Recover Overnight?
Work Interruption and Worker Productivity

This Version: March 2016

Abstract

This paper investigates the effect of work interruption on workers' subsequent productivity. We employ a data set of individual productivity and machine conditions, in which each worker faces the chance, on a daily basis, that her machine will break down randomly. Our analysis finds that compared to a workday with smooth production, experiencing a machine breakdown is associated with a decline in individual productivity the following day. We discuss possible explanations, including proficiency loss and negative emotion induced by work interruption, for the observed effect. Our findings shed light on the importance of understanding and managing interruptions in the workplace, and contribute to a growing literature on the determinants of productivity at the micro level.

Keywords: Productivity, Labor supply, Work interruption
JEL Classification: D84, J22, M11

1 Introduction

Interruptions are common in workplaces and costly to individuals, firms, and organizations. From equipment breakdowns to unscheduled meetings and communication requests, unplanned breaks from the ideal smooth and continuous production process lead to losses in work hours and reductions in worker productivity. For instance, according to Spira and Feintuch (2005), the direct costs of unnecessary interruptions have been estimated at an average of 28% of daily time for knowledge workers in the United States. Nowadays, with firms adopting modern organization and communication technologies—open-plan offices, e-mails, instant messages, etc.—managing interruptions has become increasingly important and challenging for businesses and their workers.

In addition to impairing direct productivity, work interruption may have spillover effects on subsequent production. On the one hand, it has been widely recognized that interruptions could have negative spillovers: Workers may need extra time to warm up and resume full engagement and concentration; productivity may decline due to negative emotions, such as stress from time pressure and frustration about failing to meet targets (Mandler 1990).[1] On the other hand, interruptions may be beneficial for subsequent productivity, as occasional breaks could help workers alleviate fatigue and boredom (Roy 1959). Moreover, workers might try to make up for the losses caused by interruptions and exert greater effort leading to increased productivity (Camerer, Babcock, Loewenstein and Thaler 1997). In the field of management and organization sciences, both positive and negative effects of interruptions have been discussed extensively (see Jett and George 2003 for a review). Nevertheless, scant empirical work has estimated the consequences of work interruption, partly because of the difficulty of identifying each incident and establishing a causal relationship.

This paper is the first to estimate the effect of interruption on workers' subsequent productivity using exogenous incidents of work interruption. We examine the consequences of machine breakdown in a plastics-printing company in China. Manufacturing companies face the possibility of machine breakdown and concern about how much resources to expense on machine maintenance. For manufacturers in developing countries that transit from labor intensive to technology intensive structures, the cost of inadequate maintenance is particularly important yet not well understood.

Specifically, we collect a data set of worker-level daily output and machine conditions. Our sample spans a period of 473 days, during which a total of 273 workers labored on 75 machines for 25 different products. As the workers are paid by piece rate, we have an

[1] In the workplace, it has been observed that negative shocks, such as failed salary negotiations (Mas 2006), bonus payments falling short of individually assigned bonus targets (Ockenfels, Sliwka and Werner 2014), and bereavement and family illness (Oswald, Proto and Sgroi 2014), affect workers' subsequent productivity.

accurate record of workers' daily output levels. On a daily basis, team managers assign each worker to a machine and a product type, and each worker faces the chance that the machine will break down. On average, machines have a breakdown frequency of 14%, and each breakdown takes about 4 hours to repair. [2]

Our primary empirical strategy compares two workers operating the same machine during two different shifts on the same day. The idea is that the worker who operates the machine smoothly constitutes a good control for the worker who operates the same machine on the same day, but experiences the shock of machine breakdown. More specifically, we estimate how machine breakdown affects a worker's productivity the following day, and find that it leads to a 2% decline in the worker's daily output. This result is consistent across alternative estimation methods and several robustness checks.

Our findings demonstrate a robust, negative spillover effect from work interruption. We provide further discussions about possible sources of the observed negative effect. First, it is possible that workers become less proficient and less engaged after the interruption; for example, they need to start over with the techniques and procedures. Second, the observed effect may come from a negative response to or emotion about the negative shock. We conduct further empirical analyses and show evidence consistent with the proficiency hypothesis. We further explore worker heterogeneity, and find that the effect is larger among females than males, and among older than younger workers. In addition, we find that when more peers are experiencing interruptions, the negative effect on a worker's subsequent productivity is larger, lending support to the literature on peer effects in the workplace.

Our study contributes to a growing literature on the economic consequences of work interruption. Researchers have investigated various forms of interruptions in workplaces, including worker absence, menstruation, weather, pollution, and multitasking. Herrmann and Rockoff (2012b) use unique data on worker absence to estimate the effect of work absences on productivity. They find that expected loss in daily productivity from employing a temporary substitute is on par with replacing a regular worker of average productivity with a worker of $10-20\%$ worse productivity. Ichino and Moretti (2009) observe that menstruation as a form of interruption contributes to gender gaps in absenteeism and earnings, while Herrmann and Rockoff (2012a, 2013) find little support for the role of menstruation in explaining the gender gap in earnings. Connolly (2008) explores the effect of exogenous variation in daily weather on labor supply, and finds that men shift on average 30 minutes from leisure to work on rainy days, giving rise to a rough estimate of the intertemporal elasticity of labor supply at around 0.01. Coviello et al. (2014) propose a model of task

[2] Our data contain the information on the duration of machine breakdown, but not its exact time, which precludes us from examining the effect of breakdown on productivity in the same day.

juggling, in which a worker is interrupted and switches from one project to another too frequently. Subsequently, Coviello et al. (2015) estimate the causal effect of an exogenously induced increase in parallel working of some judges in Italy and show that juggling causally and substantially lowers their productivity.

Our findings are also related to income targeting and labor supply, which has been extensively investigated in the recent literature. In a seminal study by Camerer et al. (1997), the authors find that cabdrivers work more on the days when the transient salary is low, which suggests that they might have a daily income target and quit working once they reach that target. Several follow-up studies use field evidence and observational data to investigate the daily income-targeting hypothesis. The evidence appears to be mixed. Several studies find positive evidence for the daily income-targeting hypothesis (Fehr and Goette 2007; Crawford and Meng 2011), while some suggest otherwise (Farber 2005, 2008; Andersen et al. 2014). As our data do not contain same-day productivity before and after interruption (see data section for details), we cannot directly test daily income targeting. Alternatively, our setting could be interpreted as a test of cross-day income-targeting behavior, in which workers would work harder on the second day to make up for losses on the first day. However, our results do not support this prediction, suggesting that income targeting behavior is more likely to be "one day at a time."

