PEDL Research Papers

This research was partly or entirely supported by funding from the research initiative Private Enterprise Development in Low-Income Countries (PEDL), a Department for International Development funded programme run by the Centre for Economic Policy Research (CEPR).

PEDL Twitter

Knowledge Exchange and Productivity Spill-overs in Bangladeshi Garment Factories

Andreas Menzel*

CERGE-EI, Prague†

May 2020

Abstract

Knowledge sharing between employees has long been viewed as a major driver of firm productivity growth, and has commonly been measured by productivity spill-overs within firms. Using data from three Bangladeshi garment factories, I first find that spill-overs occur within organizational sub-divisions of the factories, but not across. I then show that a management intervention that routinely brought together team leaders producing the same garments to exchange production knowledge further strengthened spill-overs within sub-divisions, but not across, when it was implemented in randomly selected sub-divisions. These findings suggest that boundaries between sub-divisions pose strong frictions to knowledge sharing within firms.

Keywords: Organizational Learning, Productivity, Firms, Randomized Trials

JEL Code: D2, L2, M5, O3

Declarations of interest: none

*I am grateful to Chris Woodruff and Rocco Macchiavello for their support and advice throughout this project. I furthermore would like to thank Robert Akerlof, Oriana Bandiera, Michal Bauer, Sascha O. Becker, Emily Breza, Clement de Chaisemartin, Florian Englmaier, Jonas Hjort, Stepan Jurajda, Steven Levitt, Nikolas Mittag, Roland Rathelot, Eric Verhoogen, and seminar participants at Columbia University, Urbana Champaign LER, CUNEF, as well as participants at the Stanford Conference on Firms, Trade and Development 2018, the Market and Organisations in Emerging Economies Conference 2018 at Kellog School of Business, Oxford Young Economist Workshop 2014, PEDL Workshop 2016, Italian Development Association Summer School 2016, and the LMU Workshop Ohlstadt 2014 and 2016 for many useful comments. Financial help by GACR Standard Grant 17-26395S and PEDL is acknowledged.

†Address: R.330 - CERGE-EI, a joint workplace of Charles University and the Economics Institute of the Czech Academy of Sciences, Politickych veznu 7, P.O. Box 882, 111 21 Prague 1, Czech Republic. Email: Andreas.Menzel@cerge-ei.cz

1 Introduction

Organizational learning, "the creation, retention and exchange of knowledge within organizations" (Argote (2013)), has long been viewed as a major driver of firm productivity growth, and thereby overall economic growth (Arrow (1962); Lucas (1993)). Extensive literature shows that the productivity of establishments increases over time due to the accumulation and sharing of production experience among employees (Wright (1936); Benkard (2000); Hendel and Spiegel (2014)). Recently, Sandvik et al. (2020) and Papay et al. (2020) showed that experimentally pairing employees and encouraging them to meet and exchange job skills increases their productivity, especially of initially less productive employees. Other papers have shown that proximity of employees within buildings supports knowledge exchange (Catalini (2018); Battiston et al. (2017)). This paper combines these literatures by showing first that also in garment factories in Bangladesh, experimentally pairing workers and encouraging them to share production knowledge leads to productivity increases. Second, these increases are much stronger when employees from the same sub-divisions inside the factories ("production floors") are paired with each other. Meetings of employees from different floors were less likely to be implemented even when called for by the experimental design. Qualitative evidence from the three participating factories indicates that information and coordination frictions between the different floor-managements prevented the implementation of more of the planned cross-floor pairings. Finally, the effects of the knowledge sharing intervention are mostly visible when the workers start producing newly allocated garment orders, by shortening the learning curve process that their productivity goes through.

Garment factories in Bangladesh are organised into dozens, or sometimes hundreds, of separate production lines. Almost daily, these factories roll out the production of new garment styles, ordered by a diverse set of international brands, across their lines.[1] On average, production lines in the three factories studied in this paper switch every 12 days to new styles allocated to them by the factory's planning departments. This requires constant dissemination of production knowledge on the new styles among the workers on the different lines. Using two year's worth of daily productivity information, I first show that production lines have higher

[1] Style is the industry term for the technically-differentiated garment designs that the buyers order. For example, two dress-shirts from the same brand but with different designs would be two different styles.

productivity during the first days they produce a newly allocated style if other lines on the same production floor have already gained production experience with the same style, as shown in Figure 1. These productivity spill-overs are similar to those that have been used in previous literature as evidence for knowledge exchange between workers (Benkard (2000); Thompson and Thornton (2001); Levitt et al. (2013); Egelman et al. (2016)). However, these effects are not visible when only lines located on other floors have already produced the style. On the other hand, I do not find that the effect is stronger between lines located closer to each other on the same floor. Given that floors are the primary organizational sub-divisions of the factories, this suggest organizational barriers as a main friction to knowledge exchange in these firms.

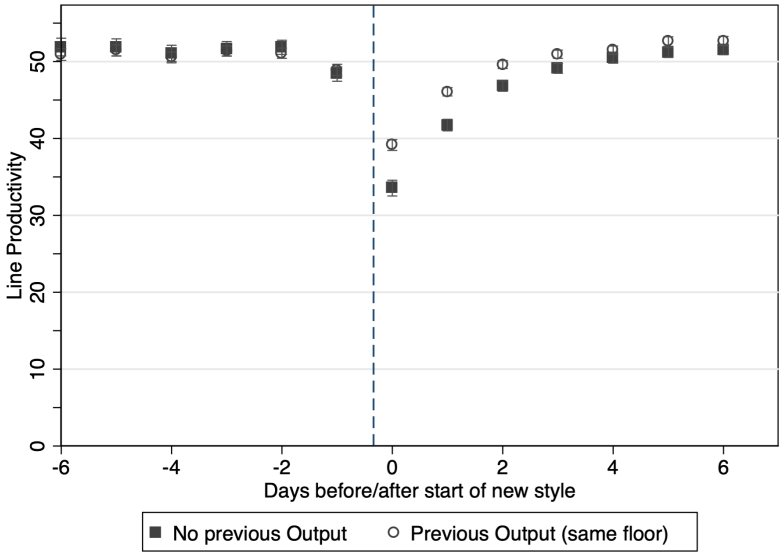

Figure 1: Sewing Line Productivity in days before and after Start of New Style. Dashed line denotes switch to new style, and "Day 0" the first day of production of new style. Capped bars represent 95% confidence intervals. Productivity is daily piecewise output multiplied by style specific SMV (Standard Minute Value), divided by daily line-wise labor input in worker-minutes (see Section 2 for more detail). Productivity controlled for Style fixed effects

To understand whether knowledge sharing levels between lines are efficient in the factories, or whether these levels can be increased by basic and inexpensive management routines, I implemented with the managements of the three factories a randomized worker pairing and communication intervention, similar to those studied in Sandvik et al. (2020) or Papay et al. (2020). The intervention mandated that whenever production lines were allocated a new style, their line supervisor was brought together with another line supervisor who had already pro-

duced the same style, in the event that such a supervisor was available in their factory. The experienced supervisor was given up to 30 minutes to share the production knowledge he or she had gained with the particular style with the supervisor starting to produce it. For the purpose of the evaluation shown in this paper, the intervention was first implemented only on randomly selected floors in the factories for a period of four months.

I find that the intervention amplified existing patterns; spill-over strength within floors increased relative to control floors during the four months in which the intervention was run on these floors. Basic calculations show that the productivity gains from the intervention are significantly larger than their estimated costs. However, despite the explicit incorporation of knowledge exchange between supervisors from different floors in the design of the intervention, I detect very little increases in spill-overs across floors. Using log-books documenting the implementation of the intervention in the factories, I find that compliance levels with implementing cross-floor briefings was lower than those for implementing within-floor briefings. Post-intervention surveys with the factory managements suggest that floor level managements were often unaware that experienced supervisors with the style were available on other floors, while the central management lacked man-power to organise the frequent cross-floor briefings. Such information frictions between organizational sub-divisions of factories could explain why young small firms often exhibit higher productivity growth, as opposed to larger and older firms (Haltiwanger et al. (2016)).

The intervention differs in two important ways from those studied Sandvik et al. (2020) or Papay et al. (2020). First, while in these interventions pairings were stable over time and workers met repeatedly with the same peer, in my intervention, supervisors were potentially matched for each briefing with a different peer, depending on which other supervisors had already produced the style they were starting. Second, while these interventions focused on transmission of general skills to become a more productive employee overall, my intervention focused on transmission of production knowledge on specific products, that may be obsolete once workers switch to producing a different style again. Thus, this paper studies whether knowledge exchange interventions are effective in transmitting detailed, product-specific knowl-

edge, without the repeated, relationship building component that the other two studies entailed.

Bangladeshi garment factories are a uniquely suitable setting to study organizational learning for a number of further reasons. First, the factories are organized into large numbers of autonomous production lines. Given large order sizes, however, the same style is often produced by more than one line, though these lines typically do not start producing the style on the same day since they first have to finish different previously-allocated styles. This generates significant variation in the amount of a style already produced on other lines when a new line starts producing the style. The three factories in my data have a total of 260 lines, located on 23 floors, and produce around 2,400 different styles in the time covered by my data, around 1,300 of which are produced on more than one line. Therefore, this paper first presents observational evidence on spill-overs from the complete dataset I collected, before studying the effects of the randomized management intervention, which was implemented at the end of the time covered by the data. This allows me to benchmark the effects of the intervention against the spill-over patterns observed at baseline, and estimate the effects of the intervention in a difference-in-difference framework. This approach, and the similarity of observational and experimental results, is also reassuring in light of some imbalances in observable characteristics of production floors randomly allocated to treatment and control for the intervention, which I discuss further below.

A second advantage of my setting is that the factories work with a basic, labour-intensive technology. Therefore, the results are less dependent on idiosyncratic features of the production process, supporting their external validity. This study is thus also complementary to many previous studies on productivity spill-overs between production units that focus on high technology or capital-intensive industries, such as electronics, automobile, and airplane plants, or ship-yards (Benkard (2000); Thompson and Thornton (2001); Levitt et al. (2013); Egelman et al. (2016)). Furthermore, McKinsey (2011) for the Bangladeshi garment sector, and Bloom and Reenen (2012) and Bloom et al. (2013) for manufacturing in South Asia more broadly, suggest that management quality is sub-optimal. This leaves scope to test whether targeted management interventions increase productivity in the sector, and what the limitations to this

approach may be. Finally, this paper connects to Atkin et al. (2017), who study communication incentives between employees on different hierarchy levels in similar export oriented factories in Pakistan, while I study the effectiveness of horizontal communication between lower-level managers in firms.[2]

This paper proceeds as follows. The next section provides more information about the setting and the collected data. Section three presents the non-experimental results on productivity spill-overs. Section four provides more details on the randomized management intervention and shows and discusses its effects, while section five concludes.

2 Background and Data

This study was conducted at three large garment factories in Bangladesh. Both ownership and management of the factories are domestic, and all output is produced for export. The factories produce mainly t-shirts, polo shirts, dress shirts and trousers, employing ca. 1,200 to 5,000 workers in their sewing sections, on which this paper focuses (other sections being, e.g., knitting, dyeing, cutting, and finishing sections). The smallest is Factory 2, with 15 sewing lines on three sewing floors, while Factory 3 is the largest, with 183 lines located on 14 floors. Factory 3 also uses fewer line supervisors (or "Line Chiefs"); in most cases, two lines share one line chief, and in some cases even four lines. Factory 1 has 59 lines on six floors; each line has its own line chief. Factory 2 has more workers per line, stemming from its specialization into woven garments, while Factory 1 and 3 produce knit garments.

Sewing lines are organized as assembly lines in which each worker only does one sewing operation. Additionally, each line has one to three quality inspectors, and garments found with quality defects are not counted in the line's output. The factories' central planning departments

[2]By using detailed data on worker characteristics and productivity, this paper also connects to a broader and rapidly growing literature on the interplay between management, worker characteristics and productivity. Adhvaryu et al. (2018) and Adhvaryu et al. (2019) show that line manager characteristics have significant effects on line productivity in Indian garment factories, which are very similar to the ones studied in this paper from Bangladesh. Hjort (2014) provides a case study of a Kenyan flower packaging factory showing that ethnically more diverse work teams have lower productivity. Similarly, Kato and Shu (2011) use data from a Chinese garment factory to show that the effect of team incentives to increase productivity depends on the composition of work teams out of urban and rural migrant workers. Similarly, Hamilton et al. (2003) study the introduction of team work in a U.S. garment factory with a similar set-up as the factories I study in Bangladesh, and find a significant increase in productivity.

Table 1: Factory Characteristics

	Factory 1	Factory 2	Factory 3
Nbr. Sewing Floors	6	3	14
Nbr. Sewing Lines	59	15	183
Nbr. Workers in Sewing Section	ca. 2000	ca. 1200	ca. 5000
Nbr. Workers in whole Factory	ca. 5000	ca. 2000	ca. 9000
Nbr. Buyers	25	64	9
Nbr. Styles in Data	757	791	824
Avg. Nbr. Lines /Style	3.13	1.47	3.65
Avg. Nbr. Days /Style & Line	16.1	9.4	8.1
Avg. Nbr. Workers /Line	31.0	72.4	23.2
S.Dev. Nbr. Workers /Line	8.0	11.0	7.1

Notes: All information from production data collected from factories, except for 'Nbr. Workers in ...' which is from surveys of factory management.

allocate workers, supervisors, and styles to sewing lines. Sewing lines are kept homogeneous in terms of size and productivity within the factories, and workers are not sorted to lines according to experience or productivity.[3] The reason for this lies in the high flexibility required in operations. Buyers place orders with low predictability, particularly regarding the order's specifications, and with close delivery deadlines. Furthermore, frequent disruptions to the production process (power failures, unrest outside factories, problems in supply and delivery chains, missing inputs) often require reallocations and re-prioritization of orders to lines. Therefore, it is not optimal to have differentiated lines specializing in certain types of garments. According to the production engineers in the planning departments, incoming orders are prioritized based on the importance of the buyer to the firm, and how close the delivery date is, and are then essentially allocated to the next free line.

