumers. *Manufacturing & Service Operations Management* **21**(2) 308–326.

Van Loocke, Eddy, Vital Put. 2011. The impact of performance audits: A review of the existing evidence. *Performance Auditing* .

Vansteenwegen, Pieter, Wouter Souffriau, Dirk Van Oudheusden. 2011. The orienteering problem: A survey. *European Journal of Operational Research* **209**(1) 1–10.

Widjajanto, Sugiri, Humiras Hardi Purba, Sansuri Choesnul Jaqin. 2020. Novel poka-yoke approaching toward industry-4.0: A literature review. *Operational Research in Engineering Sciences: Theory and Applications* **3**(3) 65–83.

World Bank. 2021. A year in the lives of smallholder farmers. *World Bank* .

Xu, Wenzheng, Zichuan Xu, Jian Peng, Weifa Liang, Tang Liu, Xiaohua Jia, Sajal K Das. 2020. Approximation algorithms for the team orienteering problem. *IEEE INFOCOM 2020-IEEE Conference on Computer Communications*. IEEE, 1389–1398.

Zhang, Dennis, Yuval Salant, Jan A Van Mieghem. 2018. Where did the time go? on the increase in airline schedule padding over 21 years. *On the Increase in Airline Schedule Padding Over* **21**.

This page is intentionally blank. Proper e-companion title page, with INFORMS branding and exact metadata of the main paper, will be produced by the INFORMS office when the issue is being assembled.

Online Appendix

O.1. Field Visit Details

Different stakeholders in Jaipur Rugs' supply chain were interviewed during 3 separate trips in March, June, and December of 2022. The objective was to visit villages and better understand Jaipur Rugs' supply chain operations. The visits involved semi-structured interviews and focused on, (i) understanding the process of rug weaving; (ii) collecting perspectives on supervisor-weaver interactions in the supply chain. In total, the team visited 6 different locations and interacted with more than 20 weavers, 5 supervisors, and 2 branch managers during these visits. Key insights from the interactions are summarized below.

Weaver Interviews: Each interview with the weavers lasted for approximately 30 minutes. While some weavers were interviewed individually in their courtyards, others were interviewed in groups at the training centers of Jaipur Rugs. The findings from interviews conducted with weavers for the project revealed that the majority of weavers are women who work from 9 am to 5 pm, treating rug making as a regular job and relying on it as their primary source of income. Several key factors were identified that affect the speed and quality of rug making. The experience was noted as a crucial factor, with an estimated time of 3 years to achieve full speed. Additionally, unforeseen deaths or diseases in the family that require the weavers' attention were found to have an impact on their productivity. Most defects in rug making were attributed to issues such as wrong thread colors, loose strings, and slanted designs, particularly in designs with similar nearby colors. Furthermore, harvest season and rains were identified as potential factors that may affect the speed of rug making, as some artisans may also work on farms during these periods. These insights highlight the complexities and challenges faced by weavers and provide valuable information for understanding the dynamics of rug making as a livelihood for women artisans.

Supervisor Interviews: Supervisors were interviewed individually and asked about key aspects of supervision and scheduling their routes. Insights obtained from interviews with supervisors revealed their significant role in identifying defects in a timely manner by actively inspecting and communicating with the weavers. Additionally, supervisors also provide transparency on defects found in previous rugs at the finishing center, which serves as motivation for weavers to be more careful in their work. However, it was noted that more experienced weavers may become over-confident, resulting in more defects despite taking less time. Furthermore, supervisors visit each loom once every 4-5 days, covering 15-20 looms in a day. They decide their schedules by picking an area and covering all looms in that area on a particular day, or sequencing across different areas based on requirements. These insights highlight the crucial role of supervisors in quality control and their strategic approach to managing multiple looms and weavers for efficient rug production.

O.2. Additional Tables and Figures

We present results from the first stage for our two instrumental variables used in Table O.1. The results confirm that we find strong first-stage results for both instruments.

Table O.1 First-stage Results of Instrumental Variables

Variable	IV1 Avg. Days Between Visits	IV2 Avg. Days Between Visits
Driving Distance	0.011***	–
	(0.003)	
Avg. Days Between Visits (Previous Rug)	–	0.071***
		(0.010)
Product Cubage	-0.001	-0.001
	(0.001)	(0.001)
Knot Density	0.004**	0.003*
	(0.001)	(0.001)
Number of Colors	-0.014*	-0.013*
	(0.007)	(0.007)
Avg. temperature	0.061***	0.063***
	(0.007)	(0.003)
Size of Loom	-0.021	0.003
	(0.025)	(0.026)
City FE	Y	Y
Supervisor FE	Y	Y
Time FE	Y	Y
Experience FE	Y	Y
Cragg-Donald F statistic	16	66
Observations	8,006	7,150

Notes. "–" means the variable is not present in the model. Standard errors (in parentheses) are clustered at the loom level. ***: $p < 0.01$; **: $p < 0.05$; *: $p < 0.1$.