The rest of the paper proceeds as follows. In Section 2 we describe our research setting, including the institutional background and a stylized model. In Section 3 we discuss the data, main variables, and identification strategy. Section 4 presents our empirical estimates and robustness checks. Section 5 concludes.

2 Context

2.1 Workplace

We collect a data set of worker-level daily output from a leading company that produces double-wall paper cups in Fujian Province of China. Established in 2003, the company has become a major supplier of paper cups for the food and beverage industry. In 2011, the company accounted for around 15% of the national market. By 2013, it had $7 million in total assets, $5 million in annual revenue, and about 300 employees.

Our data focus on production workers in the molding and wrapping divisions. Their main task is to operate the molding machines to seal the cup body and bottom, then use wrapping machines to bond sleeves for the paper cups. Workers are hired from both within and outside Fujian Province. Each day, workers are divided into day and night shifts, typically working for 12 hours. Specifically, the day shift is from 6am to 6pm, while the night shift is from 6pm to 6am the next day. In our sample, 7,331 employee-shift observations were during the day shift, and 7,633 during the night shift. Within a shift, each worker is assigned a machine and a product type. A machine can generally be configured for multiple types of products, and production efficiency may differ across products.

Notably, workers themselves do not decide on their product or machine. Each day, a general production manager determines the composition of the production teams, which consist of a team manager and four or five workers. The composition varies each day, and is largely unknown to the team managers and workers until they arrive at the site. Team managers assign their member workers to machines and product types. Once assigned, workers cannot change the machine or the product. The only choice variable a worker has is how much effort to exert in running the machine and preparing materials for the next stage of production. Moreover, workers rotate tasks and machines, reflecting (1) the firm's aim to ensure fairness, as machine efficiency varies across products, and (2) the fact that there is little room for human capital accumulation in the low-skilled jobs.

Once production begins, team managers prepare packaging boxes labeled with workers' names and products, and provide assistance to smooth operations if necessary. A worker's output depends on the normal operation and functioning of the machine. If a machine breaks down, the worker who operates it calls maintenance staff for repairs and the maintenance staff records the duration of the breakdown. The worker can neither change to a different machine nor leave the site, which means that the worker's production is completely interrupted until the machine is fixed.[3]

Team managers are paid a flat rate, while workers are paid a two-tier piece rate based on their daily output and the specified target level. Specifically, on a typical day and for each product-machine pair, the general production manager specifies a target, denoted by $\bar{Q}_{pm}$, which depends on the features of product p and machine m. When they meet the target, workers are paid a piece rate denoted by α_1 per unit of output. When output exceeds the target, a higher rate, denoted by α_2 (where $\alpha_2 > \alpha_1$), applies to each unit above the target. When a worker fails to meet the benchmark, the unfulfilled units are also deducted by the rate α_2. Formally, given the worker's daily production Q_{ipmd}, the salary is determined as $\alpha_1\bar{Q}_{pm} + \alpha_2(Q_{ipmd} - \bar{Q}_{pm})$.[4] The target will be proportionally adjusted if machine breakdown occurs; i.e., given a breakdown of x hours, the worker's new target becomes $\tilde{Q}_{pm} = \bar{Q}_{pm}\frac{12-x}{12}$.

[3]Whether the machine is fixed or not by the end of the workday, the next day the worker always rotates machines and products, as other workers do.

[4]In principle, the daily salary could be negative if output level is very low, i.e., $Q_{ipmd} < \frac{\alpha_2-\alpha_1}{\alpha_2}\bar{Q}_{pm}$. However, since there is no such case in our data, presumably the specified target output level $\bar{Q}_{pm}$ is at a reasonable level and workers exert sufficient effort.

2.2 A Stylized Model

We construct a parsimonious and stylized model to illustrate our research setting. Consider that the company has N workers indexed by i. The production function is denoted as $Q_{id} = Q(e_{id}, x_{id})$, where e_{id} is the effort chosen by worker i on day d; and x_{id} are exogenous variables that affect daily output, such as worker proficiency, product features, machine conditions, and workplace environment (e.g., weather and pollution). $Q(.)$ has the features of a standard production function: Output is an increasing and concave function of worker effort, and there is complementarity between worker effort and external productive factors—i.e., $Q_e > 0$, $Q_{ee} < 0$, $Q_x > 0$, and $Q_{ex} > 0$. A worker i's cost function is $C_{id} = C(e_{id}, w_{id})$, where w_{id} are external factors such as the mental status of the worker, and the regular assumptions for a cost function apply, namely, $C_e > 0$, $C_{ee} > 0$, $C_w < 0$, and $C_{ew} < 0$.

The optimal effort chosen by worker i on day d is given by

$$\max_{e_{id}} Q(e_{id}, x_{id}) - C(e_{id}, w_{id})$$

$$\Rightarrow e^*_{id} = e(x_{id}, w_{id}).$$

Linking the optimal output Q^*_{id} to the external variables (x_{id} and w_{id}), we have

$$\frac{\partial Q^*_{id}}{\partial x_{id}} = Q_e \frac{\partial e^*_{id}}{\partial x_{id}} + Q_x = Q_e \frac{Q_{ex}}{C_{ee} - Q_{ee}} + Q_x > 0, \tag{1}$$

and

$$\frac{\partial Q^*_{id}}{\partial w_{id}} = Q_e \frac{\partial e^*_{id}}{\partial w_{id}} = Q_e \frac{C_{ew}}{Q_{ee} - C_{ee}} > 0. \tag{2}$$

Next, we use equations (1) and (2) to illustrate how machine breakdown on day $d-1$ affect Q_{id} and the potential bias from omitted variables. We also extend the framework to incorporate heterogeneous effects across workers.