For these reasons, also workers have fixed lines to which they are allocated, and usually change lines only as a result of promotions to new positions. Workers with production experience on some styles are generally not reallocated to other lines if these lines also start producing the same style. Given that on any typical day, several lines switch to newly allocated styles, with the identity of these lines often being little predictable due to the reasons mentioned above,

[3] An exception are the "sample lines", on which experienced workers produce samples of new orders from buyers during the negotiation processes. Sample lines are not included in my dataset.

constantly re-optimizing the allocation of workers across lines (and positions within lines) would be logistically extremely challenging. Thus, it is unlikely that such reallocations of workers drive the observed productivity spill-overs across lines producing the same style. Furthermore, Appendix A.1 presents the results of a placebo test suggesting the same. In all three factories in my sample, ordinary workers receive hourly pay, while line supervisors receive fixed monthly wages. No workers in the factories receive piece rate payments or performance bonuses, and promotions occur rather infrequently on the supervisor level, as further up positions, such as floor managers, are scarce. Thus, it seems unlikely that explicit career or money concerns provide strong incentives for or against knowledge sharing in this setting.

The lack of specialization of lines into types of garments can be seen in the stacked bar charts for each of the three factories in Figure 2, in which each bar represents a sewing line, and the wider spaces between the bars separate lines from different floors. The differently shaded parts of the bars represent the shares of different garment types (e.g. t-shirts, polo shirts, pants) among all styles the lines produce. The graphs show few patterns of lines specializing in certain types of garments. Lines also do not specialize on being brought in earlier or later in the roll-out of styles across lines. Figure 3 shows a similar stacked bar chart as Figure 2, but this time the different parts of the bars indicate the share of styles the line produced with different "start-ranks". This is, whether the line produced it as first (orange), second (light blue), or third or later line (dark blue) in the factory. Again, few obvious patterns of lines being more often allocated styles early or later on can be seen.[4]

The main dataset used for the analysis contains daily, line-level production data for all lines in the factories for 30 consecutive months from Factories 1 and 2, and for eight consecutive months from Factory 3 (this factory was recruited for this project only at a later point in time). This dataset includes: daily sewing line productivity; daily production hours and number of workers for each line, the identifier of the style being produced by the line on that day; its buyer; and the Standard Minute Value (SMV) of the style. The SMV is the sum of the time,

[4] At Factory 3, the six floors to the left of the graph produce for one large buyer, while the other floors produce for other buyers. Because the orders from this buyer are larger, they are produced on average on more lines. As a result, lines on these floors are on average less often the first line to produce a given order, and instead are more often second or later.

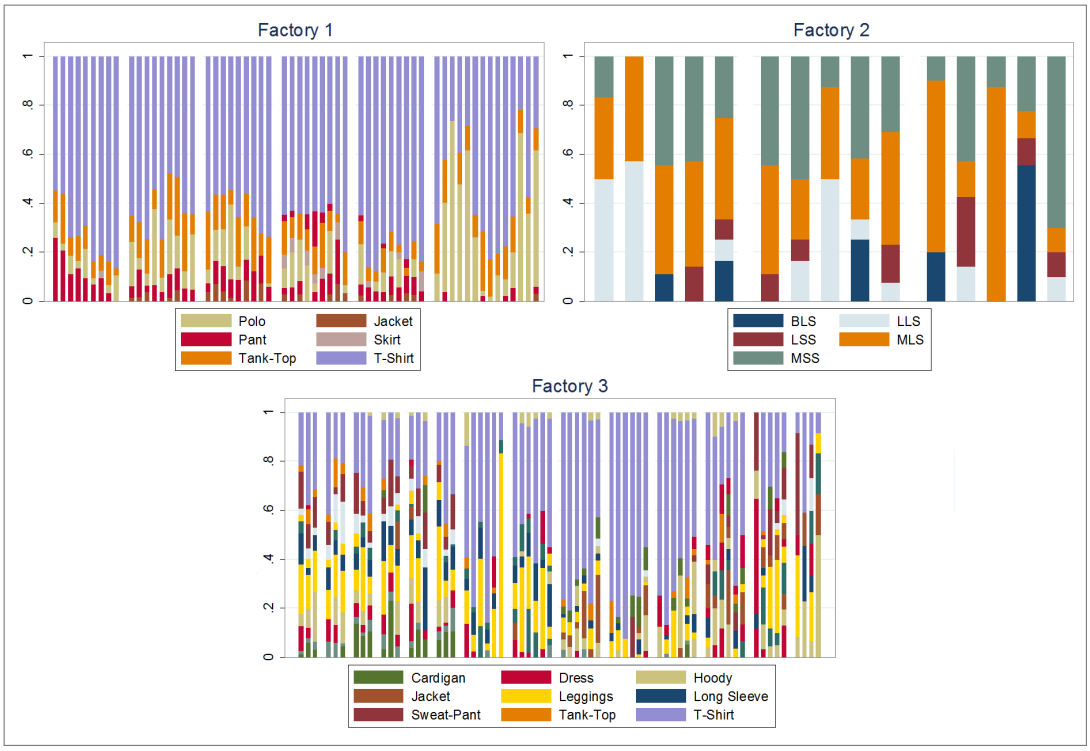

Figure 2: Garment Types Produced on Different Sewing Lines. The graphs represent the types of garments produced by different sewing lines at the three factories. Each bar in the graphs represents a sewing line, and the wider spaces between bars separate sewing lines from different floors. Legends show most common garment types only, for illustration. In sub-graph of Factory 3, each bar represents a line chief instead of a line (line chiefs at this factory look after 2 or 4 lines), to keep the number of bars in the graph parsimonious.

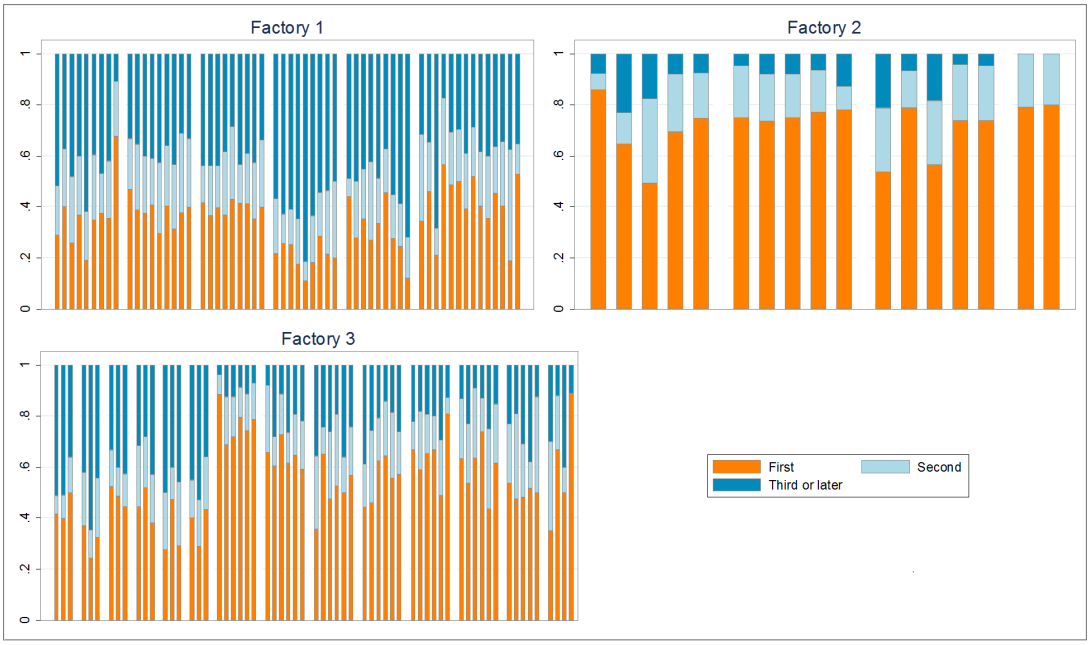

Figure 3: Start Ranks of Styles produced on different Sewing Lines. Graphs show for each sewing line in the three factories for which share of the styles they produce the lines are the first (orange), the second (light blue), or the third or later line (dark blue) to produce that style in the factory. Each bar represents one line, and the wider spaces between bars separate lines located on different floors. In sub-graph of Factory 3, each bar represents a line chief instead of a line (line chiefs at this factory look after 2 or 4 lines).

in seconds, it takes to perform each sewing operation to assemble one piece of the style under ideal production conditions. To calculate the daily line-productivity measure, daily piecewise output is multiplied by the style-specific SMV, and then divided by total labor input on that line and day measured in worker-minutes:

$$\text{Productivity} = \frac{\text{Daily Output} * \text{SMV}}{\#\text{Worker} * \text{Daily Hours} * 60^{mins}}$$

I also conducted a survey of all line chiefs at the three factories, who are on average around the age of 30, and only two of 128 line chiefs interviewed were female. They have worked as line chiefs on average for 1.8-3.5 years at the factories, and on average for more than one year on the line they were supervising at the time of the survey. This fits with the accounts from the factory management and the production data; line chiefs generally have a fixed line and are rarely reallocated.[5]

[5] Among ordinary workers, the average share of female workers on the sewing lines is around 80 percent. However, very few women are promoted to these positions, with reasons for this gender inequality discussed in Macchiavello et al. (2015)

3 Observational Evidence on Productivity Spill-over

Before providing the results from the experimental management intervention, this section explores the strength of productivity spill-overs at the three factories using observational variation in the amount of a style already produced by other lines in the factories when additional lines start its production as well. This allows the use of the large amount of data in the dataset which stems from the time before the start of the intervention, and thus allows to understand how the novel intervention altered the baseline situation. I first show the observational results and then briefly discuss at the end of this section the necessary assumptions to identify spill-over strength from the observational variation in previous output. In the overall dataset, I observe 2,372 different styles being produced, 1,306 of which have been produced on more than one production line in the factories.[6] Given that styles produced on only one line do not give me variation to study spill-overs across lines, I exclude them from the analysis (17.3% of observations).

Denote the nth day a line produces a style as the nth "style day."[7] I first estimate the model:

$$y_{fisnt} = \sum_n \beta_n^A ln(A_{is}) + \sum_n \beta_n^F ln(F_{is}) + \alpha_{fin} + \gamma_{ftn} + \epsilon_{fisnt} \tag{1}$$

Productivity y_{fisnt} of production line i in factory f producing style s in month t on its nth 'style day' is regressed on output A_{is} of the same style that has already been produced on all other sewing lines in the factory up to, but excluding, the first day on which line i started producing style s. Excluding output from the first day line i produces the garment prevents bias in the estimation of the spill-over parameters β_n^k from factory-wide productivity shocks that would affect both productivity of line i and the cumulative output on other lines, under the assumption that these productivity shocks are independent across days.

I interact this previous output from other lines with fixed effects for style day n. Thus, the

[6] Note that even if two factories in my dataset would produce identical garments, my data would not allow me to identify them, as factories use their own style identifiers in the data they shared with me. Note that the three factories are at least 10km apart from each other, so spill-overs between them are unlikely.

[7] When lines return to the production of a style they already had produced earlier, I do not consider it a new style-start, and do not start counting style days from 1 again, but from the level at which it stopped when the line switched away to another style the last time. Furthermore, at Factory 3, most line chiefs supervise two or four lines. As I am mainly interested in cross-worker spill-overs, and given that the management intervention studied in the next section is an intervention between line chiefs, I include from this factory only the observations of a line chief producing a certain style from the first line he/she produces it on.

effect of previous output of the same style is estimated separately for each 'style day' n. This allows to explore non-parametrically for how long previous output on other lines affects the productivity of a line producing a new style. I use the log of previous output on other lines as I expect each additional piece of previous output to have a diminishing marginal effect on the stock of knowledge with the style held by workers on other lines. Since Wright (1936), the literature on learning within firms has frequently used the log function to approximate knowledge with cumulative output.[8]

I also include output F_{is} of the same style produced on all other lines *on the same floor*, with its additional effect estimated again separately for each style-day n. Sewing lines located on the same floor are located in parallel and only two to three meters apart from each other. Furthermore, each sewing floor has a floor manager, who is the direct superior to the line chiefs, and who could transfer knowledge from one line to another on his floor. For these reasons we could expect the effect of previous output by lines on the same floor to differ from that of output on other floors. Meanwhile, styles produced on more than one floor as opposed to only on a single floor do not differ on observables like technical complexity (SMV), or on how many styles of the same type (t-shirt, pants,...), or from the same buyer the factory produces (Appendix A.2). In total, 533 out of the 1,306 styles produced on more than one line are produced on more than one floor.

I control for fixed effects α_{fin} on the line chief - style day level. Thus, I estimate the effect of previous output of the same style (on the same floor) as a deviation of line productivity from learning curves estimated separately and non-parametrically for each line chief.[9] Furthermore, I include time fixed effects γ_{ftn} on the factory - month - style day level. Standard errors in this section are clustered at the line chief level.[10]

[8] More precisely, I use the log of previous output on other lines plus one. Alternative functional forms for past output are discussed, e.g., in Thompson (2007)

[9] The later results on the experimental intervention use line chief fixed effects because the intervention treats line chiefs. Thus, for consistency, I also use line chief fixed effects in this section. All results do not differ qualitatively when using line fixed effects instead of line chief fixed effects.

[10] Results are qualitatively the same when clustering at the style level, or two-way on style and line chief level.

3.1 General Results

Column 1 of Table 2 shows the results from estimating equation 1. We can see a significant effect of previous output on all other lines for at least the first eight days a line produces a garment. Due to space constraints, I only show the coefficients for the first 10 production days, or 'style days' on a line; effects for later production days are not significant anymore. Looking at the coefficients for previous output from lines on the same floor indicates that spill-overs within floors are two to three times stronger, particularly during the first three production days. After about seven days, it no longer makes a difference.