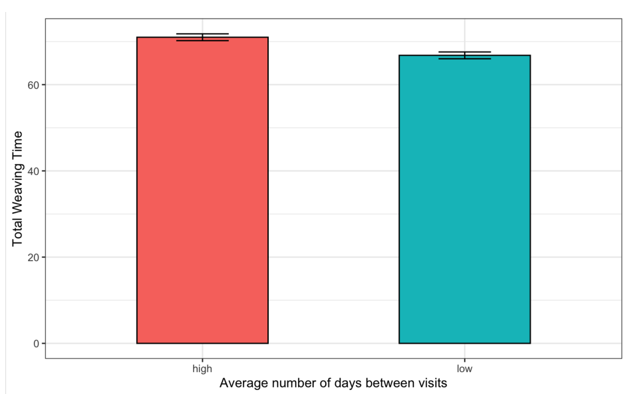
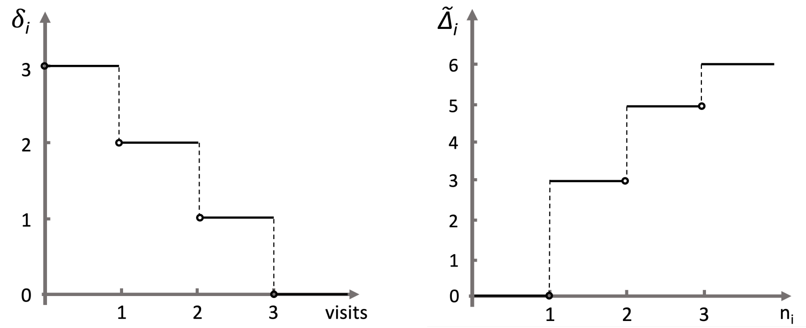


Figure O.1 Average weaving times for rugs with above-median (below-median) average number of days between supervisor visits are in the red (blue) bar.



(a) Profit of each visit to village v_i

(b) Profit of visiting village v_i, n_i times

Figure O.2 **Profit function of visiting village v_i when the target visit $\Omega_i^T = 3$. Figure O.2a demonstrates the profit gain from each additional visit to village v_i. For example, the first visit to village v_i generates profit $\delta_i = 3$ since v_i has never been visited before and $c_i = 0$. Similarly, $\delta_i = (3-1)^+ = 2$ for the second visit because the number of visits prior to the current visit is 1. Figure O.2b illustrates the definition of total profits collected at village v_i, $\tilde{\Delta}_i$ when village v_i is visited n_i times (Equation (7)).**

O.3. Results from Robustness Checks

We perform a number of robustness checks to further strengthen our results. Our results remain consistent under, (i) alternative dependent variable; (ii) alternative experience definitions; (iii) alternative data preparation; (iv) additional controls; (v) village-level aggregation to account for spillover concerns; (vi) alternative clustering; (vii) data filtering to tackle COVID-19 interference; (viii) using payments data instead as an alternate data source; (ix) redoing the analysis at an individual visit level.

Alternate Dependent Variable We consider the average number of knots weaved per day for each rug as an alternate measure of productivity and use it as our dependent variable. The average number of knots weaved per day for each rug is calculated by dividing the total number of knots in each rug by the total weaving days. We expect β_1 to be negative and statistically significant since increase in the average number of days between visits should be associated with a lower number of average knots weaved per day. The results from the specification are shown in Column (1) of Table O.2 and confirm that our results are consistent with the results from the main model.

Alternate Experience Definitions We consider an alternate experience measure based on the literature (Akşin et al. 2021). In particular, we find the log-transformed cumulative cubage weaved and use that as an alternate covariate to proxy experience in our specification. The results from the specification are shown in Column (2) of Table O.2 and confirm that our results are consistent with results from the main model. In particular, β_1 is again positive and statistically significant.

Alternate Data Preparation In order to check the robustness of our results against reverse causality concerns, we re-estimate Equation (1) by re-preparing the data in the following manner.

It is likely that (if) supervisors increase visits on a loom, they do so after the rug is already delayed. In order to minimize such concerns, for each rug we first calculate the expected date of completion without delay by using the expected daily weaving rate. Next, we recalculate the average number of days between visits by only using data from visits that happened *before* the expected date of completion. We also update the dependent variable to be the average number of knots weaved per day *before* the expected date of completion. The results from the analysis are in Column (3) of Table O.2 and are consistent with the results from the main model. In line with our hypothesis, we again find that a one-day increase in average days between visits leads to a 2.2% decrease in the average number of knots weaved per day.

Alternate Control Variables First, we additionally control for the design code fixed effects in this specification. Since the total number of colors and knot density will be identical for all rugs with the same design code, the effects of these variables are absorbed in the design code fixed effects. Column (4) of Table O.2 shows results that are consistent with the main model. Second, we control for month-fixed effects instead of season-fixed effects and confirm that the results remain consistent (Column (6), Table O.3).