Effect of Machine Breakdown on Day $d-1$. Machine breakdown on day $d-1$ can affect the output Q^*_{id} on day d through production factor x_{id}, or cost factor w_{id}, or both. Consider first the possibility that breakdown affects x_{id}. For example, after a substantially long interruption, workers may become less proficient and less engaged in the task, and need to start over with the techniques and procedures (Jett and George, 2003). In our context, it is possible that workers take time to configure and tune their machines, prepare relevant materials and work logs, and adjust the pace according to their own and their peers' productivity. The process of reaching maximum proficiency imposes a fixed cost on productivity; once it is interrupted, workers have to start over again and incur such a cost.

Formally, let B_{id-1} denote that worker i experienced a machine breakdown on day $d-1$. The proficiency effect means that $\frac{\partial x_{id}}{\partial B_{id-1}} < 0$, which causes $\frac{\partial Q^*_{id}}{\partial B_{id-1}} < 0$.

The other possibility is that machine breakdown affects the external factor of workers' cost of effort w_{id}. For example, workers may feel frustrated by machine breakdown, because it is unpleasant, it goes against their expectations about the production process, or they are disappointed by the fact that their daily salary will be compromised due to the lost hours. The negative emotions and reactions make working hard more costly, or $\frac{\partial w_{id}}{\partial B_{id-1}} < 0$, which causes $\frac{\partial Q^*_{id}}{\partial B_{id-1}} < 0$.

Omitted Variables Bias. The daily rotation of workers among machines and products suggests that machine breakdown for a given worker i on a particular day d is arguably random. However, there could be other factors associated with, or even causing, machine breakdown on day $d-1$. If such factors persist to affect x_{id}, e.g.,the efficiency of production in day d, which in turn influences the output level Q^*_{id}, then the estimated effects of machine breakdown on output would be biased. We discuss a primary threat arising from team managers' manipulation of production assignments in the next section, and conduct several validity checks in Section 4.1.

Heterogeneous Effects. In the baseline setting, the production function $Q(\cdot)$ and cost function $C(\cdot)$ are the same for all workers on the site. This implies that our estimate is the average effect across all workers. However, worker heterogeneity is commonly observed in the workplace. We can extend the framework and allow for heterogeneity in both the output and cost function—i.e., $Q_g(.)$ and $C_g(.)$, where g indexes the worker type. Doing so allows us to estimate potentially heterogeneous effects from machine breakdown. In further analyses, we consider differential effects across several demographic characteristics—gender, age, and local versus migrant workers

3 Data, Variables, and Estimation Strategy

We use data on worker output and machine conditions for each workday from October 2012 to April 2014. The sample covers 473 workdays, 273 workers, 75 machines, and 25 product categories.

The effect of machine breakdown on a worker's subsequent productivity comes from the comparison of output at day d between workers with and without machine breakdown at day $d-1$. An unbiased estimation requires that machine breakdown randomly occurs to some workers but not others at day $d-1$. There are reasons to challenge this assumption.

Specifically, if machine breakdown on day $d-1$ were manipulated and the factors of manipulation continued to influence production on day d, the estimates would be biased. A primary concern arises from possible manipulation by team managers: A team manager, for instance, could assign his favorite worker—one he is socially connected with or who stands out in performance—to a machine that is in good condition. [5]

To overcome this identification problem, we implement a single difference design in which we compare a pair of workers with similar job assignments except for the experience of machine breakdown. Specifically, there are two job shifts every day and during each shift one worker operates the machine. We therefore focus on the two workers who operates the same machine during different shifts of day $d-1$, in which the machine breaks down during one shift (treatment group) but not the other (control group). If there were any manipulation by team managers in worker-machine assignment, the two workers assigned to the same machine on the same day, although during two different shifts, should be reasonably similar.

To verify the identifying assumption, we conduct a balancing test of the two workers' predetermined characteristics. Specifically, we compare their productivity, work participation, working hours; and whether the experienced machine breakdown on day $d-2$; and whether they produced the same products on two consecutive days (day $d-2$ and day $d-1$). In a reduced sample of workers with demographic information, we examine whether gender, age, residential status (i.e., local versus migrants), and job shift are also balanced between the two groups of workers. As shown in Table 1, except for worker age, none of the mean comparisons show any statistical or economic significance. For worker age, despite its statistical significance, the mean difference is only 0.706—very small compared with the group means (31.154 for the group experiencing machine breakdown and 31.861 for the group without machine breakdown). Test results suggest that our intended treatment and control groups are quite balanced.

[Insert Table 1 here]

We estimate the effect of machine breakdown on a worker's subsequent productivity based on the following single difference framework:

$$y_i = \beta Treatment_i + \varepsilon_i. \tag{3}$$

[5]Another possibility is manipulation by workers. However, this may not be a large concern in our research setting, for several reasons. First, workers are paid by their output (piece rate), and therefore have no incentive to sabotage the machine intentionally. Moreover, the company has implicit policies about deliberate machine sabotage, which is costly in repair expenditure and output losses. Team managers patrol machines to ensure that they are being operated properly.

where y_i denotes the productivity of worker i. To account for variations in machine efficiency across product types, we define worker productivity y_i as over-target percentage of output; specifically, $y_i \equiv \left(Q_i - \bar{Q}\right)/Q_i$, where $\bar{Q}$ is the targeted output. As workers are paid by daily output, both the firm and the workers make an effort to record output precisely, so measurement errors are not of huge concern. $Treatment_i$ is a dummy variable that indicates whether worker i's machine broke down one day before. This is the coefficient of our interest. Standard errors are clustered at the machine level to adjust for any heteroskedasticity and serial correlation.

To further strengthen our identification, we use two alternative strategies—one that controls for the nonrandom components of machine breakdown and the other a DD analysis—and conduct two placebo tests. We also check sample selection bias. Details are provided in Section 4.1.

4 Empirical Findings

Table 2, column 1, reports the regression results of baseline specification (3). The coefficient of $Treatment$ is negative and statistically significant, indicating that compared to his counterpart, who operated the same machine on the same day, the worker who experienced a machine breakdown had lower output in the following day.

The estimate suggests that the adverse effect is about a 2 percentage point drop in the over-target percentage of output. Evaluated at the mean, or 9.8 percentage points over-target of output for the control group (workers without machine breakdown on the previous day), machine breakdown reduces over-target output by $2/9.8 = 20.41\%$, or gross daily output by $2/92.2 = 2.17\%$.[6]

[Insert Table 2 here]

In the next subsections, we conduct several validity checks and examine possible interpretations.