These effects, however, could be due to selection; e.g. styles that can be produced more efficiently are rolled out across more lines. Therefore, Column 2 repeats Column 1, but controlling for style fixed effects interacted with 'style day' fixed effects. Now, over the first three production days, productivity is enhanced only when the style has previously been produced on the same floor, but not on others. On the fourth day of production, the effect for overall previous output is statistically significant, but this could be a statistical anomaly, given that for the fifth day again, output from the same floor has a significant effect. From the sixth day onwards, there are no statistically significant effects anymore.

To better gauge the size of these effects, Columns 3 and 4 replicate Columns 1 and 2, respectively, but use dummies for positive previous output of the same style (on the same floor) instead of log previous output (on the same floor). The estimated effects are qualitatively very similar, especially when using style fixed effects (Column 4). If another line on the same floor has previously produced the same style, productivity is increased on average by a bit more than four productivity units, or about a quarter of the average productivity drop of about 16-17 productivity units on the first day a line produces a new style, as shown in Figure 1. The results also do not seem to be driven by later lines being under higher production pressure from management. While I do not have data on delivery deadlines for the styles, controlling for basic proxies for these deadlines does not change the results qualitatively, as shown in Appendix B.1.[11]

[11] The strength of the spill-overs do not differ with respect to whether or not other lines are producing the style *on the day* line i starts the style (results available on request).

Table 2: General Productivity Spill-Over

	(1)		(2)		(3)		(4)	
Specification of Indep. Variables:	Log Prev. Outp.		Log Prev. Outp.		Prev. Outp. >0		Prev. Outp. >0	
Cumul. Previous Output x ...								
Day 1	0.296**	(0.12)	0.138	(0.14)	1.380	(1.20)	0.975	(1.15)
Day 2	0.292***	(0.09)	0.173	(0.11)	1.609**	(0.81)	1.284	(0.90)
Day 3	0.204**	(0.09)	0.138	(0.09)	1.019	(0.79)	0.838	(0.73)
Day 4	0.327***	(0.09)	0.218**	(0.10)	2.337***	(0.85)	1.673**	(0.77)
Day 5	0.229***	(0.07)	-0.041	(0.10)	1.477**	(0.72)	-0.904	(0.84)
Day 6	0.215**	(0.09)	0.070	(0.11)	1.521	(0.81)	0.738	(0.96)
Day 7	0.161*	(0.10)	0.074	(0.13)	1.038	(0.91)	0.439	(0.92)
Day 8	0.256**	(0.11)	0.132	(0.13)	1.876*	(1.10)	0.811	(1.01)
Day 9	0.150	(0.11)	-0.019	(0.14)	0.505	(0.99)	-1.205	(1.22)
Day 10	0.248**	(0.12)	0.223	(0.16)	1.464	(1.11)	1.140	(1.39)
Cumul. Previous Output Same Floor x ...								
Day 1	0.516***	(0.13)	0.392***	(0.12)	4.746***	(1.22)	3.259***	(1.03)
Day 2	0.291***	(0.10)	0.185*	(0.10)	2.729***	(0.84)	1.714**	(0.82)
Day 3	0.258***	(0.09)	0.198**	(0.09)	2.516***	(0.78)	2.004***	(0.72)
Day 4	0.161*	(0.09)	0.091	(0.09)	1.487*	(0.79)	0.755	(0.81)
Day 5	0.214**	(0.08)	0.220**	(0.10)	1.882**	(0.75)	2.001**	(0.86)
Day 6	0.218**	(0.11)	0.172	(0.12)	1.825**	(0.91)	0.915	(0.95)
Day 7	0.193*	(0.10)	0.079	(0.12)	1.877**	(0.91)	0.658	(0.96)
Day 8	0.144	(0.11)	0.094	(0.12)	1.299	(0.96)	0.803	(0.98)
Day 9	0.000	(0.11)	0.102	(0.14)	0.014	(0.94)	0.859	(1.19)
Day 10	0.124	(0.12)	0.051	(0.16)	1.211	(1.07)	0.247	(1.50)
N	30,392		30,392		30,431		30,431	
Controls	YES		YES		YES		YES	
Factory-Month FE	YES		YES		YES		YES	
Line Chief FE	YES		YES		YES		YES	
Style FE			YES				YES	

Notes: Column 1 reports effects of regressing daily line productivity on the first 10 days a line produces a new style (first 10 "style days") on log cumulative previous output of the same style on other lines, and on other lines located on the same floor, interacted with fixed effects for the first 10 style days. Column 2 adds style fixed effects interacted with style day fixed effects. Columns 3 and 4 replicate Columns 1 and 2, respectively, but use instead dummies indicating positive amounts of previous output as independent variables. Controls are SMV, daily runtime, and number of workers on line. All Controls and FE in turn interacted with style day fixed effects. Due to computational constraints, regressions estimated separately for each of the ten 'style days' that a line produces a new style. Note that this does not affect the coefficient estimates as all independent variables are interacted with style day fixed effects. 'N' refers to summed 'N' for all ten of these regressions. Standard errors clustered at the line chief level in brackets: *** p<0.01, ** p<0.05, * p<0.1.

Due to the large number of estimated coefficients, gauging the aggregate effect of previous output over the learning curve is challenging. Furthermore, the treatment effects are not independent across style-days, making obtaining correct standard errors a concern. For these reasons, Appendix B.2 shows the results from fitting a basic non-linear model of learning curves to the data, which shows that the total productivity loss over the learning curve phase is reduced by 11 percent if some other line in the factory has produced the style before, and by an additional 32 percent if that line is located on the same floor.

3.2 Same Floor and Neighboring Lines

The results from Table 2 indicate that productivity spill-overs are a phenomenon that mainly occur within floors. My data does not allow me to study cases where more than one floor fall

under one common floor level management, to test whether spill-overs can be sustained across floors if they are under a common management. But I can test whether conditional on lines being located on the same floor, spill-overs are stronger if they are located more closely to each other, for example as neighboring lines. Table 3, Columns 1 and 2 repeat Columns 1 and 2 from Table 2, but add a third set of regressors, log previous output of the same style produced on neighboring lines on the same floor. The results are somewhat inconclusive. While output from the same floor seems to play a larger role during the first one to two style days, output from neighboring lines seem to play a stronger role in subsequent days. But when controlling for style fixed effects (Column 2), the effects from neighboring lines are never significant at the five percent level.

The inclusion of log previous output of the style in the whole factory, from the same floor, and from neighboring lines leads to the presence of three sets of highly correlated variables, making the estimation more unreliable. To mitigate against this concern, Column 3 and 4 of Table 3 repeat Columns 1 and 2, but only on the subset of styles that are produced on only one floor. Thus, previous output in the whole factory and on the same floor are the same and one of the two can be dropped. In this specification, output from neighboring lines is never significant, while previous output from the same floor is highly significant, at least on the first two days of production when controlling for style fixed effects (Column 4). Results are qualitatively the same when using the dummy specification for positive amount of previous output instead of its log (Column 5). It thus seems that the crucial threshold in proximity for the strength of spill-overs is to be on a the same floor. This could either suggest that organizational proximity is the more important dimension along which spill-overs vary, or spatial proximity matters discontinuously at the floor boundary. Any more precise disentangling of the two dimensions would have to be left to future research.

3.3 Identifying Assumptions

Given the use of line-chief fixed effects, spill-overs in this section are identified solely from within line chief variation in productivity. The identifying assumption then is that line chief - garment type effects on productivity are uncorrelated with start-ranks of lines in the roll-out of styles

Table 3: Productivity Spill-over, with Neighbouring Lines

	(1)		(2)		(3)		(4)		(5)	
Specification of Indep. Variables:	Log Prev. Outp.		Log Prev. Outp.		Log Prev. Outp.		Log Prev. Outp.		Prev. Outp. >0	
Cumul. Previous Output x ...										
Day 1	0.295**	(0.12)	0.147	(0.14)						
Day 2	0.292***	(0.09)	0.174	(0.11)						
Day 3	0.204**	(0.09)	0.150*	(0.09)						
Day 4	0.327***	(0.09)	0.221**	(0.10)						
Day 5	0.228***	(0.07)	-0.032	(0.10)						
Day 6	0.216**	(0.09)	0.087	(0.11)						
Day 7	0.161*	(0.10)	0.081	(0.12)						
Day 8	0.256**	(0.11)	0.140	(0.13)						
Day 9	0.150	(0.11)	-0.020	(0.14)						
Day 10	0.246**	(0.12)	0.225	(0.16)						
Cumul. Previous Output, Same Floor x ...										
Day 1	0.383**	(0.15)	0.299**	(0.14)	0.844***	(0.20)	0.838***	(0.18)	6.755***	(1.66)
Day 2	0.210*	(0.12)	0.168	(0.11)	0.482***	(0.16)	0.391***	(0.13)	3.063***	(1.07)
Day 3	0.162	(0.10)	0.101	(0.10)	0.435***	(0.15)	0.265	(0.21)	1.991	(1.72)
Day 4	0.107	(0.10)	0.064	(0.10)	0.436***	(0.15)	0.280	(0.20)	1.947	(1.66)
Day 5	0.127	(0.10)	0.129	(0.12)	0.234	(0.16)	-0.029	(0.19)	-0.623	(1.58)
Day 6	0.075	(0.12)	0.043	(0.13)	0.370**	(0.17)	0.178	(0.18)	1.080	(1.54)
Day 7	0.076	(0.11)	-0.018	(0.13)	0.376**	(0.15)	-0.104	(0.19)	-0.780	(1.56)
Day 8	0.077	(0.14)	-0.029	(0.12)	0.324*	(0.18)	0.041	(0.22)	-0.021	(1.74)
Day 9	0.024	(0.12)	0.123	(0.14)	0.033	(0.17)	-0.086	(0.21)	-1.586	(1.80)
Day 10	0.045	(0.13)	-0.049	(0.15)	0.460**	(0.19)	0.193	(0.42)	-0.045	(3.19)
Cumul. Previous Output, Neighbouring Lines x ...										
Day 1	0.257*	(0.14)	0.177	(0.13)	0.296	(0.24)	0.121	(0.23)	1.013	(1.89)
Day 2	0.155	(0.11)	0.034	(0.11)	0.141	(0.17)	0.144	(0.16)	1.697	(1.34)
Day 3	0.178**	(0.09)	0.178*	(0.10)	0.190	(0.16)	0.341	(0.22)	3.176*	(1.83)
Day 4	0.100	(0.09)	0.050	(0.11)	0.109	(0.18)	0.090	(0.19)	1.187	(1.55)
Day 5	0.168*	(0.09)	0.178	(0.12)	0.034	(0.18)	-0.049	(0.26)	-0.048	(2.19)
Day 6	0.281**	(0.11)	0.247*	(0.13)	0.138	(0.20)	-0.078	(0.21)	0.143	(1.69)
Day 7	0.224**	(0.10)	0.184	(0.13)	0.025	(0.17)	0.103	(0.19)	0.833	(1.64)
Day 8	0.130	(0.12)	0.233*	(0.14)	-0.010	(0.20)	0.152	(0.27)	1.437	(2.33)
Day 9	-0.046	(0.11)	-0.040	(0.15)	0.048	(0.20)	-0.079	(0.40)	0.021	(3.38)
Day 10	0.153	(0.13)	0.190	(0.17)	-0.202	(0.31)	-0.314	(0.45)	-2.516	(3.51)
N	30,392		30,392		11,302		11,302		11,311	
Factory-Month FE	YES		YES		YES		YES		YES	
Line Chief FE	YES		YES		YES		YES		YES	
Controls	YES		YES		YES		YES		YES	
Style FE			YES				YES		YES	

Notes: Table expands on Columns 1 and 2 of Table 2, adding log previous output on neighboring lines as additional independent variables (interacted with style day fixed effects). Columns 3-5 constrain sample to styles that were produced on a single floor only. Controls are SMV, daily runtime, and number of workers on a line. All Controls and FE in turn interacted with style day fixed effects. Due to computational constraints, regressions were estimated separately for each of the first ten 'style days' that a line produces a new style. Note that this does not affect the coefficient estimates as all independent variables are interacted with style day fixed effects. 'N' refers to summed 'N' for all ten of these regressions. Standard errors clustered at the line chief level in brackets: *** p<0.01, ** p<0.05, * p<0.1.

across lines. Or, put differently, on average, lines do not produce garments which they are better at as later (or earlier) lines in the factory. We have seen before that lines mostly do not specialize in producing certain types of garments. However, this does not rule out that lines produce those types of garments earlier (or later) that they are particularly good at.

A sufficient condition for absence of such correlation is the absence of any correlation between garment types and start ranks within each line. In Factory 1 and 2, this condition is fulfilled; fixed effects on the interaction of production lines and garment types (dress-shirts, t-shirts, pants...) have no jointly significant predictive power on the start rank of styles on lines. At the third factory, the line - garment type fixed effects are jointly significant. Note, however, that this is a very strong test, testing for absence of joint significance of 515 fixed effects. And the pattern that would induce an upward bias in the estimates, that lines produce those styles they are better at *later* in the roll-out process seems conceptually unlikely. If anything, we would expect factories to allocate a style first to lines that are good at producing it. This more likely pattern would induce a downward bias in the estimates of spill-over strength, which would imply that the estimates presented here are lower bounds for the true spill-overs.[12] Note furthermore that all results of the paper are broadly the same when estimated just on the sub-set of Factory 1 and 2, where there is no correlation detected.

Most of the literature on organizational learning has used evidence of productivity spill-overs as presented in this section as evidence for knowledge exchange in firms. However, such effects could also be driven by other peer effects, such as competition between workers on different lines producing the same style. However, if the spill-overs are caused by mere differences in worker effort, we could expect such effects to appear also outside the learning curve, as the scope for effort to affect productivity should not be constrained to the learning curve phase. Finally, the next section will study the management intervention which increased the potential for knowledge exchange on randomly selected floors. If the intervention results in similar productivity spill-overs as those observed in this observational part, it would be further evidence for knowledge exchange being a main driver of the spill-overs.