Spillover Effects If weavers in the same village interact, supervisor's visit to one loom may also affect the productivity of other looms in the village. For example, if the supervisor shares knowledge about weaving errors with a weaver, weavers of neighboring looms may also be more cognizant of such errors after talking to the weaver of the focal loom. Therefore in order to control for these spillover effects from supervisor visits, we aggregate our data at the village level and re-estimate Equation (1) to check the robustness of our results. Column (5) of Table O.2 confirms that the results are consistent with the main model.

Alternate Clustering Recall that we estimate Equation (1) by clustering standard errors at the loom level. We consider a two-way clustering of standard errors at the month and loom level, and confirm that the results remain consistent under this alternative (Column (7), Table O.3).

Impact from COVID-19 To test the robustness of our results against concerns that the estimates may be driven by COVID-19, we re-estimate Equation (1) using data only from before 2019. Column (8) of Table O.3 confirms that the results are consistent with the main model.

Payments Data An alternative way to test the impact of supervisor visits on weavers' productivity is to use payment data. Results from Equation (1) suggest that if supervisors increase visit frequency in a month, weavers' productivity should increase. Increased productivity of a weaver should lead to an increase in payments made to her by Jaipur Rugs in that month since weavers are paid on a piece-rate basis. We obtain monthly payment data for weavers across branches in the state of Rajasthan between 2020-2021 and estimate the following specification to test this hypothesis.

$$P_{lmy} = \beta v_{lmy} + \phi_l + \delta_s + \psi_m + \gamma_y + \epsilon_{lmy}$$

P_{lmy} in this specification is payment made to loom l in month m and year y. v_{lmy} is the number of visits made by supervisor to loom l in month m of year y. We also control for month (ψ_m), year (γ_y), loom (ϕ_l) and supervisor (δ_s) fixed effects in this specification. We cluster standard errors at the loom level to account for correlation among observations from the same loom in this specification. Note that we cannot control for additional rug-level characteristics in this specification since weavers may work on multiple rugs in the same month. The estimates from the specification are included in Column (9) of Table O.3. We find that β is positive and statistically significant, suggesting that the increased visits are associated with an increase in payments for weavers.

Individual Visit Data Recall from §3.3 that we also observe supervisor visit dates and incremental work done since the last visit in our dataset. An alternative way to test our key hypothesis involves analyzing this dataset as follows. Suppose a supervisor s visits loom l at times t_{rli}, $i = 1, 2, 3..$ and records the incremental cubage weaved during visit i as c_{rli}. We calculate the weaving rate in period i as $s_{rli} \equiv c_{rli}/(t_{rli} - t_{rli-1})$. This rate is influenced by errors detected during the supervisor visit at period $i-1$. If the time gap between t_{rli-1} and t_{rli-2} is large, errors like slanted weaving or wrong thread colors would have persisted for a longer cubage, making them more time-consuming to rectify. As a result, a lower (higher) weaving rate, s_{rli}, is expected if the days between visits $i-1$ and $i-2$ are large (small). Let $d_{rli} \equiv t_{rli-1} - t_{rli-2}$, $i > 3$. We estimate the following model in order to test this hypothesis.

$$\log(s_{rli}) = \beta_0 d_{rli} + \beta \mathbf{X}_{rli} + \phi_l + \delta_s + \psi_m + \omega_y + \epsilon_{rli}. \tag{O.1}$$

$\mathbf{X}_{rli}$ includes time-varying rug and loom-level attributes discussed in §4. ψ_m, ω_y are month and year fixed effects during visit i. Standard errors are clustered at the loom level in this specification as before. A negative and statistically significant value of β_0 would suggest a negative correlation between days between visits, d_{rli} and weaving rates, s_{rli}. The estimates from the specification are included in Column (10) of Table O.3. We find that β_0 is negative and statistically significant, suggesting that the increased average days between individual visits are associated with a decrease in weaving rate during the subsequent visit.

O.4. Machine Learning Evaluation Metrics

We describe evaluation metrics for different ML methods for binary classification tasks:

Accuracy: The percentage of instances where predicted labels equal actual labels.

Precision: The fraction of true positives among all predicted positive label instances (Pedregosa et al. 2011).

Recall: The fraction of true positives among all positive label instances.