4.1 Validity Checks

In this subsection, adopting the methods detailed in Section 3, we report the results from several validity checks.

[6] We define over-target percentage of output as $y_i \equiv \left(Q_i - \bar{Q}\right)/Q_i$. The average gross output of the control group is then $Q_i^c = \bar{Q}/(1-0.098)$ while the average output of the treatment group is $Q_i^t = \bar{Q}/(1-0.078)$. The treatment effect is equivalent to $(Q_i^t - Q_i^c)/Q_i^c = -2.17\%$ of the gross output.

Controlling for nonrandom components of machine breakdown. First, we make use of the fact that machine breakdown is probabilistic *ex ante*, which provides the random components to support our identification. On a given workday, until workers are assigned to machines, neither the team manager nor the worker himself knows with certainty whether a machine is going to break down, even with the understanding that some machines are in worse condition than others. Figure 1 shows the frequency of breakdown for each of the 75 machines in the regression sample. There are quite a few variations across machines, with some having high breakdown frequency (above 60%) and others having low frequency (around 10%). We consider whether machine m would break down on day $d-1$, B_{md-1}, as a function of two components: $B_{md-1} = g(c_{md-1}) + \upsilon_{md-1}$. Here, c_{md-1} is a systematic, non-random component related to machine condition and is observable to team managers. The other component, υ_{md-1}, is an exogenous, random element of chance—e.g., whether the workshop has the "right" temperature and humidity that day, or the raw material has defects that could damage the machine, etc. As long as $\upsilon_{md-1} \neq 0$, machine breakdown is arguably random even in the presence of endogenous assignment (i.e., $c_{md-1} \neq 0$). This is similar to the without-full-manipulation argument that underpins regression discontinuity design, as elaborated by Lee (2008). Following this literature (Lee and Lemieux 2010), we control for $g(c)$ in the regression to isolate the random components of machine breakdown. Specifically, we proxy machine condition c using the frequency of machine breakdown in the past five days and proxy $g(.)$ using a third-order polynomial function. As reported in Table 2, column 2, we find the estimate similar to the baseline in column 1, in both statistical significance and economic magnitude. These results corroborate the evidence of the balancing test reported in Table 1: For the pair of workers operating the same machine during different shifts on the same day, the shock of machine breakdown is plausibly exogenous.

Difference-in-differences estimation. To further address the concern that workers in the treatment and control groups might differ systematically in some unobservable characteristics, we use their productivity two days previous as a further control and conduct a DD analysis. Specifically, the if machine breaks down on day $d-1$, we now define the first-difference in outcome $\triangle y_i = y_{id} - y_{i(d-2)}$, i.e., the difference between day d's productivity and day $d-2$'s. Such within-worker difference can eliminate all differences across workers—including the differences between the treatment and control groups—that do not change across two days. This could include, for example, worker ability, experience, social connections with the managers or within the firm, etc. The DD estimates are reported in Table 2, column 3. Consistent with results from the baseline specification, we continue to find a negative effect of machine breakdown on workers' subsequent productivity. The magnitude is larger, but statistically insignificant.

An identifying assumption for the DD analysis is that absent machine breakdown, the treatment and control groups would have the same productivity trends throughout the two workdays. We conduct a validity check of parallel pretreatment trends, as commonly used in DD estimations. Specifically, we compare treatment and control groups' productivity change from day $d-3$ (two days before the day of machine breakdown) to day $d-2$ (one day before the day of machine breakdown). As reported in Appendix Table A1, column 1, the coefficient of $Treatment$ is highly insignificant and small in the magnitude, suggesting that workers in the treatment and those in the control group have parallel pretreatment productivity trends.

Placebo test I: Worker pairs without machine breakdown. Our baseline estimation uses pairs of workers operating the same machine on the same day, but one experiences machine breakdown during his shift (treatment), while the other does not (control). Here, we conduct a placebo test using pairs of workers operating the same machine on the same day, and the machine breaks down in neither shift. Given that the assignment rule is the same for all workers and machine breakdown is probabilistic, such pairs of workers should not show substantial differences in productivity on the subsequent day. In other words, if our treatment and control groups differ only in the breakdown incident, then the pairs of workers—neither of whom experienced a breakdown incident—should be the same. Any significant differences in productivity would imply a misspecification of our estimation. As Table 2, column 4 reports, the coefficient of $Treatment$ is close to zero and not statistically significant, which lends further support to the validity of our estimation strategy.

Placebo test II: Random assignment of breakdown incidents. As another placebo test, we randomly assign machine breakdown to workers. Specifically, during our sample period, there are 1,783 machine-day observations in which a machine breaks down. We first randomly select 1,783 out of a total of 9,578 machine-day observations, and then randomly assign an incident of machine breakdown to one of the two shifts for each selected machine-day observation. As such, we construct a *false* treatment for the pairs of workers on the same machine on the same day, i.e., $B_i^{false} = 1$ for one worker and $B_i^{false} = 0$ for the other. We then re-estimate equation (3) using this placebo sample. The randomization process implies that B_i^{false} should have no effect on y_i; in other words, we should have $\hat{\beta}^{false} = 0$. We conduct this random data-generating process 500 times to avoid contamination by any rare event and increase the power of identification.

Figure 2 shows the distribution of estimates from the 500 randomizations and plot our benchmark estimate, -0.020, from column 1 of Table 2. We find that the distribution of estimates from random assignments is centered around zero (the mean value is 0.00005), and the standard deviation of the estimates is 0.00662, which suggests that $\hat{\beta}^{false} = 0$.

In addition, our benchmark estimate clearly lies outside the estimates from the placebo tests. Altogether, these results imply that there are no substantial omitted variables in our specification.

[Insert Figure 1 here]

4.2 Interpretation

The baseline estimates show that machine breakdown has a negative and significant effect on workers' productivity the following day. Here, we discuss possible interpretations of this effect. In particular, we test the possibility that machine breakdown affects work participation the following day, in such a way that more productive workers are too frustrated by the interruption to attend work the following day—i.e., the negative effect is the result of sample selection. After addressing this concern, we test two possible channels of negative spillover: one from the standpoint of production proficiency and the other by examining workers' emotional reactions to work interruption.