[12] Note that a direct test for correlation of average productivity on the line - garment type level with start-rank of the garments of that type on the line is not valid, as productivity spill-overs through learning would exactly induce such a correlation.

4 Randomized Communication Intervention

The randomized management intervention mandated that whenever a line on randomly selected "treatment" floors began producing a new style that had already been produced by any other line in the factory, the most senior line chief with previous experience of the style should be instructed by the production management to brief the line chief who is now also starting the same style. The 15- to 30-minute brief was to be held at the line of the line chief receiving the briefing, and was intended to explain how the earlier line had overcome initial problems, particularly at "bottleneck" operations that slowed down the new style's production.[13]

The experiment ran on the treatment floors for four months, from June to end of September 2014. The production data show 687 "non-first" style starts at the three factories during these four months; these are style starts at which some other line chief in the factory had already produced the same style and therefore the line chief now starting the style could have been briefed in the management intervention by a line chief already experienced with the style. Out of the 687 style starts, 393 occurred on lines on treatment floors, and therefore should have been accompanied by a briefing for the line chief, while the remaining 294 on control floors should have not. The treatment protocol was implemented by the production engineers from the factories. The engineers were provided with experimental logbooks to record each instance of a treatment of a style start. According to these logbooks, 154 briefings among line chiefs were conducted, of which 98 could be matched with a style start in the production data.[14] However, it is likely that compliance was higher than indicated by these numbers. The implementing engineers admitted underreporting of treatments in the logbooks. Among the actually treated style starts, there is likely to be selection into treatment of starts for which the treatment was

[13] When lines switch to a new style, the adjustment of machines is typically done by the line chiefs together with a production engineer, who also briefs the line chief on the new style. The adjustments are typically made machine after machine going down the line, so that workers on machines already changed can start producing the new style while workers on not-yet-adjusted machines still apply to last stitches to the previously allocated style. The briefings from the communication intervention were to have been done once all machines on the line had been adjusted for the new style. Its goal was to convey the knowledge held by a line chief who had already produced a style for several days (e.g. how to hold a garment in the hand when doing a certain stitch) – expertise that the production engineers do not have, who design the production line layout for a style from its design template, but do not stay on a line after the switch.

[14] 26 further briefings of the 154 recorded ones could also be matched to style starts on treatment floors in the production data. However, according to the production data, in these instances no other line chief had previously produced the style. Thus, according to the protocol, no briefing should have taken place. Possibly, in these cases, line chiefs who had produced technically similar styles were sent to give instructions.

expected to have a stronger effect.[15] The log-books show no briefings to have taken place on control floors. For these reasons, the analysis will focus on the intention-to-treat effect. Appendix C.1 shows results when regressing productivity on recorded treatments.

4.1 Randomization and Balance

Nine floors from a sample of 17 floors across the three factories were randomly selected to receive the treatment, stratified at the factory level.[16] Table 4 shows balance tests over the characteristics of lines and line chief on treatment and control floors from April and May 2014, just before the start of the intervention. Unfortunately, some imbalance is visible. Lines randomized into treatment have lower average productivity, and somewhat longer average daily runtime. Given that daily runtime enters (negatively) the productivity measures as shown in section 2, these two imbalances may be co-determined, though controlling the imbalance in efficiency for runtime does not eliminate the difference. The randomization of floors into treatment and control was done in presence of the factory managers to achieve buy-in into the experiment, which precluded potential re-randomization. I discuss strategies to account for this imbalance in the reminder of this section.

Because the intervention was conducted at the end of the time covered by the collected production data, a substantial amount of pre-intervention data is available. Figure 4 plots the average productivity over the first five days a line produces a new style for four different cases: treated lines before and during the intervention months, and control lines before and during the intervention months. I use data from the beginning of 2014 until the end of September 2014, when the initially agreed upon treatment time ended, and only include data from the non-first style starts, which were "treatable" because another line chief has previously produced the same style. Prior to the start of the intervention, from January-May 2014, compared to control lines

[15] All three factories reported that prior to the intervention, they occasionally sent line chiefs with experience on a style to other lines to help co-workers if they started producing the same style. However, this behavior was not institutionalized in any of the three factories. The factories were instructed to not change their behavior on the control floors.

[16] Factory 3 requested to include only six of its 14 floors in the sample. In fact, the sample of floors over which the randomization occurred consisted of only 15 sewing floors. However, at Factory 1 and 2, one floor was randomly chosen at each factory, split "in the middle" into two equally large halves, and then one half of the floor randomly selected into treatment. Therefore, the randomization occurred effectively across 17 units, 13 full floors, and 4 half floors.

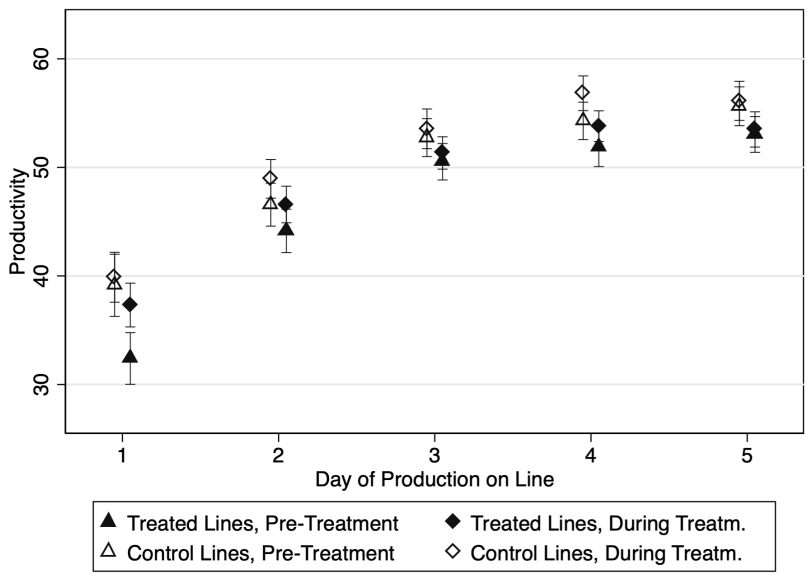

Figure 4: Pre-and Post- Treatment Learning Curves of Treated and Control Lines. Average productivity over the first five days a line produces a new style previously produced on other lines for four different cases: on treatment floors prior to start of treatment, treatment floors during experiment, control floors prior to start of treatment, and control floors during treatment. Productivity standardized at factory-level. Capped bars represent 95% confidence intervals.

at the same time, treated lines had on average lower productivity values, as already shown in Table 4. However, while productivity remains constant across pre-treatment and treatment time on control floors, treated lines experience an upward shift in their learning curves during the time of the treatment.

Using a difference-in-difference design to estimate the effects of the intervention, the identifying assumption is that absent the intervention, productivity levels on the treatment floors would have followed the same trends as on the control floors. Figure 5 plots, for the first day a line produces a new style that has already been produced by another line chief, monthly average productivity from January to September 2014, separately for lines from treatment (square symbols) and control floors (triangle symbols). Given that data from Factory 3 is only available from April 2014 onwards, the graphs are shown separately for Factory 3 on the right column of Figure 5. The solid vertical line in the graphs indicates the start of the treatment with June 2014, while the dashed line indicates the end of the intervention period at the end of September 2014. Log-book entries indicate that the briefings continued to some extent, but anecdotal

evidence suggest that from then on they were also extended to control floors.

The upper two graphs show the productivity trends when only controlling for SMV, daily runtime and number of workers on the line, while the lower two graphs also control for style fixed effects. As already suggested in Table 4, first-day productivity was systematically lower on floors selected for treatment in the months before the start of the treatment, with the difference mostly driven by Factory 3. And while the data are noisy, productivity broadly follows the same trends at all factories in the months running up to the start of the intervention. The productivity difference is then offset after the start of the intervention due to an upward shift of productivity on treatment floors. During the first month of the intervention, though, if anything, productivity on treatment floors drops relative to control floors. However, in the remaining three months, it catches up and in some months at some factories overtakes productivity on the control floors. The fact that the effects are only visible from the second month of the intervention onwards could indicate that the intervention was initially not very effective in increasing productivity, possibly because the line chiefs needed to get used to the communication intervention.[17] Appendix C.2 shows productivity trends from Factory 1 and 2 for a longer time preceding the intervention, as from these two factories, more than one year of production line data before the start of the intervention is available.

4.2 Regression Specification

To estimate the intention-to-treat effect in a difference-in-differences approach, I keep, similar as in the previous section, the observations from the first five style days from each style-start. I restrict the sample, as in Figure 4, to style-starts from observations from January to September 2014 and to style-starts for which another line in the factory has already produced the same style. Using this sample, I run the following baseline regression, resembling the specifications from the previous section:

$$y_{fisnt} = \sum_n \beta_n^T Treat_{fis} + \beta_n^X X_{fisn} + \alpha_{fin} + \gamma_{ftn} + \epsilon_{fisnt} \tag{2}$$

"$Treat_{fis}$ is a dummy variable taking value one if line i is a treatment line and if it started

[17] The logbooks indicate that the share of style starts that was treated was roughly equal across the months, except for July 2014, in which the share was roughly twice as large.

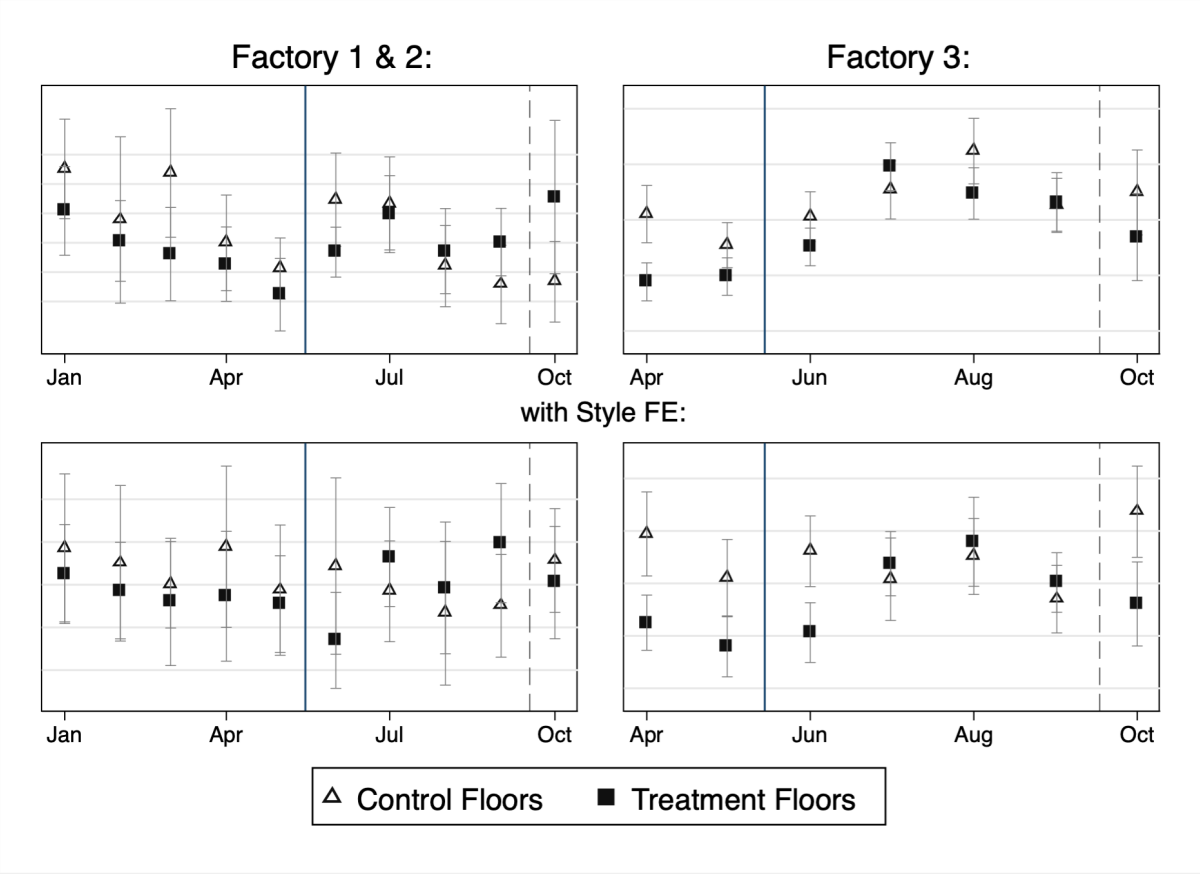

Figure 5: Pre/Post Start of Intervention Trends for First-Day Productivity. Figures shows average monthly productivity of lines on the first day they start producing a new style that has already been produced on another line in the factory, separately for lines on treatment floors (solid squares) and on control floors (hollow triangles). The solid vertical line indicates start of treatment from June 2014 on. The dashed line indicates the end of the treatment period at the end of September 2014. Capped bars represent 95% confidence intervals.

producing style s for the first time during the treatment time June-September 2014. As in the previous section, I interact this variable with 'style day' fixed effects. I also control again for fixed effects a_{fin} at the line-chief - 'style day' level, and γ_{ftn} on the factory - month - 'style day' level. X_{fisn} is a set of controls that includes SMV of the style, the number of hours the line ran on the day, and the number of workers on the line, plus a set of fixed effects for the number of line chiefs in the factory that previously produced the style. Due to the small number of 17 clusters over which the randomization was conducted, standard errors clustered on the 17 floors can be biased downwards. I follow two approaches to obtain more accurate standard errors, similar to Papay et al. (2020); wild cluster bootstrap, as suggested by Cameron et al. (2008), and randomisation based standard errors, following Young (2018).[18]

4.3 Results of Intervention

Table 5 shows the main results from the communication intervention. Column 1 shows the results from the basic regression shown in equation 2 above. Stars indicate statistical significance based on wild-cluster bootstrap, while daggers based on randomisation inference. The brackets show p-values for both bootstrap and randomisation based inference. A significant positive effect on productivity on the first day of production of a new style can be seen. The effect on the first day is roughly five productivity units, or about 35% of the remaining average productivity drop on the first day lines produce new styles they did not produce before, but had been produced on some other line before. On the second day the effect is smaller and at most only marginally significant. From the third day onwards, no effects are discernible anymore. In Column 2, I again add style fixed effects (interacted with style-day fixed effects). This is a demanding specification, as it only uses variation in productivity within the same style and 'style day', either across treatment and control floors, or across pre-treatment and treatment periods. Nevertheless, the effect remains almost unchanged and marginally statistically significant on the first two production days.