Table O.2 Robustness Checks I

Variable	(1) Alt. Dep. Variable	(2) Alt. Exp.	(3) Alt. Data Prep.	(4) Additional Controls	(5) Spillover
Avg. days between visits	-0.017***	0.028***	–	0.028***	0.032***
	(0.002)	(0.002)		(0.002)	(0.003)
Avg. days between visits prior to delay	–	–	-0.022***	–	–
			(0.002)		
Product Cubage	–	0.005***	–	0.005***	0.003***
		(0.000)		(0.000)	(0.000)
Knot Density	-0.009***	0.006***	-0.009***	–	0.005***
	(0.000)	(0.000)	(0.000)		(0.000)
Number of Colors	-0.005***	0.006***	-0.005***	–	0.004***
	(0.001)	(0.001)	(0.001)		(0.002)
Avg. temperature	0.006***	-0.006***	0.006***	-0.007***	-0.006***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
(log) Cum. Cubage Weaved	–	0.011***	–	–	–
		(0.004)			
Experience FE	Y	N	Y	Y	
Design FE	N	N	N	Y	
Other Controls	Y	Y	Y	Y	Y
Observations	8,006	8,006	7,981	8,006	6,967

Notes. "–" means the variable is not present in the model. Standard errors (in parentheses) are clustered at the loom level. ***: $p < 0.01$; **: $p < 0.05$; *: $p < 0.1$. Note that the dependent variable in Columns (1) and (3) is the average number of knots weaved per day and in all other columns is log-transformed total weaving time (in days). Other controls include loom, supervisor and time fixed effects in all specifications. For Column (5), we take the average of the independent variables wherever appropriate.

Table O.3 Robustness Checks II

Variable	(6) Month Controls	(7) Alt. Clustering	(8) COVID-19	(9) Payment	(10) Individual Visit
Avg. days between visits	0.026***	0.028***	0.023***	0.014***	-0.012***
	(0.002)	(0.007)	(0.003)	(0.006)	(0.001)
Product Cubage	0.005***	0.005***	0.006***	–	0.005***
	(0.000)	(0.000)	(0.000)		(0.000)
Knot Density	0.006***	0.006***	0.006***	–	-0.007***
	(0.000)	(0.000)	(0.000)		(0.000)
Number of Colors	0.006***	0.006***	0.005***	–	0.0001
	(0.001)	(0.001)	(0.001)		0.001
Avg. temperature	-0.009***	-0.007	-0.012***	–	-0.0002
	(0.002)	(0.007)	(0.001)		(0.0002)
Other Controls	Y	Y	Y	Y	Y
Observations	8,006	8,006	8,006	9,028	50,838

Notes. "–" means the variable is not present in the model. Standard errors (in parentheses) are clustered at the loom level. ***: $p < 0.01$; **: $p < 0.05$; *: $p < 0.1$.

Balanced Accuracy: The average recall obtained on each class.

Receiver Operating Characteristic (ROC) Curve: plots the true positive rate against the false positive rate (Pedregosa et al. 2011).

ROC AUC: ROC AUC is the Area Under the Receiver Operating Characteristic Curve (Pedregosa et al. 2011).

F1 Score: F1 score is the harmonic average of precision and recall (Pedregosa et al. 2011).

O.5. Details on the Machine Learning Framework

In this section, we briefly discuss the details of the ML framework used to predict lower-than-expected productivity rugs.

Feature engineering and selection of control features: Our objective is to predict whether a rug will have lower-than-expected productivity or not. Hence, some of the features used as a control in §4 might not be available at the time of making a delay prediction. For example, consider I_{rl} which is the average number of days between supervisor visits to rug r on loom l. Naturally, I_{rl} cannot be calculated until the rug is finished. Hence, we instead use the number of days on average between supervisor visits from the previous rug on loom l, I_{r-1l}, as a proxy for I_{rl}. Table O.4 describes different independent variables that we used as features, along with their descriptions, for the prediction task.

Benchmark algorithms: We compare state-of-art ML classification models such as Logistic Regression, XGBoost Classifier (Chen and Guestrin 2016), LGBM Classifier (Ke et al. 2017) and Random Forest Classifier (Ho 1995). While logistic regression is a parametric generalized linear classifier, the other methods use non-parametric tree-based classifiers with increased complexity. These models were mainly selected based on their superior performance and ease of usage.[12]

Test-train-validate framework and parameter tuning: To ensure there is no over-fitting, we consider a time-based split of the available data such that 70% of the data (from January 2017 to December 2020) is used for training, 10% of the data (from December 2020 to February 2021) is used to validate (tune different tuning parameters of the benchmark algorithms) and 20% of the most recent data for testing (from February 2021 to June 2021). We tune algorithm-specific tuning parameters (e.g., tree depth) independently using a uniform grid search.

O.6. Alternative Scheme of Hiring Additional Supervisors

While we discuss optimizing the routes of existing supervisors and targeting rugs predicted to have lower-than-expected productivity in §6.2, we may consider an alternative scheme and hire additional supervisors to improve weavers' productivity. More formally, recall that Ω_j^T is the pre-decided target number of supervisor visits for a village v_j over a time horizon T. We let $\Omega_j^{scheduled(T)}$ denote the scheduled number of supervisor visits for village v_j recommended by Policy MPTO without increasing the supervisor's workload. In other words, $\mathbf{\Omega}^{schedule(T)}$ is the solution of $\mathbf{OSVP(\Omega)}$ with an additional constraint $\sum_{(v_i,v_j)}(\frac{D_{ij}}{\nu}+\tau\cdot n_j)\cdot x_{ij}^t \le \bar{T}$, where $\bar{T}$ is the average daily working hours

[12] We use open-source Python implementation of these algorithms so that they can be easily used by our collaborator.