Sample selection due to worker absence. Our baseline identification uses pairs of workers who both showed up the day after a machine breakdown. However, work participation might be endogenous: For instance, if workers—especially those with high productivity—are frustrated by the experience of machine breakdown, they may be discouraged and decide to be absent from work the next day. In that case, our estimates simply reflect the sample selection due to worker absence, instead of changes in worker productivity or behavior. To check this possibility, we examine whether machine breakdown affects work participation differently for our treatment and control groups. Estimates are reported in Table 3, column 1. We find that incidents of machine breakdown on average raise worker absence of the following day by 1.9 percentage points, which implies the possibility of sample selection bias.

[Insert Table 3 here]

To address potential bias due to sample selection, we use the methodology developed by Lee (2009), the premise of which is to consider the best and worst scenarios caused by sample selection and bound the estimated effect. Specifically, in our context, consider the case in which the 1.9% higher absence in the treatment group arises from workers with the highest productivity, i.e., the worst scenario and largest bias. Then, to construct a balanced sample, we also drop the top 1.9% productive workers in the control group, and the resulting estimates constitute the lower bound of the true effect. Similarly, if we assume that in the treatment group, workers who are absent the following day are the least productive ones, then we exclude their counterparts—the bottom 1.9% of productive workers—in the control group. The estimate from the refined and balanced sample represents the upper bound of the true effect. Table A1, columns 2 and 3 report the lower and upper bounds of the effects, respectively. We find that both are negative, although the estimated lower bound is marginally insignificant. Taken together, these results suggest that the negative spillover from machine breakdown is not primarily driven by sample selection.

Proficiency. One channel through which machine breakdown may negatively affect workers' subsequent production is the proficiency loss during the interruption. It is possible that after a substantial period of pausing and waiting—especially when the interruption is unscheduled—workers become less proficient, and they need extra time to gain momentum with techniques and procedures before returning to full engagement and concentration.

In a stylized framework, we can consider production proficiency x_{id} as a function of continuous working time—for instance, cumulative working hours until day d. We assume that $x_{id} = f(H_{id-1})$, where $f' > 0$, $H_{id-1} \equiv \sum_{j=1}^{d-1} h_{ij} = h_{id-1} + H_{id-2}$, and h_{ij} is the number of hours worked on day j. Machine breakdown on day $d-1$ leads to a loss of working hours on day $d-1$, i.e., $\frac{\partial h_{id-1}}{\partial B_{id-1}} < 0$, and, in turn, lowers production proficiency—i.e., $\frac{\partial H_{id-1}}{\partial B_{id-1}} < 0 \Rightarrow \frac{\partial x_{id}}{\partial B_{id-1}} < 0$. We can infer that $\frac{\partial^2 x_{ipmd}}{\partial B_{id-1} \partial H_{id-2}} > 0$, under the condition $f'' < 0$. In other words, while work interruption reduces proficiency, it is easier to get back up to speed for workers with more intensive work schedules (i.e., larger H_{id-2}). The damage is more severe for those with a smaller stock of work hours.

To test this potential channel, we investigate whether the effect of machine breakdown on day $d-1$ is different for workers with different working hours on day $d-2$. The new estimation specification is

$$y_i = \beta Treatment_i + \gamma Treatment_i \times Hours_i + \theta Hours_i + \varepsilon_i, \tag{4}$$

where $Hours_i$ is the number of working hours by worker i on the day before the machine breakdown incidence. With the hypothesis derived from the framework above, we predict that the effect would be larger when the affected worker operated the machine for a shorter length of time, i.e., fewer work hours, on day $d-2$. Table 3, column 2 reports the results of the test. Indeed, we find that the single term of $Treatment_i$ is negative and statistically significant, whereas the interaction between $Treatment_i$ and $Hours_i$ is positive and statistically significant. These results are in line with the proficiency hypothesis—that is, that workers with more hours worked previously are less affected.

Emotional Reaction. Machine breakdown may also cause emotional reactions—frustration, disappointment—which might raise the marginal cost of effort w_{id} and, therefore, reduce output on day d. As explained in the background section, workers' payment is entirely dependent on their daily output. Machine breakdown causes direct losses in working hours and payment, so it is plausible that workers feel disappointed about failing to earn less than expected for the day.

There is evidence that the emotions that arise from external shocks affect workers' output. For example, Ockenfels et al. (2014) show that when bonus payments fall short of individually assigned bonus targets, workers are disappointed, leading to lower work satisfaction and performance. Mas (2006) finds that after New Jersey police officers lose in final-offer arbitration over salary demands, relative to when they win, arrest rates and average sentence length decline and crime reports rise. More recently, Oswald, Proto and Sgroi (2015) show that increased happiness leads to higher productivity, and decreased happiness caused by major real-world shocks, including bereavement and family illness, leads to lower productivity. [7]

To check the relevance of the potential channel through negative emotions, we investigate whether the machine breakdown effect depends on the worker's experience of similar incidents. Our premise is that negative emotions, if any, are strongest and most difficult to cope with if the worker has never experienced such incidents before. In contrast, if the worker has experienced such situations before, his reactions—and resulting productivity—should be smoother. To this end, we estimate the following equation:

$$y_i = \beta Treatment_i + \gamma Treatment_i \times First_i + \theta First_i + \varepsilon_i, \tag{5}$$

where $First_i$ indicates whether the breakdown was the first incidence experienced by worker i. Regression results are reported in Table 3, column 3. The interaction term between $Treatment_i$ and $First_i$ is not only statistically insignificant but also small in magnitude, which indicates that there is no larger effect for workers who have no prior experience of machine breakdown. To this extent, our findings do not support the channel by which interruption lowers productivity through emotions.

[7]Relatedly, negative shocks in the workplaces may induce angry and even harmful behavior by the worker—e.g., drinking, violence, or deliberate sabotage—which in turn, lowers their second-day productivity. Card and Dahl (2011) observe that losses in professional football matches increase the rate of at-home violence by men against their wives or girlfriends. Lowenstein (2000) suggests that emotions, including a wide range of visceral factors, underpin daily economic behavior. For example, angry negotiators could become obsessed with causing harm to the other party, even at the cost of their own interests.