[18]For randomisation based standard errors, I randomly reassign treatment 500 times across the 17 floors, keeping the original randomisation protocol, and use the resulting distribution of the estimates to calculate two-sided p-values, by taking twice the share of estimated coefficients of re-randomized treatments that are larger or equal to the actual treatment effect. I use coefficients of re-randomised treatments instead of t-statistics as they provide more conservative p-values in these regressions.

Table 4: Balancing of Randomization across Sewing Lines

Variable	Control	Diff.	S.E.	N
Line Characteristics:				
Nbr. Worker	29.74	-0.06	(0.76)	142
Daily Runtime	10.37	0.23*	(0.12)	142
Efficiency	65.31	-3.95***	(1.20)	142
SMV	7.985	-0.41	(0.55)	142
Start Rank	3.817	0.38	(0.25)	142
Importance Buyer	0.134	0.01	(0.01)	142
Importance Garment Type	0.459	-0.04	(0.03)	142
Line Chief Characteristics:				
Age	29.18	0.24	(1.04)	84
Seniority Factory	49.68	4.51	(8.47)	77
Seniority as Supervisor	31.69	0.69	(5.36)	84
Sen. as SV on curr. line	22.14	0.64	(4.73)	74
External Arrival as SV	0.414	-0.00	(0.11)	77
Nbr. Social Ties	2.645	0.48	(0.61)	84
Education	15.19	-0.32	(0.48)	77

Notes: Line Chief characteristics from line chief surveys. Line characteristics from production data. 'Importance buyer' is share of styles ordered by buyer among all styles factory produces. 'Importance Garment Type' is share of styles of same type among all styles factory produces. "Control" column shows average values from control floors from April and May 2014. "Diff. " column shows deviation of average values on treatment floors from those from control floors, controlling for factory fixed effects: *** p<0.01, ** p<0.05, * p<0.1.

Table 5: General Treatment Effects

	(1)		(2)		(3) Reweighted:	
Treatment x						
Day 1	5.440***†	(0.01/0.08)	5.428*†	(0.08/0.07)	6.151**††	(0.04/0.03)
Day 2	3.769*	(0.08/0.16)	3.227**†	(0.04/0.06)	3.362**†	(0.03/0.07)
Day 3	2.838	(0.40/0.42)	3.462	(0.32/0.12)	3.015	(0.44/0.34)
Day 4	1.915	(0.50/0.54)	4.061	(0.20/0.22)	4.438	(0.17/0.23)
Day 5	1.587	(0.73/0.71)	4.351	(0.27/0.20)	4.600	(0.28/0.24)
N	4,946		4,946		4,615	
Controls	YES		YES		YES	
Factory-Month FE	YES		YES		YES	
Line Chief FE	YES		YES		YES	
Style FE			YES		YES	

Notes: The table shows results from regressing an ITT treatment dummy of the line chief receiving a briefing when starting a new style, on the productivity on each of the first five days his/her line produces the new style. Sample only includes observations when line chiefs start styles that had already been produced by another line chiefs before. Column 2 and 3 controls for style fixed effects. Column 3 uses reweighting technique based on DiNardo et al. (1996) to control for differential pre-treatment productivity of treatment and control lines (see Appendix C.3). Controls are SMV, daily runtime, and number of workers on line. All controls and fixed effects interacted with factory and style day fixed effects. 'N' shown is sum across the individual 'N' for these five regressions. Note that this does not affect the coefficient estimates as all independent variables are interacted with style day fixed effects. Brackets show (two sided) p-values based on wild-cluster bootstrap (*) and randomisation tests (†): ***/††† p<0.01, **/†† p<0.05, */† p<0.1.

As pointed out before, the fact that the effects of the intervention, similar to the general spill-over effects in the previous section, are only visible during the first days a line produces a new style supports the view that the intervention effects are indeed due to knowledge exchange, and not due to alternative mechanisms, such as the briefing simply informing the line chief that other lines also produce the style, thus triggering competitive peer effects. Such peer effects could be expected to continue for longer, particularly if the management discounts the first production days when comparing productivities of different lines producing the same style. Also, line chiefs are generally aware of the styles being produced by other line chiefs, at least by those on the same floor, given that the style currently produced on a line is usually displayed prominently at different places on the floor. This speaks again against the hypothesis that the intervention merely informed the supervisors that other lines are also producing the same style.

The lower baseline productivity of lines on treatment floors could imply that the results are caused by some other form of catch-up of productivity on treatment floors relative to control floors, which coincided with the start of the intervention. To address this concern, I apply the reweighting approach by DiNardo et al. (1996). I use the approach in a similar way as Duflo et al. (2013), who reweight observations from a controlled experiment such that independent variables that were not balanced pre-treatment between treated and control units become balanced after the reweighting. As shown in Column 3 of Table 5, my results, when reweighting the data such that average productivity on the first day lines produce new garment do not differ anymore between treatment and control floors, do not differ qualitatively from the ones in the previous columns. Appendix C.3 provides more technical details on this reweighting approach.

Figure 6 shows distributions of line productivities on the first day lines produce new styles already produced on other lines, for treatment lines before (Jan-May 2014) and during (Jun-Sep 2014) the implementation of the experiment, and control lines during the same time periods. The increase in first-day productivity of treatment lines during the intervention is driven by a strong reduction in the left tail of the productivity distribution, which is indicative of the individual treatments being enacted specifically when very low productivity could have been expected. This fits with the fact that fewer treatments were reported in the logbooks than

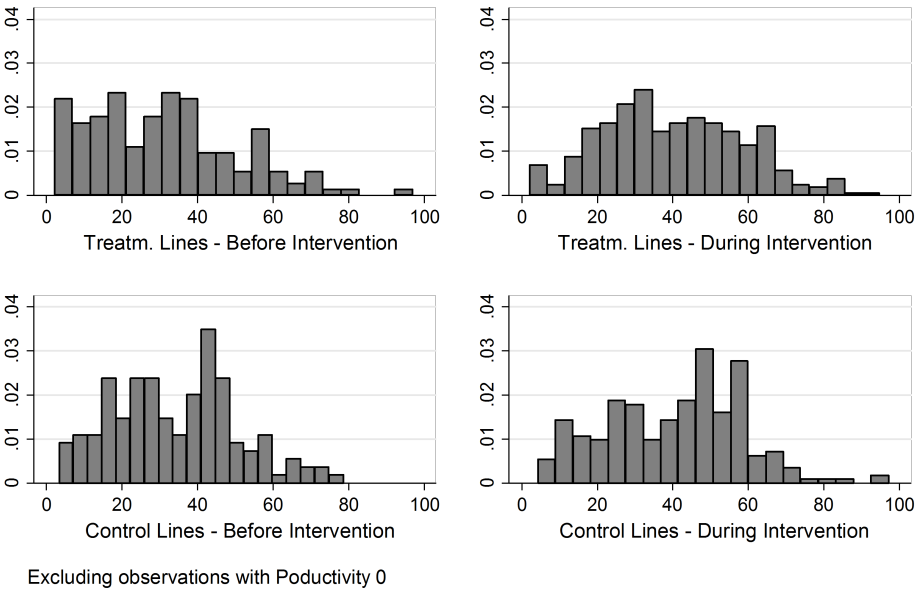

Figure 6: First Day Productivity Distribution before and during Intervention. The graph shows distribution of productivity on first day sewing lines produce new styles previously produced on some other line in the factory, on treatment floors, before (Jan-May 2014) and during (Jun - Sept 2014) implementation of intervention (top row), and on control floors, before and during the implementation (bottom row).

should have been according to the production data.

While the intervention was randomly allocated to floors, one possibility is that the allocation of styles to lines reacted endogenously to the treatment during the four month treatment period. To test for this possibility, I replicated Table 5, but with style complexity (SMV), share of style's type among total output of factory, and share of style's buyer among total output, as dependent variable. No equivalent "treatment-effects" can be found for these style characteristics (results available on request).

4.4 Briefings from Supervisors from Same Floor vs Other Floors

The results from the observational analysis using the overall dataset have shown that productivity spill-overs are predominantly occurring within floors. We can test whether this result is also found in the randomized intervention. However, given the incomplete documentation of the actual briefings, it is often not clear which line chief conducted a briefing. The instruction given to the management was that the most senior of the line chiefs that previously produced

Table 6: General Treatment Effects, with Other Floors

	(1)		(2)		(3) Reweighted	
Treatment x						
Day 1	7.811***†	(0.00/0.06)	10.303**	(0.02/0.14)	11.955***††	(0.01/0.02)
Day 2	7.008***†	(0.01/0.06)	6.409**	(0.01/0.10)	7.629**††	(0.02/0.04)
Day 3	6.758**	(0.04/0.13)	8.623**††	(0.02/0.05)	8.594**†	(0.02/0.07)
Day 4	3.386	(0.26/0.42)	6.017	(0.11/0.28)	6.953*	(0.07/0.29)
Day 5	1.714	(0.76/0.62)	3.414	(0.38/0.45)	3.569	(0.42/0.43)
Treatment x Other Floors x						
Day 1	-4.552**†	(0.04/0.10)	-9.145**	(0.02/0.15)	-10.491**†	(0.01/0.10)
Day 2	-6.410**	(0.03/0.21)	-6.229	(0.15/0.34)	-8.122**	(0.04/0.23)
Day 3	-8.037***†	(0.01/0.07)	-11.210***††	(0.01/0.02)	-11.740**††	(0.02/0.02)
Day 4	-3.307	(0.11/0.44)	-5.000*	(0.09/0.21)	-6.207**†	(0.02/0.09)
Day 5	-0.290	(0.91/0.65)	1.929	(0.26/0.10)	1.919†	(0.35/0.09)
Other Floors x						
Day 1	1.185	(0.54)	4.017	(0.17)	4.961*	(0.08)
Day 2	0.736	(0.77)	3.177	(0.48)	4.238	(0.33)
Day 3	1.497	(0.38)	4.274	(0.11)	5.381*	(0.06)
Day 4	1.611	(0.23)	3.497***	(0.01)	4.054**	(0.01)
Day 5	0.338	(0.88)	1.381	(0.52)	1.886	(0.43)
N	4,946		4,946		4,615	
Controls	YES		YES		YES	
Factory-Month FE	YES		YES		YES	
Line Chief FE	YES		YES		YES	
Style FE			YES		YES	

Notes: Table redoes Table 5, but interacting each of the five treatment effects for the first five style days with an dummy variable "Other Floors" that indicates that all lines that had produced the style before were located on other floors. Thus, treatment (briefing) must have been conducted by a line chief from another floor. As in Table 5, sample only includes observations when line chiefs start styles that had already been produced by other line chiefs before. Regressions controls for fixed effects for number of line chiefs who already produced same style. Further controls are SMV, daily runtime, and number of workers on line. All controls and fixed effects interacted with factory and style day fixed effects. Brackets show (two sided) p-values based on wild-cluster bootstrap (*) and randomisation tests (†): ***/††† p<0.01, **/†† p<0.05, */† p<0.1. No randomisation based standard errors for 'Other Floors (x Day X)', as it was not a randomly assigned variable in intervention.

the garment should conduct the briefing. But among the documented briefings, it is shown that the management deviated at times from this instruction, and could have done so selectively. For example, they might have sent line chiefs from other floors only if they expected a large effect of the briefing. This would bias upwards the estimate of the treatment effect in these cases. For this reason, I interact the ITT treatment variable with dummy variable indicating that lines that have previously produced the garment are all located on other floors. In that case, the briefing could have only been done by a line chief from another floor, which was the case for 165 out of the 393 briefings that should have taken place on treatment floors.

Table 6 shows the results from this interaction. It replicates Table 5 but adds dummies for each of the first five production days indicating that all lines that previously produced the garment are on other floors ("Other Floors x Day X"), and interactions of these dummies with the ITT treatment variable for each of the first five production days ("Treatment x Other Floors x Day X"). The overall treatment effect on the first days a line produces a garment that has already been produced on the same floor now increases to around seven to 12 productivity units, depending on the different specifications in Columns 1 to 3, or 50-80% of the average productivity drop when lines switch to new styles previously produced on other lines before. The effect of the intervention now remains significant for around the first three production days, similar the lengths of the average learning curve when lines start new garments. However, the interaction effect with the line chief having to come from other floors offsets 65-100% of this treatment effect. This indicates that the communication intervention was little effective when the garment was previously produced only on other floors.

I also adapt the estimation of the non-linear model laid out in the section on observational results to the estimation of the treatment effects of the randomized intervention. The details of the estimation and the results are laid out again in Appendix B.2. The estimates shown in Column 3 of Table B.2 suggest that the treatment reduced the productivity loss over the whole learning curve by around 41 percent. Column 4 interacts the treatment indicator variable T again with an indicator variable O for all lines that already produced the style being located on other floors. The results are consistent with the main ones from Table 6; if the briefing could

have been made by a line chief from the same floor, the productivity losses over the learning curve are reduced by around 60%, while if only line chiefs on other floors have produced the style before, two thirds of this effect are offset again. These results show that the treatment effects are not just artefacts of estimations with large amount of fixed effects, but are also visible when fitting a highly stylised model of learning curves to the data.