Table O.4 Feature Names and Definitions

Feature Name	Definition
Loom Size	Size (in ft) of loom
Prod. Length	Length (in ft) of rug
Prod. Cubage	Cubage (in sqft) of rug
QS	Quality supervisor name
Quality A	Quality code segment that represents color
Quality B	Quality code segment that represents style and knots density
Branch location	Branch location name
Loom Count	Number of looms at the same location
Order Priority	Internal company priority code for rugs
Month Issued	Issue month of rug
Year Issued	Issue year of rug
Order to Delivery	Days between issue date and estimated finish date
Previous Lead Time	Lead time (in days) of rug that is weaved on the same loom
Previous Avg. iat	Average supervisor inter-arrival time (in days) for previous rug

Table O.5 Prediction Results from Different Machine Learning Methods

Model	Accuracy	Balanced Accuracy	ROC AUC	F1 Score
XGBClassifier	0.72	0.71	0.71	0.72
LGBMClassifier	0.72	0.71	0.71	0.71
RandomForestClassifier	0.72	0.70	0.70	0.72
SVC	0.72	0.70	0.70	0.71
LogisticRegression	0.68	0.68	0.68	0.68

of the supervisor. We capture the marginal effect of hiring an additional supervisor by solving the optimization problem $\mathbf{OSVP}(\mathbf{\Omega} - \mathbf{\Omega}^{schedule})$, where $\mathbf{\Omega} - \mathbf{\Omega}^{schedule}$ represents the unfulfilled target visits given the visiting schedule of the existing supervisor.

Figure O.3 illustrates the cumulative average additional village visits per week facilitated by recruiting more quality supervisors. Note that while an additional supervisor increases the cumulative village visits, we observe diminishing marginal returns from supervisor recruitment and newly hired supervisors may be left idle, thus leading to another source of inefficiency (for example, the seventh supervisor only visits 2 villages throughout a weekly schedule). Recall from §4.3 that a one-day decrease in the average number of days between supervisor visits can decrease the weaving times by 2.8% -14.1%. Using these estimates, we can calculate the effect on weaving times from increased number of available supervisors and more frequent supervisor visits. In particular, Figure O.4 demonstrates that recruiting additional supervisors can decrease weaving times and subsequently reduce the number of rugs with lower-than-expected productivity by 1.56%-9.97% conservatively (7.91%-38.89% optimistically).

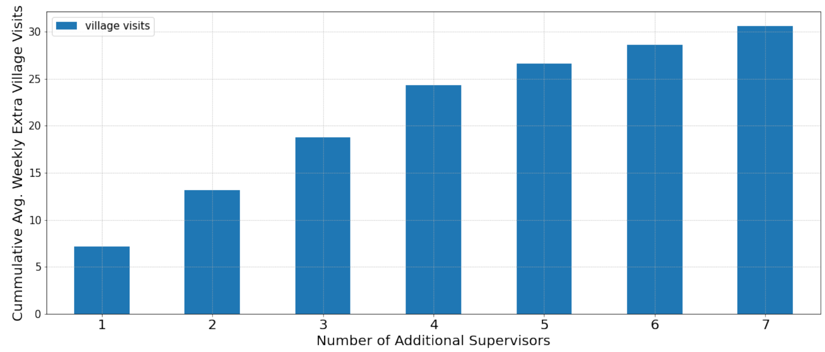


Figure O.3 Cumulative average weekly extra village visits from hiring additional supervisors

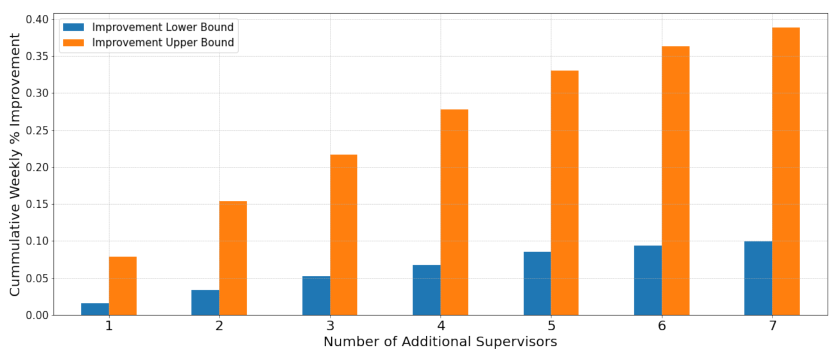


Figure O.4 Cumulative reduction in rugs with lower-than-expected productivity from hiring additional supervisors, the optimistic estimate is in orange and conservative estimate is in blue in the figure.