4.3 Heterogeneity and Peer Effects

We now explore the heterogeneous effects across worker characteristics. The literature suggests that there is substantial heterogeneity in economic behavior, which is partially accounted for by demographic information—e.g., gender, age, socioeconomic background, etc. For instance, Dohmen et al. (2011) conduct a study with a representative sample of roughly 22,000 individuals in Germany, and find that willingness to take risks is negatively related to age and gender, and positively related to height and parental education. Similarly, it has been suggested that gender plays an important role in economic preference (Croson and Gneezy 2009), and that for some decision behaviors, elderly individuals are less biased than younger individuals (Kovalchik et al. 2005).

Motivated by these studies, we examine whether the effect of machine breakdown differs by workers' gender, age, and place of residence. Table 4 reports the results. Columns 1 and 2 show results for males and females, respectively. We find that for both men and women, machine breakdown has negative and statistically significant effects on worker productivity on the subsequent day, but the effect is much larger for females than for males.

[Insert Table 4 here]

Columns 3 and 4 report, respectively, effects for young and older workers, relative to the sample mean (31.5). While the negative effects of machine breakdown are observed for both age groups, the effect is much larger for older individuals than younger ones.

Columns 5 and 6 present, respectively, the results for local and immigrant workers—i.e., those who are not registered as Fujian residents. The estimates appear to be similar for the two groups, which suggests that local residency does not play an important role in coping with the shocks of work interruption.

Finally, we check whether the effect is reinforced or mediated by peer effect. Peer effects in the workplace have been extensively explored in the literature. Mas and Moretti (2009) observe that a 10% increase in coworker productivity is associated with a 1.5% increase in a worker's productivity, which suggests that workers are motivated by social pressure and mutual monitoring. De Grip and Sauermann (2012) exploit a field experiment and observe that a 10 percentage points increase in the share of treated peers improves an individual's performance by 0.51%. Overall, evidence in the literature suggests a positive spillover from peers.

In our context, such spillover would imply that lower productivity by peers would reduce the examined worker's productivity as well. For a worker experiencing machine breakdown, it is possible that his reaction to the shock, and therefore his output, hinges on his peers' exposure to the shock and productivity. To take into account the influence of peers, we examine how individuals' responses vary by the percentage of machines that break down on the same day. For the same worker whose machine breaks down, if more coworkers are affected by the negative shock, the peer effect would enhance the effect and lead to a further reduction in output. We divide the sample in two based on whether the percentage of machines that broke down on day $d-1$ is above or below sample median. Indeed, we find—as shown in columns 7 and 8 of Table 4—that the effect is much larger when more machines on the site break down , relative to days on which only a few such incidents occur. Our results provide support for literature on the peer effect in the workplace.

5 Conclusion

This paper investigates how machine breakdown affects workers' subsequent productivity. Using daily output data, we show that individual productivity declines following a workday with machine breakdown. The adverse effect cannot be explained by either traditional theories of labor supply or income targeting behavior. In the sense that the marginal return on effort is larger when production is restored, workers should work harder on the following day. Or, if the worker has a targeted level of income that guides his labor supply, we would also observe higher subsequent output from the worker to compensate for his income losses during the interruption. By investigating possible channels, we find evidence that interruption may lower workers' proficiency, which, in turn, lowers their next-day productivity.

Our findings document a hidden cost of work interruption: the cost is not limited to the hours and productivity lost to the interruption, but may also spill over to subsequent production and can last for days. The hidden but economically significant cost is relevant to firms' decisions about how much resources to expend on managing interruptions, including the costs of maintaining equipment and arranging for standbys. A further implication lies in the remedies for interruptions. We find evidence that interruption causes subsequent output loss because of proficiency decline. One implication, therefore, is that managers may want to keep the affected individuals working, for instance, on other machines instead of letting them "cool down."

More generally, our study highlights work interruption as yet another determinant of productivity along with external factors such as incentive schemes, peer effects, weather and pollution, etc. It would be interesting to compare effects from other work interruptions or how the effect of machine breakdown can be generalized to other types of interruptions—especially those caused by communication devices and social media, as they are increasingly common and have become a concern for modern firms and organizations. These would be fruitful avenues for future research and contribute to better understanding of the causes and consequences of, and remedies for interruptions in the workplace.

References

Andersen, Steffen, Alec Brandon, Uri Gneezy, and John A List, "Toward an Understanding of Reference-Dependent Labor Supply: Theory and Evidence from a Field Experiment," Technical Report, National Bureau of Economic Research 2014.

Camerer, Colin, Linda Babcock, George Loewenstein, and Richard Thaler, "Labor supply of New York City cabdrivers: One day at a time," *The Quarterly Journal of Economics*, 1997, pp. 407–441.

Connolly, Marie, "Here Comes the Rain Again: Weather and the Intertemporal Substitution of Leisure," *Journal of Labor Economics*, 2008, *26* (1), 73–100.

Coviello, Decio, Andrea Ichino, and Nicola Persico, "Time Allocation and Task Juggling," *The American Economic Review*, 2014, *104* (2), 609–623.

_ , _ , and _ , "The Inefficiency of Worker Time Use," *Journal of the European Economic Association*, 2015, *13* (5), 906–947.

Crawford, Vincent P and Juanjuan Meng, "New York City Cab Drivers' Labor Supply Revisited: Reference-Dependent Preferences with Rational-Expectations Targets for Hours and Income," *The American Economic Review*, 2011, pp. 1912–1932.

Croson, Rachel and Uri Gneezy, "Gender Differences in Preferences," *Journal of Economic literature*, 2009, pp. 448–474.

Dohmen, Thomas, Armin Falk, David Huffman, Uwe Sunde, Jürgen Schupp, and Gert G Wagner, "Individual Risk Attitudes: Measurement, Determinants, and Behavioral Consequences," *Journal of the European Economic Association*, 2011, *9* (3), 522–550.

Farber, Henry S, "Is Tomorrow Another Day? The Labor Supply of New York City Cabdrivers," *Journal of Political Economy*, 2005, *113* (1).

_ , "Reference-Dependent Preferences and Labor Supply: The Case of New York City Taxi Drivers," *The American Economic Review*, 2008, *98* (3), 1069–1082.

Fehr, Ernst and Lorenz Goette, "Do Workers Work More if Wages Are High? Evidence from a Randomized Field Experiment," *The American Economic Review*, 2007, pp. 298–317.