4.5 Possible Mechanisms

While the larger effects of the intervention in case the briefings could be done within floors appear consistent with the generally stronger spill-overs within floors as shown in the observational part (Table 2), possible mechanisms that cause these stronger treatment effects are difficult to pin down with the available data. Ex-ante, the direction of the interaction effect of the treatment was ambiguous. If the weaker spill-overs across floors were solely due to communication across floors being more costly, or due to the fact that line chiefs were less likely to know if lines on other floors produce the same style, then this intervention would be ideally suited to overcome these frictions, as it induced communication across floors regardless of the costs for line chiefs or their lack of knowledge. In this case, we would expect the intervention to have stronger effects for briefings *across* floors. Given that the empirical results go in the opposite direction, they speak against these hypotheses.

Unfortunately, I lack data to precisely pin down the mechanisms that can explain why we observe that the experiment further amplifies spill-overs within floors as opposed to across floors. But two pieces of information could point towards possible explanations. First, it may be that such an intervention is only effective between workers that already share some form of social capital or trust. The interventions studied in Sandvik et al. (2020) and Papay et al. (2020) could have produced their positive effects because they induced co-workers to meet repeatedly and thus build social ties between them. I can investigate this hypothesis by leveraging social network data that I collected from all line chiefs in the factory before the start of the experiment. Line chiefs are substantially more likely to report social connections (kinship, being 'friends', eating lunch together...) to line chiefs that work on the same floor as opposed to those working

on other floors. I run two tests. First, I test whether the stronger within floor treatment effects are explained by the density of social networks within floors (share of reported links over all possible links). However, variation in this density across floors cannot explain variation in the strengths of within floor spill-overs. Second, I test if the treatment effects are stronger if among the line chiefs that previously produced the style (and thus could have done the briefing) are any that are socially connected to the line chief to receive the briefing. This would have given the management the opportunity to select a connected line chief to do the briefing had they deemed it beneficial. Indeed, I find that in these instances, the treatment effects were stronger, but only when not controlling for style fixed effects. Thus this effect is not very robust across specifications (results available on request).

The second piece of information that can point towards an explanation for the further amplification of the spill-overs within floors by the intervention is that in the log-book data, those briefings that could have only been done by a line chief from another floor were less likely recorded to have taken place, as shown in Table 7. If briefings were less likely implemented across floors, it would be a direct explanation for the weaker treatment effects across floors. This is further corroborated by the fact that the lower likelihood of briefings across floors being recorded was driven by Factory 3 (Column 4, Table 7), which is also the one that drives the result on the lower treatment effects across floors from Table 6, as shown in Appendix C.4.[19]

Why would the briefings be less implemented across floors? Ex-post surveys with the managements indicate that the day-to-day implementation of briefings that involved supervisors from different floors were more difficult to coordinate. Floor level managements preferred their line supervisors to stay on their floors, and often lacked information about which styles were already produced on other floors. Therefore, they were not very pro-active in coordinating briefings with supervisors from other floors. Relative to floor level managers, the central management and production engineering department lacked the man-power to overcome this lack of initiative from floor level managements. This may reflect an optimal allocation of most of the day-to-day production management capacity to the floor level, instead of keeping it at the fac-

[19]Note that the overall report rate of briefings was particularly low at Factory 1, which makes it more difficult to test whether briefings across floors were less likely reported to be happening than briefings that could be done within floors at that factory.

Table 7: Recorded Treatments, within & across Floors

	(1)	(2)	(3)	(4)
	Intervention Reported			
Other Floors Only	-0.166***	-0.168**	-0.184**	
	(0.055)	(0.065)	(0.073)	
Other Floors Only x Factory 1				0.013
				(0.073)
Other Floors Only x Factory 2				0.041
				(0.116)
Other Floors Only x Factory 3				-0.271***
				(0.092)
Mean Reported Rate	0.196			
Mean Reported Rate x Factory 1				0.044
Mean Reported Rate x Factory 2				0.422
Mean Reported Rate x Factory 3				0.211
Observations	393	393	351	351
Factory FE	YES			
Floor FE		YES	YES	YES
Controls			YES	YES

Notes: Table shows results when regressing an indicator variable that a briefing was reported to have taken place according to the experimental log-books, on an indicator variable that at the time when the line chief should have received the briefing, only line chiefs on other floors so far had produced the style, and thus could have provided the briefing. The sample consists of all instance where according to the production data and the research design, a briefing should have occurred. Controls are Line Chief age, education, seniority as line chief, Nbr of social ties to other line chiefs, average productivity of his line, and SMV of style. Standard errors clustered on line level of line chiefs receiving briefings: *** p<0.01, ** p<0.05, * p<0.1

tory level. However, this could come at the expense of cooperation across these organizational sub-divisions, resulting in lack of knowledge exchange, and ensuing productivity losses.

4.6 Return on Intervention

To obtain an idea of the overall effect of this intervention on factory productivity and profits, I make the conservative assumption that the intervention only affected line productivity on the first day the line produced a new style, increasing productivity on average by 5 units on that day, as suggested by the results from Table 5. Lines on average switch to a new style every 12 days, and at roughly every second start, another line has previously produced the style. Average daily productivity in the three factories is 47.4 efficiency points; thus, a very basic back-of-the envelope calculation shows that output was increased by $5/(12 * 2 * 47.4) = 0.44$ percent.

Labor costs make up around 12 percent of revenue on average in these factories, while the

profit margin is about 6 percent. If we assume that the intervention would save 0.44 percent of labor costs, this would translate into an increase in profits of almost 0.9 percent. On the other hand, the pure monetary costs of the intervention are low. The hourly wage of a line chief in the factories is about US$1; therefore, the wage cost of a half-hour briefing is about $0.50. In the largest of the three factories, with more than 180 sewing lines, roughly 3,000 briefings per year would be required to treat lines whenever they start a style that is novel to them but has previously been produced by another line. Thus, the yearly monetary cost of the intervention would be $1,500. With labor costs commonly estimated to make up around 12 percent of revenue and average monthly wages in the sector of around $100, we can estimate conservatively that a factory with 5,000 employees has revenues of around $50 million per year. Using the commonly referenced margin of 6 percent yields estimated profits of aroundt $3 million. A 0.9 percent increase would thus imply an increase in profits of around $27,000, making the estimated costs small in comparison. For an upper bound on the costs, we could conservatively assume that the factory level production manager also needs to dedicate half an hour of their work time per briefing, to keep track of when briefings should be done and to induce the line chiefs to conduct them. The median wages of production managers are about twice as high as for line chiefs, implying that the costs of the intervention would triple to $4,500, still very low relative to the estimated gains. These high estimated returns are also in line with anecdotal reports from the production managers in the post-intervention interviews, that the monetary costs of the intervention were not deemed a hindrance to its implementation.

5 Conclusion

This paper presents novel evidence on organizational learning from three large Bangladeshi garment factories, using a randomized management intervention to generate exogenous variation in knowledge exchange across different parts of the factories, similar to recent field experiments that paired workers to encourage communication of job-related skills (Sandvik et al. (2020); Papay et al. (2020)). Differently to these papers, which had a strong focus on repeated meetings, relationship building and exchange of broader job-specific knowledge, this intervention induces one-off briefings between workers to exchange knowledge on specific products the factories are producing. The intervention has positive effects on productivity in the first days lines produce

the new products, inducing steeper learning curves. However, it is mainly effective in increasing productivity spill-over within floors, less so across floors. These findings resemble the baseline situation in the factories before the start of the intervention, in which production lines had higher productivity when other lines on the same floor had already produced the same product before, but not lines on other floors. The results are most in line with coordination and information frictions between floors preventing the same amount of information exchange across floors than is possible within floors. This points to an important trade-off that factories may have to face; as they grow in size, and thus require a more complex organizational structure with a larger number of sub-divisions, they may loose their ability to seamlessly transfer important production knowledge between all employees in the firm, with potentially negative effects on their productivity.

References

Adhvaryu, A., N. Kala, and A. Nyshadham (2019). Management and Shocks to Worker Productivity. mimeo, University of Michigan.

Adhvaryu, A., A. Nyshadham, and J. Tamayo (2018). Managerial Quality and Productivity Dynamics. mimeo, University of Michigan.

Amodio, F. and M. M. Carrasco (2018). Input Allocation, Workforce Management and Productivity Spillovers: Evidence from Personnel Data. *The Review of Economic Studies 4*(85), 1937–1970.

Argote, L. (2013). *Organizational Learning*. Springer US.

Argote, L., S. l. Beckman, and D. Epple (1990). The persistence and transfer of learning in industrial settings. *Management Science 36*(2), 140–154.

Arrow, K. (1962). The Economic Implications of Learning by Doing. *The Review of Economic Studies 29*(3), 155–173.

Atkin, D., A. Chaudhry, S. Chaudry, A. Khandelwal, and E. Verhoogen (2017). Organizational Barriers to Technology Adoption: Evidence from Soccer-Ball Producers in Pakistan. *Quarterly Journal of Economics 132*(3), 1101–1164.

Battiston, D., J. B. i Vidal, and T. Kirchmaier (2017). Face-to-Face Communication in Organisations. Working Paper, London School of Economics.

Benkard, L. (2000). Learning and Forgetting: The Dynamics of Aircraft Production. *American Economic Review 90*(4), 1034–1054.

Black, S. E. and L. M. Lynch (2004). What's Driving the New Economy?: The Benefits of Workplace Innovation. *Economic Journal 114*(493), 97–116.

Bloom, Nicholas, B. E., D. McKenzie, A. Mahajan, and J. Roberts (2013). Does Management Matter: Evidence from India. *Quarterly Journal of Economics 128*(1), 1–51.

Bloom, N., R. Sadun, and J. V. Reenen (2016). Management as a Technology? Working Paper, Stanford University.

Bloom, Nicholas, R. S. and J. V. Reenen (2012). The Organization of Firms across Countries. *Quarterly Journal of Economics 127*(4), 1663–1705.

Breshnahan, T. F., E. Brynjolfsson, and L. M. Hitt (2002). Information Technology, Workplace Organizatio, and the Demand for Skilled Labor: Firm-Level Evidence. *Quarterly Journal of Economics 117*(1), 339–376.

Cameron, A. C., J. Gelbach, and D. Miller (2008). Bootstrap based Improvements for Inference with Clustered Errors. *Review of Economics and Statistic 90*(3), 414–427.

Catalini, C. (2018). Microgeography and the Direction of Inventive Activity. *Management Science 64*(9), 4348–4364.

Combs, J., Y. Liu, A. Hall, and D. Ketchen (2006). How much do High-Performance Work Practices Matter? A Meta-Analysis of their Effects on Organizational Performance. *Personnel Psychology 59*(3), 501–528.

David, G. and T. Brachet (2011). On the Determinants of Organizational Forgetting. *American Economic Journal: Microeconomics 3*(3), 100–123.

DiNardo, J., N. Fortin, and T. Lemieux (1996). Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semi-Parametric Approach. *Econometrica 64*(5), 1001–1044.

Duflo, E., M. Greenstone, R. Pande, and N. Ryan (2013). Truth-telling by Third-party Auditors and the Response of Polluting Firms: Experimental Evidence from India. *Quarterly Journal of Economics 128*(4), 1499 – 1545.

Egelman, C. D., D. Epple, L. Argote, and E. R. Fuchs (2016). Learning by Doing in Multiproduct Manufacturing: Variety, Customization, and Overlapping Product Generations. *Management Science 63*(2), 405–423.

Haltiwanger, J., R. S. Jarmin, R. Kulick, and J. Miranda (2016). High Growth Young Firms: Contribution to Job, Output and Productivity Growth. in: Measuring Entrepreneurial Businesses: Current Knowledge and Challenges. National Bureau of Economic Research, Inc.

Hamilton, B., J. Nickerson, and H. Owan (2003). Team Incentives and Worker Heterogeneity: An Empirical Analysis of the Impact of Teams in Productivity and Participation. *Journal of Political Economy 111*(3), 465–497.

Hendel, I. and Y. Spiegel (2014). Small Steps for Worker, a Giant Leap for Productivity. *American Economic Journal: Applied Economics 6*(1), 73–90.

Hjort, J. (2014). Ethnic Divisions and Production in Firms. *Quarterly Journal of Economics 129*(4), 1899–1946.

Huselid, M. A. (1995). The Impact of Human Resource Management Practices on Turnover, Productivity, and Corporate Financial Performance. *Academy of Management Journal 38*(3), 635–672.

Ichniowski, C., K. Shaw, and G. Prennushi (1997). The Effects of Human Resource Management Practices on Productivity: A Study of Steel Finishing Lines. *American Economic Review 87*(3), 291–313.

Kato, T. and P. Shu (2011). Competition, Group Identity, and Social Networks in the Workplace: Evidence from a Chinese Textile Firm. IZA Discussion Paper 6219.

Levitt, S., J. List, and C. Syverson (2013). Towards and Understanding of Learning by Doing: Evidence from an Automobile Plant. *Journal of Political Economy 121*(4), 643–681.

Lucas, R. (1993). Making a Miracle. *Econometrica 61*(2), 251–272.

Macchiavello, R., A. Menzel, and C. M. Woodruff (2015). Challenges of Change: An Experiment Training Women to Manage in the Bangladeshi Garment Sector. mimeo, CERGE-EI Prague.

McKinsey (2011). Bangladesh's Ready Made Garments Landscape: The Challenge of Growth. McKinsey and Company, Apparel, Fashion and Luxury Practice.

Osterman, P. (1994). How common is workplace transformation and who adopts it? *ILR Review 47*(2), 173–188.

Papay, J. P., E. S. Taylor, J. H. Tyler, and M. Laski (2020). Learning Job Skills from Colleagues at Work: Evidence from a Field Experiment using Teacher Performance Data. *American Economic Journal: Economic Policy 12*(1), 359–388.

Sandvik, J. J., R. E. Saouma, N. T. Seegert, and C. T. Stanton (2020). Workplace knowledge flows. *Quarterly Journal of Economics , forthcoming.*

Thompson, P. (2007). How Much Did the Liberty Shipbuilders Forget? *Management Science 53*(6), 908–918.