O.7. Continuous Model of Supervisor Visits (CMSV)

To increase supervisor visits on looms predicted to have lower-than-expected productivity, instead of having a binary model of pre-decided target number of supervisor visits $\Omega_j^{binary(T)}$[13] (as specified in Equation 4), we may alternatively consider a continuous function for village v_j, $\Omega_j^{cont(T)} = \gamma_j \sum_i (h_{ij} - f_{ij})^+$, where h_{ij} is the standard weaving rate for the rug on loom i in village v_j and f_{ij} is the corresponding machine-learning prediction of the weaving rate. Note that γ_j is a normalizing constant and it depends on the number of rugs being weaved in village v_j, n_j. We further define x_j as the integer decision variable indicating the number of supervisor visits to village v_j over the next T days, and we find $\boldsymbol{\Omega}^{cont}$ by solving an integer program given by,

$$\mathbf{CMSV} = \max_{\mathbf{x}} \quad \sum_{v_j} r_j x_j \tag{O.2a}$$

[13] We use $\Omega_j^{binary(T)}$ to denote the binary target number of supervisor visits Ω_j^T in Equation 4 for ease of reference in this section.

$$\text{s.t.} \quad x_j \geq \underline{c}, \ \forall v_j \tag{O.2b}$$

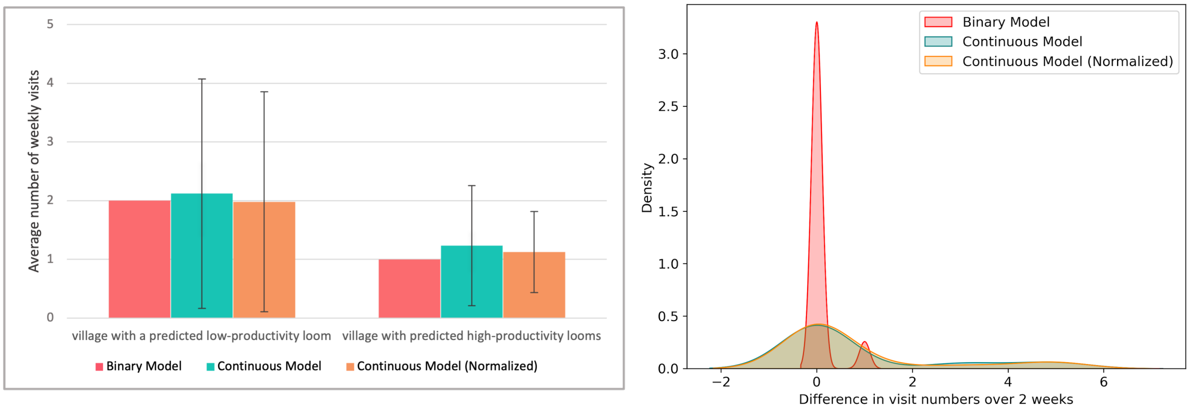
$$x_j \leq T, \ \forall v_j \tag{O.2c}$$

$$\sum_{v_j} x_j \leq \bar{c} \cdot |V|, \tag{O.2d}$$

$$x_j \in \mathbb{Z}, \ \forall v_j, \tag{O.2e}$$

where $r_j := \frac{\sum_i (h_{ij} - f_{ij})^+}{n_j}$ is the average delay in weaving rates across all rugs in village v_j. Therefore **CMSV** would prioritize visiting villages with lower than expected productivity, while ensuring that (i) each village is visited at least $\underline{c}$ times and at most T times (once per day)(O.2b and O.2c), and (ii) the total number of visits is upper bounded by the product of $\bar{c}$ and the number of villages (O.2d).

To better compare $\Omega_j^{binary(T)}$ and $\Omega_j^{cont(T)}$, we further consider a normalized continuous model, where we solve **CMSV** with an additional constraint $\sum_{v_j} x_j = \sum_{v_j} \Omega_j^{binary(T)}$, which enforces the same supervisor workload as in the binary model. Figure O.5 illustrates the results of solving **CMSV** for $\mathbf{\Omega}$ using real-world data from Jaipur Rugs when T is one week, $\underline{c} = 1$ and $\bar{c} = 2$. On the left (Panel a), we plot the average weekly target visits for villages with a loom that is predicted to have lower-than-expected productivity and villages with no such looms. For both the sets of villages, we compare the average number of weekly visits in the binary model (Equation (4)), the continuous model (**CMSV**) and the normalized continuous model, respectively. The error bars plot the standard deviation in visits scheduled from the corresponding model. By definition (Equation (4)), the binary model has constant weekly visit numbers across villages in each group. We find that, compared to the benchmark binary model, the continuous model slightly increases average weekly visits to villages with lower-than-expected productivity looms (villages with higher-than-expected productivity looms) by 5.5% (23.14%), but significantly increases the variability in the number of visits across villages. These results suggest that formulating a continuous model of supervisor visits leads to a highly variable number of visits across villages, making it difficult to implement in practice. The density plot on the right (Panel (b)) illustrates the distribution of differences in visits over 2 weeks within the same village. While the binary model has a maximum change of one visit, there is high variability in visits over time for the continuous model. In particular, for 21.74% of villages, the total number of visits more than doubled over the previous week. Similarly, for 10.14% of the villages, visits changed from one extreme to the other over two weeks (a difference of 5 visits week-over-week). Our calculations thus indicate that the continuous formulation results in supervisor visits that are more sensitive to the machine-learning predictions, and in turn highly variable across villages and time. Therefore we consider the binary objective visit formulation specified in Equation (4).