Grip, Andries De and Jan Sauermann, "The Effects of Training on Own and Co-worker Productivity: Evidence from a Field Experiment*," *The Economic Journal*, 2012, *122* (560), 376–399.

Herrmann, Mariesa A and Jonah E Rockoff, "Does Menstruation Explain Gender Gaps in Work Absenteeism?," *Journal of Human Resources*, 2012, *47* (2), 493–508.

_ and _ , "Worker Absence and Productivity: Evidence from Teaching," *Journal of Labor Economics*, 2012, *30* (4), 749–782.

_ and _ , "Do Menstrual Problems Explain Gender Gaps in Absenteeism and Earnings?: Evidence from the National Health Interview Survey," *Labour Economics*, 2013, *24*, 12–22.

Ichino, Andrea and Enrico Moretti, "Biological Gender Differences, Absenteeism, and the Earnings Gap," *American Economic Journal: Applied Economics*, 2009, *1* (1), 183–218.

Jett, Quintus R and Jennifer M George, "Work Interrupted: A Closer Look At the Role of Interruptions in Organizational Life," *Academy of Management Review*, 2003, *28* (3), 494–507.

Kovalchik, Stephanie, Colin F Camerer, David M Grether, Charles R Plott, and John M Allman, "Aging and Decision Making: A Comparison between Neurologically Healthy Elderly and Young Individuals," *Journal of Economic Behavior & Organization*, 2005, *58* (1), 79–94.

Lee, David S, "Randomized Experiments from Non-Random Selection in US House Elections," *Journal of Econometrics*, 2008, *142* (2), 675–697.

_ and Thomas Lemieux, "Regression Discontinuity Designs in Economics," *Journal of Economic Literature*, 2010, *48*, 281–355.

Mandler, George, "Interruption (Discrepancy) Theory: Review and Extensions," *On the move: The psychology of change and transition*, 1990, *13*, 32.

Mas, Alexandre, "Pay, Reference Points, and Police Performance," *Quarterly Journal of Economics*, 2006, *121* (3).

_ **and Enrico Moretti**, "Peers at Work," *American Economic Review*, 2009, *99* (1), 112–145.

Ockenfels, Axel, Dirk Sliwka, and Peter Werner, "Bonus Payments and Reference Point Violations," *Management Science*, 2014.

Oswald, Andrew J, Eugenio Proto, and Daniel Sgroi, "Happiness and Productivity," *Journal of Labor Economics*, 2014. forthcoming.

Roy, Donald F, "" Banana Time": Job Satisfaction and Informal Interaction," *Human organization*, 1959, *18* (4), 158–168.

Spira, Jonathan B and Joshua B Feintuch, *The Cost of Not Paying Attention: How Interruptions Impact Knowledge Worker Productivity*, Basex New York, NY, 2005.

Table 1. Balance Test

	(1) Treatment			(2) Control			(3) Difference (1)-(2)	
	Mean	S.D	#Obs	Mean	S.D	#Obs	Mean	SE
Over-target output	0.114	0.007	1293	0.098	0.015	1315	0.016	0.017
Absence	0.275	0.011	1783	0.262	0.01	1783	0.0123	0.015
Work hours	10.751	0.066	1293	10.899	0.064	1315	-0.148	0.092
Machine breakdown	0.247	0.012	1293	0.221	0.011	1315	0.0262	0.017
Molding division	0.666	0.013	1293	0.666	0.013	1315	-0.0003	0.018
Reduced sample								
Age	31.154	0.263	1122	31.861	0.247	1169	-0.706**	0.361
Gender (male=1)	0.439	0.015	1122	0.428	0.014	1169	0.0117	0.021
Local Resident	0.397	0.015	1122	0.424	0.014	1169	-0.0277	0.021
Day shift	0.504	0.016	1011	0.498	0.015	1073	-0.007	0.022

Notes: This table compares the means of workers’ productive and demographic characteristics for the treatment (column 1) and control groups (column 2), separately. For each variable, we report the mean and standard errors from a t-test of equal means (column 3).

Table 2. Baseline Specifications

	(1)	(2)	(3)	(4)
	Baseline	Nonrandom components	DD	Placebo
Treatment	-0.020*	-0.022*	-0.031	-0.002
	(0.011)	(0.011)	(0.022)	(0.003)
Breakdown frequency in past 5 days		-0.126		
		(0.093)		
(Breakdown frequency in past 5 days)^2		0.053		
		(0.162)		
(Breakdown frequency in past 5 days)^3		-0.012		
		(0.073)		
Constant	0.142***	0.169***	0.031**	0.186***
	(0.004)	(0.009)	(0.015)	(0.002)
Observations	2,943	2,834	2,136	13,261

Notes: The dependent variables for columns (1), (2) and (4) is over-target output percentage. The dependent variable for column (3) is the difference in output between the considered workday and two days prior. All standard errors are clustered at the machine level.

Table 3. Interpretations

	(1) Work participation	(2) Accumulated work hours	(3) First breakdown
Treatment	-0.019*	-0.040**	-0.020*
	(0.010)	(0.018)	(0.011)
Working hours in day d-2		-0.002**	
		(0.001)	
Treatment*(work hour)$_{d-2}$		0.003**	
		(0.001)	
First time breakdown			-0.027
			(0.032)
Treatment* First time breakdown			0.002
			(0.032)
Constant	0.184***	0.156***	0.156***
	(0.009)	(0.008)	(0.008)
Observations	3,566	2,943	2,943