Thompson, P. and R. Thornton (2001). Learning from Experience and Learning from Others: An Exploration of Learning and Spill-overs in Wartime Shipbuilding. *American Economic Review 91*(5), 1350–1368.

Wright, T. (1936). Factors affecting the Cost of Airplanes. *Journal of Aeronautical Sciences 3*(4), 122–128.

Young, A. (2018). Channelling fisher: Randomization tests and the statistical insignificance of seemingly significant experimental results. *Quarterly Journal of Economics , forthcoming.*

Appendix A.1: Worker Movement

While sewing workers are allocated to fixed lines, they do at times switch lines on a day-to-day basis to replace absent workers. With average daily absenteeism rates in the sample factories between three to five percent, there is scope that within these reallocations, enough workers with relevant production knowledge on specific styles are moved across lines to drive the observed productivity spill-overs.

To test for this possibility, a tentative test is conducted. If workers with experience of the styles are reallocated across lines, the lines from which the workers are taken should experience a negative effect on their productivity if they still produce the style when further lines also start producing the style, and workers are reallocated away to these lines. Table A.1 shows the results when regressing the daily productivity of the first lines that produced a style in the factory on line chief - style day fixed effects, factory-week fixed effects, and on a dummy indicating that meanwhile, additional lines have also started producing the same style. The results show that instead of a drop in productivity, if anything, the first line experiences an increase in productivity when other lines also start producing the style. The source of these positive effects are not immediately clear; they could be due to reverse knowledge spill-overs from the additional lines back to the first line, or due to other forms of peer effects, such as competition. However, these results are not in line with what could be expected if systematic movement of workers with production experience of certain styles caused the observed productivity spill-overs.

Appendix A.2: Styles produced within vs across multiple floors

Table A.2 shows results from tests on whether styles produced on multiple versus a single floor differ on observable style characteristics. The available observables are the technical complexity of the style as measured by its SMV, the importance of the style's buyer to the firm (the share of all styles produced by the factory ordered by this buyer), the importance of the style's garment type to the firm (the share among all styles produced by the factory that are of the same type - t-shirts, polo-shirts, pants, etc - as this style), and the share among all styles the factory produces of the same type - buyer combination as this style. Styles produced on multiple floors do not differ significantly on these dimensions from those styles produced only on one floor.

Table A.1: Worker Movement

	Efficiency
Additional Lines	1.421**
	(0.714)
Observations	19,972
R^2	0.426
Controls	YES
Factory-week FE	YES
L.Chief-Style.Day FE	YES

Notes: Table shows results from lines producing styles that no other line in the factory had yet produced when they started the style, when regressing daily productivity of these first lines on an indicator variable that by the day of the observation other lines had also started to produce the same style ("Additional Lines"). Regressions control for factory-week fixed effects, as well as line chief fixed effects, SMV, runtime and worker on line, all interacted with factory - 'style day' fixed effects. Standard Errors clustered at the line chief level in parentheses: ** p<0.05

Table A.2: Styles produced within vs across Floors

	(1)	(2)	(3)	(4)
	Complexity	Buyer Share	Garment Type Share	Buyer-Garment Type-Buyer Share
Produced on multipl. floors	-0.178	0.006	-0.010	-0.002
	(0.289)	(0.004)	(0.014)	(0.003)
Observations	2,369	2,345	1,583	1,579
Factory & Year FE	YES	YES	YES	YES

Notes: Table shows the result of regressing, in the sample of all styles produced on more than one line, different style observables on a dummy indicating that the style was produced on more than one floor. All regressions control for factory and year fixed effects.

Table B.1: Productivity Spill-Over & Delivery Pressure

	(1)		(2)		(3)		(4)	
Specification of Indep. Variables:	Log Prev. Outp.		Log Prev. Outp.		Prev. Outp. >0		Prev. Outp. >0	
Cumul. Previous Output x ...								
Day 1	0.317***	(0.12)	0.037	(0.15)	1.372	(1.17)	0.064	(1.30)
Day 2	0.288***	(0.10)	0.244**	(0.12)	1.754**	(0.87)	1.671*	(1.00)
Day 3	0.171*	(0.09)	0.197*	(0.10)	0.826	(0.87)	1.111	(0.83)
Day 4	0.290***	(0.09)	0.197*	(0.12)	2.250**	(0.91)	1.374	(0.90)
Day 5	0.112	(0.08)	-0.088	(0.13)	-1.454	(0.75)	0.756	(1.05)
Day 6	0.130	(0.09)	0.065	(0.12)	1.013	(0.80)	0.605	(1.05)
Day 7	0.128	(0.10)	0.092	(0.15)	1.002	(0.90)	0.636	(1.06)
Day 8	0.251**	(0.11)	0.193	(0.14)	1.999*	(1.10)	1.451	(1.13)
Day 9	0.049	(0.11)	0.048	(0.15)	-0.225	(1.01)	-0.665	(1.19)
Day 10	0.159	(0.14)	0.257	(0.17)	0.966	(1.21)	1.650	(1.47)
Cumul. Previous Output Same Floor x ...								
Day 1	0.396***	(0.13)	0.349***	(0.13)	3.597***	(1.17)	2.790**	(1.13)
Day 2	0.248**	(0.10)	0.136	(0.10)	2.171**	(0.84)	1.296	(0.87)
Day 3	0.250**	(0.10)	0.180*	(0.10)	2.402***	(0.87)	1.945**	(0.81)
Day 4	0.151	(0.10)	0.137	(0.09)	1.117	1.226	(0.84)	(0.80)
Day 5	0.235***	(0.08)	0.221**	(0.11)	1.834**	(0.75)	2.038**	(0.94)
Day 6	0.217*	(0.11)	0.214*	(0.12)	1.602*	(0.95)	1.244	(1.01)
Day 7	0.177*	(0.10)	0.070	(0.12)	1.628*	(0.90)	0.542	(0.94)
Day 8	0.110	(0.11)	0.066	(0.13)	0.842	(1.01)	0.405	(1.07)
Day 9	0.011	(0.12)	0.083	(0.15)	-0.081	(1.03)	0.675	(1.29)
Day 10	0.093	(0.13)	-0.022	(0.16)	0.534	(1.11)	-0.620	(1.55)
N	27,661		27,661		27,700		27,700	
Controls	YES		YES		YES		YES	
Factory-Month FE	YES		YES		YES		YES	
Line Chief FE	YES		YES		YES		YES	
Style FE			YES				YES	

Notes: Table replicates Table 2 but controlling for log of remaining days to last observed day of production of same style, as a proxy for remaining time to delivery deadline. Observations from last month of data from factory excluded. Controls are SMV, daily runtime, and number of workers on line. All fixed effects and controls, including log remaining days of production, interacted with style day fixed effects. Due to computational constraints, regressions were estimated separately for each of the first ten days a line produces a new style. 'N' refers to summed 'N' for all ten of these regressions. Standard errors clustered at the line chief level in brackets: *** p<0.01, ** p<0.05, * p<0.1.

Appendix B.1: Productivity Spill-Over and Delivery Pressure

Table B.1 replicates Table 2, but controlling for a proxy for delivery pressure the line faces on the production day. As I do not observe the delivery date for an order, I proxy delivery date by the last day I observe a style being produced.[20] For each observation of daily output of a line, I calculate the difference of days to the last day of production of the same style, and take the log of one plus that number, as the marginal effect of one day less to delivery date on production pressure should be decreasing in the number of days left (controlling linearly for remaining days to delivery date yields the same results). I then interact this variable, as all other right hand side variables in these regressions, with style day - factory fixed effects.

[20]For the styles that lines produce on the last day of production data that is available from a factory, we do not know if that is the last day of production of the style, or if production continued. To remedy this concern in the construction of the proxy variable for delivery deadline, I exclude the last 30 days of production data from each factory from the estimations in Table B.1.

Appendix B.2: Non-Linear Model Estimation

By estimating independent effects of previous output for each of the first ten "style-days", the results from Table 2 provide a non-parametric description of how the effects evolve over the learning curve when a line starts producing a new style. However, due to the large number of estimated coefficients, gauging the aggregate effect of previous output over the learning curve is challenging. Furthermore, the treatment effects are not independent across style-days, making obtaining correct standard errors a concern. For these reasons, Table B.2 shows the results of an alternative estimation strategy, in which the parameters of a basic non-linear model of a learning curve are fitted to the data. While this approach is based on a number of modelling assumptions, it provides an intuitive and parsimonious approach to estimate the total reduction of the productivity loss during the learning curve due to productivity spill-overs across the production lines.

Specifically, the model that I fit assumes that once the workers on a line have fully solved all problems in the production process of a given style, they produce the style with a steady state productivity Ω. However, non-solved problems induce a negative gap between actual productivity Y and Ω. On average, if no other line has produced the style before, productivity Y is only a share $1 - w$ of Ω on the first day of production of a new garment. Subsequently, each day of production the gap "depreciates" by a factor θ, or a share θ of the previous day's gap remains on the next day. Finally, previous output on another line reduces the initial gap w by a factor κ_A, and an additional factor κ_F if the line was located on the same floor. Thus, productivity Y_{nift} on the n'th style-day a line chief i produces a garment is given by

$$Y_{nift} = \left(1 - [1 - \mathrm{I}(A_{ift} > 0)\kappa_A - \mathrm{I}(F_{ift} > 0)\kappa_F] * w_f \theta_f^{(n-1)}\right) \ \Omega_{ft} + \epsilon_{nift} \tag{3}$$

I add factory f and month t subscripts, as I allow steady state productivity Ω to vary at the factory-month level, to allow for factory specific time trends and seasonality in productivity. Similarly, I also allow for factory specific values for the start gap w_f, and the rates θ_f with which the gaps close over time. I(.) is an indicator function, while ϵ_{nift} an additive residual term.

Table B.2: Non-Linear Model Estimation

	(1)		(2)		(3)		(4)	
	Observational Data				**Randomized Intervention**			
w_1: Start Gap. Fact. 1	0.374***	(0.02)	0.352***	(0.02)	0.222***	(0.00)	0.220***	(0.000)
θ_1: closing factor, Fact. 1	0.609***	(0.02)	0.585***	(0.02)	0.452***	(0.00)	0.478***	(0.000)
w_2: Start Gap. Fact. 2	0.195***	(0.03)	0.190***	(0.02)	0.141***	(0.00)	0.113***	(0.000)
θ_2: closing factor, Fact. 2	0.682***	(0.06)	0.693***	(0.06)	0.627*	(0.05)	0.611	(0.100)
w_3: Start Gap. Fact. 3	0.532***	(0.02)	0.482***	(0.02)	0.392***	(0.00)	0.346***	(0.000)
θ_3: closing factor, Fact. 3	0.627***	(0.02)	0.546***	(0.02)	0.482***	(0.00)	0.488***	(0.000)
κ_A: Other Lines	0.185***	(0.05)	0.111**	(0.06)				
κ_F: Other Lines, Same Floor	0.270***	(0.05)	0.319***	(0.06)				
γ: Treatment					0.408**††	(.010/.025)	0.622***	(.005/.12)
γ x O: Treatm. x Other Floors							-0.404***†††	(.00/.00)
O: Other Floors							-0.239***	(0.000)
N	30,392		30,392		7,933		7,933	
Factory-Month FE	YES				YES		YES	
Line Chief FE			YES		YES		YES	
Treatm.Line & Treatm.Time					YES		YES	

Notes: Columns 1-2 show estimated parameter values of the basic non-linear model of learning curves and productivity spill-overs, as shown in equation 3, with column 1 allowing steady state productivity to vary on the factory-month level, and column 2 on the line chief level. Columns 3-4 show estimated parameters from non-linear model laid out in equation 4. Standard errors in columns 1-2 clustered at the line chief level in brackets. Brackets in columns 3 and 4 showing (two sided) p-values based on wild-cluster bootstrap (*) and randomisation tests (†), clustered on floor level. ***/††† p<0.01, **/†† p<0.05, */† p<0.1

Column 1 of Table B.2 shows the results from estimating the model, using again the productivity observations from the first 10 days a line produces a style. If no other line has produced the style before, productivity is 20 percent (Factory 2) to 53 percent (Factory 3) lower than the steady state productivity. The "discount factor" θ_f with which the gap closes with each day of production is between 61 and 68 percent for all three factories. Finally, the initial drop is reduced by 19 percent if some other line in the factory has produced the style before, and by another 27 percent if that line is located on the same floor. In an alternative specification shown in Column 2, I allow steady state productivity Ω to vary at the line chief level. Here the combined reduction of the initial drop when another line on the same floor has produced the garment is 43%. The equivalent linear specification shown in Table 2, Column 3, suggests a reduction of around six productivity units in that case, or a bit below 40% of the initial drop of around 17 units, as gleaned from Figure 1. Thus, the results from estimating the non-linear model are highly consistent with those from estimating spill-over strength using basic linear regressions, as done in the main text.

Non-Linear Model for Treatment Effects of Randomized Intervention

Columns 3-4 of Table B.2 show the results from adapting the estimation of the non-linear model to the estimation of the treatment effects of the randomized intervention. Specifically, the following equation is fitted to the data:

$$Y_{nift} = \left(1 - \left[1 - T_{ift}\gamma - T_t^D \gamma_f^D - T_{if}^L \gamma_f^L\right] \; w_f \theta_f^{(n-1)}\right) \; (\Omega_{ft} + \Omega_i) + \epsilon_{nift} \tag{4}$$

While w_f continues to capture the initial drop, or "start gap" of productivity on the first day a line starts producing a new style, and θ_f the depreciation factor of the start gap with each production day, T_{if} now captures the ITT treatment, the relative reduction of the start gap from lines being randomized into receiving briefings. I add equivalent indicator variables T_t^D that the style start occurred during the time the briefing was conducted on any floor that participated in the randomized trial, and T_{if}^L that the style start occurred on a treatment line either during or before the trial period, to obtain a complete difference in difference specification, with γ_f^D and γ_f^L allowed to vary at the factory level. Ω is allowed to vary on the factory-month level ft and on the line chief level i, mirroring the line chief and factory-month fixed effects from the linear specifications from Table 5.[21]

As shown in Column 3 of Table B.2 from the estimated parameters of the model, the start gap is now between 14 percent (Factory 2) and 39 percent (Factory 3) of steady state productivity. These values are a bit smaller than the equivalent values in Columns 1-2, which is intuitive as the sample used for this estimation is the same as in Table 5, excluding observations from lines starting styles that no line has produced before. We would expect the start gap to be larger in these instances; by dropping them from the sample, the estimated average start gaps should be smaller. The treatment reduces the remaining start gap by around 41 percent, roughly in line with the results from Table 5. Column 4 interacts the treatment indicator variable T with an indicator variable O for all lines that already produced the style being located on other floors. The results are consistent with those in Table 6; if the briefing could have been made by a line chief from the same floor, the remaining productivity drop is reduced by around 60%, while if only line chiefs on other floors have produced the style before, two thirds of this effect are offset again.