(a) Supervisor objective visits

(b) Visits variability across weeks

Figure O.5 **Average weekly objective visits for villages with a predicted low-productivity loom and villages with predicted high-productivity looms when $\underline{c}=1$ and $\bar{c}=2$ (Panel (a)), where the error bars plot the standard deviation in visits from the corresponding policy. Density plot of difference in visits to villages over 2 weeks (Panel (b)). Binary Model is in red, Continuous Model is in green and Normalized Continuous Model is in orange in the figure.**

O.8. Alternative Objective Formulation

While the formulation of **OSVP(Ω)** in (5a)-(5k) aims at minimizing the difference between the target and scheduled supervisor visits, we may also want to ensure that the newly proposed routes remain as close as possible to current supervisor routes for easier acceptance and real-world implementation. We account for this objective and present an alternative formulation.

As before, we let $y_j^t = 1$ if village v_j is visited on day t, and 0 otherwise and $x_{ij}^t = 1$ if arc (v_i, v_j) is traversed on day t, and 0 otherwise. We further define $x_{ij}^{existing(t)} = 1$ if the current supervisor routes traverse across village v_i and v_j on day t, and 0 otherwise. Note that the value of $\mathbf{x}^{existing}$ is fixed and given and thus we have,

$$\min_{\mathbf{x},\mathbf{y},\mathbf{w},\mathbf{s}} \quad \lambda \sum_{v_j} w_j + (1-\lambda)\sum_t s_t \tag{O.3a}$$

$$\text{s.t.} \quad \sum_{(v_i,v_j)} x_{ij}^{existing(t)} - x_{ij}^t \le s_t, \ \forall\, t \tag{O.3b}$$

$$\sum_{(v_i,v_j)} -x_{ij}^{existing(t)} + x_{ij}^t \le s_t, \ \forall\, t \tag{O.3c}$$

$$s_t \ge 0, \ \forall\, t \tag{O.3d}$$

$$(5b) - (5k), \tag{O.3e}$$

where $0 < \lambda < 1$ is a tuning parameter, and the additional constraints (O.3b)- (O.3d), together with the auxiliary variable s_t, set limits on the proposed changes to existing routes.

O.9. Details on Estimated Impact of Targeted Supervisor Visits

In this section, we give detailed steps of how we derive the estimated reduction in rugs with lower-than-expected productivity.

Recall that our datasets include rug-level production data with features of rug weaving time (in days), baseline expected completion time (in days) used by Jaipur Rugs, and the supervisor visit data with the exact dates on which the supervisor visited a rug. A rug is considered to have lower-than-expected productivity if its weaving time exceeds the baseline expected completion time. With the datasets and features above, we are able to calculate (i) the actual number of rugs with lower-than-expected productivity, (ii) the actual average number of days between supervisor visits for any rug, and (iii) the scheduled average number of days between supervisor visits under Policy MPTO for any rug. Since a one-day decrease in the average number of days between supervisor visits can decrease the weaving times by 2.8% -14.1% (see §4.3), we obtain the new rug weaving time under Policy MPTO by taking the difference between (ii) and (iii) and then multiplying by the value of improvement rate (2.8% -14.1%). Following the same definition of rug productivity, we derive the updated number of rugs with lower-than-expected productivity by comparing the new weaving time with the baseline expected completion time.

O.10. Proofs

Proof of Proposition 1 To show the existence of a computationally tractable algorithm for solving the supervisor visit optimization problem $\mathbf{OSVP}(\mathbf{\Omega})$ when $\underline{c}=0$, we prove the equivalence of $\mathbf{OSVP}(\mathbf{\Omega})$ and $\mathbf{TOP}(\tilde{\mathbf{\Delta}}, G(V))$ so that the proposition follows immediately by applying the result of Xu et al. (2020).