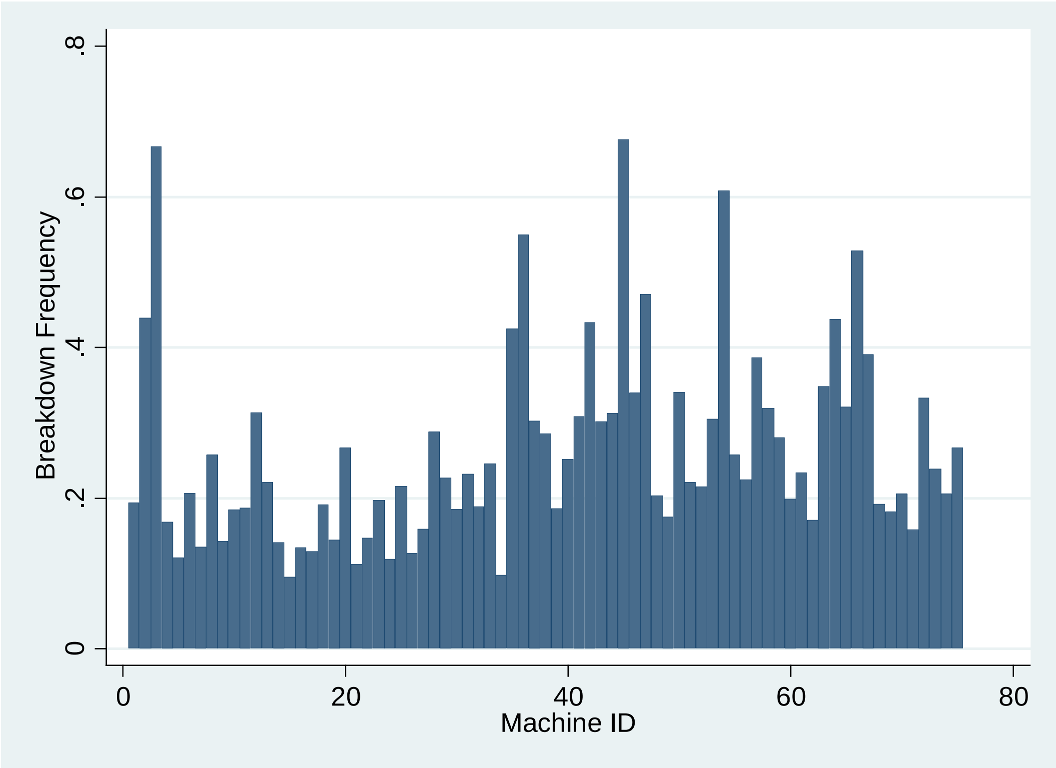
Notes: The dependent variable in column (1) is a dummy variable indicating whether the worker is present at work in day d. The dependent variable in columns (2) and (3) is the over-target output percentage. All standard errors are clustered at the machine level.

Table 4： Heterogeneity effect

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Male	Female	Older worker	Younger worker	Local (Fujian)	Migrants	Peer: High breakdown ratio	Peer: Low breakdown ratio
Treatment	-0.016**	-0.028***	-0.030***	-0.016**	-0.020**	-0.024***	-0.025*	-0.006
	(0.008)	(0.007)	(0.007)	(0.007)	(0.008)	(0.006)	(0.014)	(0.012)
Constant	0.145***	0.176***	0.175***	0.151***	0.169***	0.158***	0.144***	0.137***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.009)
Observations	842	1,084	937	989	783	1,143	2,130	813

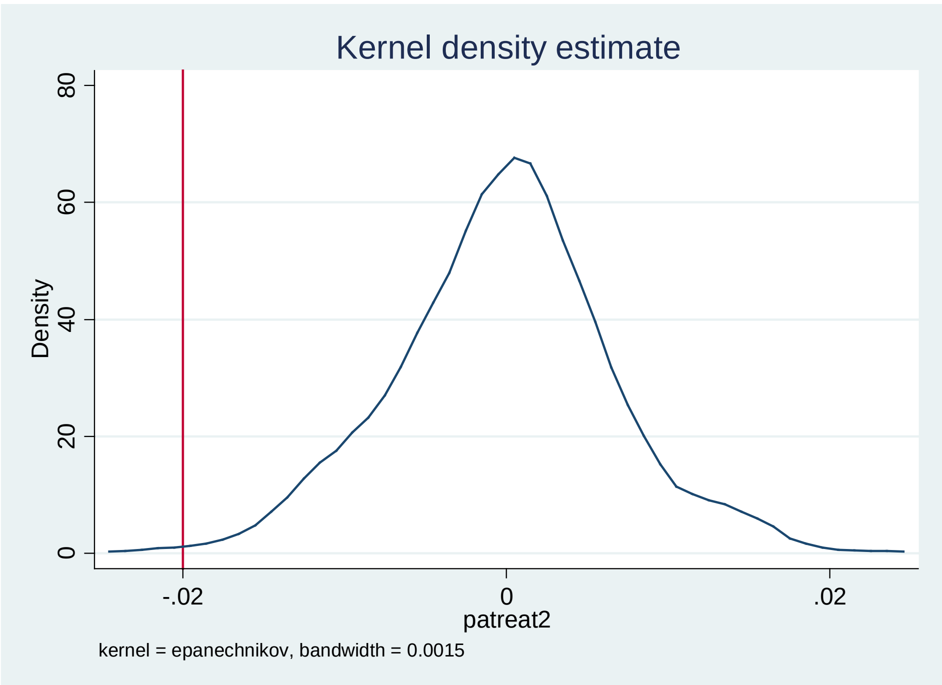
Notes: The dependent variable is the over-target output percentage. Columns (1) and (2) are estimated using subsamples of male and female workers, respectively. Columns (3) and (4) use samples of workers above and below the median age, respectively. Columns (5) and (6) use workers who have household registration status in the local province and those who do not, respectively. Columns (7) and (8) use observations in which the percentage of broken machine on the site is above or below the median, respectively. All standard errors are clustered at the machine level.

Figure 1: Breakdown frequency for all machines



Notes: This graph presents the breakdown frequency for all 75 machines in our regression sample. The horizontal axis is an ID assigned to each machine; the vertical axis is the frequency that it breaks down during the sample period.

Figure 2: Placebo test: Randomly assign breakdown incidents



Appendix Table A1: Robustness Checks

	(1) Pre-trends	(2) Correct for selection: lower bound	(3) Correct for election: upper bound
Treatment	0.014	-0.014	-0.036***
	(0.020)	(0.011)	(0.010)
Constant	-0.013	0.136***	0.159***
	(0.012)	(0.004)	(0.003)
Observations	2,163	2,902	2,915

Notes: Column (1) presents the results of test for parallel pre-trends. The dependent variable is the difference in worker output between day d-2 and d-3, i.e., one and two days prior to the considered date of machine breakdown. Column (2) and (3) report the lower and upper bound of estimates after adjusting for selection bias.