[21] Due to computational constraints, it was not possible to let Ω vary also at the style level.

Appendix C.1: Results using Recorded Treatments

The main analysis of the randomized communication intervention focuses on the Intention-to-Treat effect, assuming that any style start which was randomly selected for treatment was treated. This appendix presents results when regressing productivity on the first five days a line produces a new style on a dummy variable (interacted with 'style day' fixed effects) indicating that, according to the logbooks, the line chief received a briefing for the style by another line chief who had already produced it. As already mentioned in the main part of the paper, the production data records 393 starts of new styles on treatment lines which should have been treated, while 98 of these starts could be matched with recorded treatments in the log-books. However, given that the factory managements admitted that not all treatments were recorded in the logbooks, the number of actual treatments is likely to be larger. Note, however, that there are no recorded treatments on control lines, or on dates before the time the intervention was implemented at the factories. Thus, when dropping style starts from the sample in which no other line has yet produced the style (as was done for the same sample used in Table 5 and in Figures 4- 5), the only type of non-compliers left in the sample are style starts which should have been treated but were not.

Table C.1 shows the results when replicating Table 5, but with treatment now indicating recorded treatments instead of ITT treatment. Using recorded treatments, we only find statistically insignificant and very small effects. The point estimates are even negative (though statistically completely insignificant) when not including style fixed effects (Column 1). When using style fixed effects, the point estimates for the effect becomes positive, but remains very small and completely statistically insignificant.

The lack of effects of recorded treatments are puzzling in light of the significant intention-to-treat effects. The following points could provide explanations for this pattern. First, as indicated by the factory managements, more style starts were likely treated than indicated as "treated" in the data. This would dilute the estimated effect of the treatment, particularly if the

Table C.1: General Treatment Effects, Recorded Treatments

	(1)		(2)		(3) Reweighted	
Recorded Treatment x						
Day 1	-2.035	(0.32)	0.500	(0.85)	0.462	(.83)
Day 2	-0.941	(0.51)	-1.730	(0.55)	-2.393	(.45)
Day 3	2.437	(0.31)	0.917	(0.48)	0.126	(.94)
Day 4	0.634	(0.71)	-0.969	(0.57)	-0.842	(.69)
Day 5	-1.076	(0.82)	-2.832**	(0.04)	-2.232**	(.04)
N	4,920		4,920		4,594	
Controls	YES		YES		YES	
Line Chief FE	YES		YES		YES	
Factory-Month FE	YES		YES		YES	
Style FE			YES		YES	

Notes: Table repeats Table 5, but instead of the ITT specification, it uses actual recorded treatments. Columns 2 and 3 add style fixed effects interacted with style day fixed effects, while Column 3 uses the reweighting technique based on DiNardo et al. (1996) to control for differential pre-treatment productivity of treatment and control lines. Controls are SMV, runtime and number of workers on a line. All controls and fixed effects interacted with factory and style day fixed effects. Wildcluster bootstrap based p-values clustered at 17 randomization units shown in brackets: *** p<0.01, ** p<0.05, * p<0.1.

likelihood of the treatment not being recorded was uncorrelated (or even negatively correlated) with the effect of the particular treatment on productivity. Second, actually administered treatments were probably directed towards style starts at which the factory management expected poor initial productivity. Therefore, the estimated effect of the recorded treatments is small, as the right counterfactual is not the productivity at other untreated style starts on the same line or on control lines, but the (unobserved) productivity if these treatments would not have been conducted. This also fits with the effects becoming slightly larger when using style fixed effects as compared to when not, as well as with the reduction in the left tail of the productivity distribution on treated lines with the onset of the treatment as shown in Figure 6.

Appendix C.2: Long Term Productivity Trends, Factory 1 and 2

At Factory 1 and 2, we have more than one year of continuous production data before the start of the intervention. We can therefore investigate for a longer time period whether productivity on randomly selected treatment and control floors followed parallel trends, which would support the identifying assumption of the difference-in-difference estimation of the treatment effect of

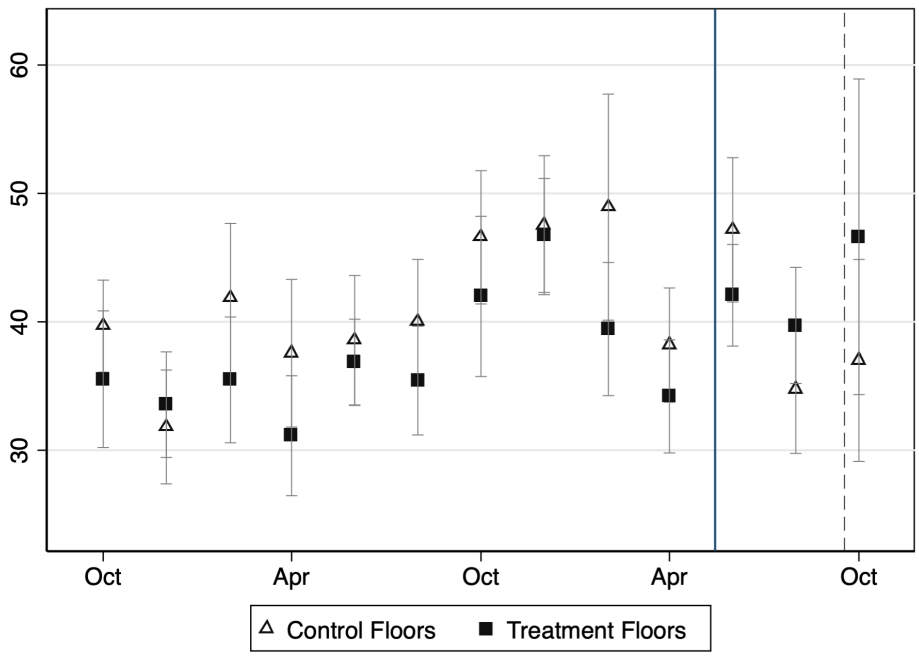

Figure C.1: Pre/Post Intervention Start Trends for First Day Productivity, Factory 1+2.

the intervention. Figure C.1 shows aggregate production data over two-month intervals from the first days lines produced styles that had already been produced by other line chiefs. I use two-month intervals instead of one-month intervals to reduce the number of intervals shown, due to the longer time period included. Figure C.1 shows the graph when only controlling for SMV, runtime and number of workers on line (interacted with factory FE). Figure C.2 additionally controls for style fixed effects. The figures indicate that also over the longer time period considered here, first day line productivity largely follows parallel trends on treatment and control floors, particularly when controlling for style fixed effects and looking at the year before the start of the intervention.

Appendix C.3: Application of Reweighting Approach from DiNardo et al. (1996)

The implementation of the approach of DiNardo et al. (1996) requires the estimation of two probit models: first, of a dummy indicating whether a unit i in the sample is selected for treatment ($T_i = 1$) on the unbalanced variable z_i, and, second, of a dummy indicating whether the unit is

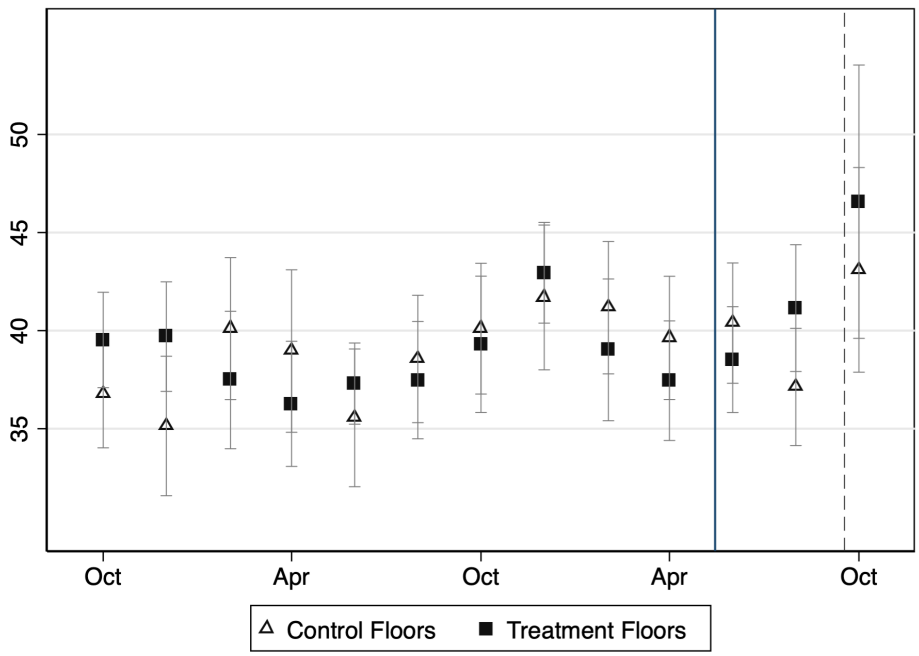

Figure C.2: Pre/Post Intervention Start Trends for First Day Productivity, Factory 1+2. Figures C.1 and C.2 show average productivity within two-month intervals of lines on the first day they start producing a new style that other line chiefs have previously produced ('non-first style starts'), separately for lines from treatment floors (solid squares) and control floors (hollow triangles). The solid vertical line indicates the start of the intervention from June 2014 onwards, while the dashed line shows the end of the four months treatment period on treatment floors. Capped bars represent 95% confidence intervals.

selected as control ($T_i = 0$) on z_i. The predicted probabilities for each unit i, $P(T = 1|z_i)$ and $P(T = 0|z_i)$, and the unconditional probabilities $P(T = 1)$ and $P(T = 0)$ of being selected into as treatment or control unit, respectively, are then used to calculate weights w_i for each unit i according to:

$$w_i = \frac{P(T = 0|z_i)P(T = 1)}{P(T = 1|z_i)P(T = 0)} \tag{5}$$

To implement the approach, I first regress, on a sample of all sewing lines, a dummy indicating that a sewing line is located on a treatment floor on the line's average productivity on the first days they produced new styles that have previously been produced on other lines, during the pre-intervention time April and May 2014. I control for factory fixed effects. The predicted values of this regression for each sewing line yield $P(T = 1|z_i)$. Similarly, I also regress a dummy indicating that a line is located on a control floor on its average first-day productivity during April and May 2014, to obtain $P(T = 0|z_i)$. I then use these weights for each production line located on treatment floors in Column 3 of Tables 5 and 6 to re-run the specification from Column 2 in the same table, only with the data from treatment floors reweighted by these weights (control units are not reweighted in this approach, therefore weights w_i for lines from control floors are set to 1).

Appendix C.4: Cross Floor Treatment Effects, by Factories

Table C.2 below replicates Column 1 of Table 6, but with the sample split between Factory 1 and 2 (Column 1), and Factory 3 (Columns 2). While the interaction with the variable indicating that all previous lines that already produced the style were on other floors shows no pattern for Factories 1 and 2, it is strongly negative for Factory 3. This fits with the fact that at Factory 3, briefings in which the line chief would have had to come from another floor were less likely reported in the log-books than those where a line from the same floor was available. Note, however, that the standard errors in the regressions from Table C.2 are clustered on the level of only 11 floors (Factory 1 and 2, Columns 1 and 3) and 6 Floors (Factory 3, Columns 2 and 4), which can cause the standard errors to be downward biased (Cameron et al. (2008)).

Table C.2: Treatment Effects with Other Floors, by Factories

	(1) Factory 1 + 2		(2) Factory 3	
Treatment x				
Day 1	5.888*	(0.07)	9.556**	(0.03)
Day 2	6.662	(0.18)	7.753***	(0.00)
Day 3	1.737	(0.57)	10.511**	(0.03)
Day 4	-1.757	(0.48)	7.221*	(0.06)
Day 5	-3.071	(0.13)	5.104	(0.53)
Treatment x Other Floors x				
Day 1	-0.221	(0.92)	-7.121***	(0.00)
Day 2	-2.400	(0.68)	-8.431***	(0.00)
Day 3	-4.248	(0.29)	-10.945**	(0.03)
Day 4	-2.008	(0.53)	-6.804***	(0.00)
Day 5	1.458	(0.79)	-2.061	(0.53)
Other Floors x				
Day 1	-2.445	(0.29)	3.453*	(0.09)
Day 2	-1.549	(0.50)	2.134	(0.56)
Day 3	-2.681	(0.28)	4.363**	(0.03)
Day 4	-0.456	(0.84)	3.016*	(0.09)
Day 5	-0.174	(0.96)	0.907	(0.69)
N	2,224		2,722	
Controls	YES		YES	
Factory-Month FE	YES		YES	
Line Chief FE	YES		YES	

Notes: Tables replicates Column 1 of Table 6, but separately on the samples from Factory 1 and 2 (Column 1), and Factory 3 (Column 2). Controls are SMV, runtime and number of workers on a line. All controls and fixed effects interacted with factory and 'style day' dummies. Wildcluster bootstrap based p-values clustered at floor level shown in brackets: *** p<0.01, ** p<0.05, * p<0.1.

The focus of the Table should therefore be on the point estimates of the interaction effects, which are negative and large for Factory 3, but not for Factory 1 and 2.