Assume $\underline{c}=0$. We claim that the $\mathbf{OSVP}(\mathbf{\Omega})$ problem with time horizon T is equivalent to $\mathbf{TOP}(\tilde{\mathbf{\Delta}}, G(V))$ with K vehicles if $T=K$. To see this, we establish a one-on-one correspondence between the conditions of the $\mathbf{OSVP}(\mathbf{\Omega})$ problem and the $\mathbf{TOP}(\tilde{\mathbf{\Delta}}, G(V))$ problem. We first show that the village-level profit function $\tilde{\Delta}_i(\cdot)$ has two properties. (i) the non-decreasing property: suppose $1 \le n_i \le n_i'$ are two integers, then we have $\tilde{\Delta}_i(n_i) = \sum_{c_i=0}^{n_i-1}(\Omega_i^T - c_i)^+ \le \sum_{c_i=0}^{n_i'-1}(\Omega_i^T - c_i)^+ = \tilde{\Delta}_i(n_i')$, since each term $(\Omega_i^T - c_i)^+$ is non-negative; (ii) the submodularity property: suppose $0 \le n_i \le n_i'$, then for any non-negative integer δ, we have $\tilde{\Delta}_i(n_i+\delta) - \tilde{\Delta}_i(n_i) = \sum_{c_i=n_i}^{n_i+\delta-1}(\Omega_i^T - c_i)^+ \ge \sum_{c_i=n_i'}^{n_i'+\delta-1}(\Omega_i^T - c_i)^+ = \tilde{\Delta}_i(n_i'+\delta) - \tilde{\Delta}_i(n_i')$ because each term $(\Omega_i^T - c_i)^+$ is non-increasing with respect to the visit number c_i.

Next, recall that in the $\mathbf{OSVP}(\mathbf{\Omega})$ problem, we have binary decision variables x_{ij}^t indicating whether arc (v_i, v_j) is traversed on day t or not. We note that the set of active arcs $(\{x_{ij}^t = 1\})$ satisfying the valid arc flow constraints ((5e)-(5f)) and the no subtour constraint (5g) translates directly to the decision variable P_k of valid path starting and ending at v_0 in the $\mathbf{TOP}(\tilde{\mathbf{\Delta}}, G(V))$

problem, and vice versa. With the assumption that $\underline{c}=0$, constraint (5b) becomes trivial and constraint (5c) is automatically satisfied in the $\mathbf{TOP}(\tilde{\mathbf{\Delta}}, G(V))$ problem since the marginal profit beyond $\mathbf{\Omega}$ is set to 0. The cost budget B_k in Xu et al. (2020) corresponds to $T_{\max}$ in constraint (5d). The cost function $\xi(P_k)$ defined as $\sum_{v_i \in P_k} h_k(v_i) + \sum_{(v_i,v_j)\in P_k} c_k(v_i,v_j)$ in Xu et al. (2020), where h_k is the node service cost and c_k is the edge travel cost, maps exactly to constraint (5d), where we have the loom inspection time cost $\tau \cdot n_j$ at village v_j and travel time cost $\frac{D_{ij}}{\nu}$ per arc (v_i, v_j), respectively. We therefore conclude that the feasible regions of the $\mathbf{OSVP}(\mathbf{\Omega})$ problem and $\mathbf{TOP}(\tilde{\mathbf{\Delta}}, G(V))$ problem are equivalent. For the objective functions, note that closing the gap between scheduled visits and target visits (as in $\mathbf{OSVP}(\mathbf{\Omega})$) is the same as allocating reward for each lower-than-target visit and maximizing total collected rewards (as in $\mathbf{TOP}(\tilde{\mathbf{\Delta}}, G(V))$). Hence, $\mathbf{OSVP}(\mathbf{\Omega})$ is equivalent to $\mathbf{TOP}(\tilde{\mathbf{\Delta}}, G(V))$ when $\underline{c}=0$.

Having established the equivalence of $\mathbf{TOP}(\tilde{\mathbf{\Delta}}, G(V))$ and our supervisor scheduling and routing optimization problem, we may leverage the polynomial run-time algorithm developed by Xu et al. (2020) to solve $\mathbf{OSVP}(\mathbf{\Omega})$ efficiently. In particular, this implies that there is a $(1-(\frac{1}{e})^{\alpha})$-approximation algorithm for $\mathbf{OSVP}(\mathbf{\Omega})$ following Theorem 1 of Xu et al. (2020), assuming that there is a α-approximation solution for the classic Orienteering Problem (Golden et al. 1987), $0<\alpha<1$. Applying the state-of-art $\frac{1}{2+\epsilon}$-approximation algorithm for the Orienteering Problem due to Chekuri et al. (2012), where ϵ is any fixed constant with $0<\epsilon\leq 1$, we state that our supervisor visit optimization problem $\mathbf{OSVP}(\mathbf{\Omega})$ can be solved efficiently with competitive ratio of $1-(\frac{1}{e})^{\frac{1}{2+\epsilon}}$.