Lifting the Veil: The Face of TFP in an Indian Rail Mill[1]

Sanghamitra Das

Indian Statistical Institute, Delhi

Kala Krishna

Pennsylvania State University and NBER

Sergey Lychagin

The Pennsylvania State University

Rohini Somanathan

Delhi School of Economics

November, 2009.

[1]We are grateful to Ana Cecilia Fieler, George Alessandria for comments. Sanghamitra Das passed away unexpectedly in December 2008 and is greatly missed.

Abstract

The importance of Total Factor Productivity (TFP) in explaining output changes is widely accepted, yet its sources are not well understood. We use a proprietary data set on the floor-level operations at the Bhilai Rail and Structural Mill (RSM) in India to understand the determinants of changes in plant productivity between January 2000 and March 2003.

During this period there was a 35% increase in output with minimal changes in the stock of physical capital or the number of employees, but sizable reductions in the number and duration of various types of production delays. We model interruptions to the production process as a function of worker characteristics and find that a large part of the avoidable delay reductions are attributable to training. Overall, changes in all delays account for over half the changes in productivity.

Our results provide some explanation for the large within-industry differences in productivity observed in developing countries and also suggest that specific knowledge-enhancing investments can have very high returns. Our approach also provides an example of how detailed data on production processes can be fruitfully used to better understand TFP changes, which have typically been treated as residuals in growth-accounting exercises.

Keywords: Total Factor Productivity (TFP), Plant level data, Competitiveness and trade.

JEL Classification: D24, J24, L23, L61, M53.

1 Introduction

This paper studies a remarkable improvement in productivity that occurred at the Bhilai Rail and Structural Mill (RSM) in India over the period 1999-2003. Prior to 1999, the Bhilai RSM was the sole supplier of rails to Indian Railways. A series of train wrecks, culminating in a major accident that killed 210 people in 1998, led to investigations. The accidents were found to be due to sub-standard rails. These rails were produced using steel with high hydrogen content, which made the steel brittle, and the rails substandard. This finding resulted in the suspension of rail purchases by the Indian Railways from the Bhilai RSM and new, more stringent, specifications for future purchases were introduced.[1] The government was under pressure to import rails in 1999, and in fact did so. Questions were raised in parliament and the government was contemplating allowing private players into rail production. In this setting, the RSM faced not just competition, but a threat to its very existence. A productivity surge ensued. How was this surge obtained? What lessons can be drawn from their experience? These are our main questions.

Total Factor Productivity (TFP) is widely used as a measure of performance in firm, industry and more aggregate country-level studies. By definition, this is the increase in output that cannot be attributed to changes in observable inputs. Despite it being a mere residual, it has been found to be important in explaining income differences across countries. For example, Hall and Jones (1999) find that in the 35-fold difference in output per worker between the United States and Niger, TFP differences explain about twice as much as differences in physical and human capital. Within-country differences in TFP have also been shown to be large for developing countries. Hsieh and Klenow (2008) use plant level data to show that TFP in India and China is much more dispersed than in the US. Rationalization of production that would bring this dispersion down to US levels could raise output by 30-50% in China and 40-60% in India.

A large range of institutional and policy variables could lie behind these TFP patterns, such as access to credit, physical and social infrastructure, technological spillovers and managerial practices to name just a few. While the above studies and much of the literature treats TFP as a black box and evaluates its response to policy changes through its correlation with a handful of observed variables, more recent work tries to go even deeper.

[1] Hindu Business Line, June 9, 2000, Indian Express, December 12, 2004.

This line of work is aided by the fact that very detailed data are collected by the businesses themselves as a part of their normal practice and are increasingly accessible to researchers.[2] The use of such information could help shed light on the types of allocative inefficiencies that underlie variations in productivity. However, developing ways of using such data is a challenge in itself! We utilize a simple and convenient framework that could be applied to a wide range of industries using the information available in floor-level data.

The dataset that guides us in this effort documents daily operations at the Bhilai RSM. The mill operates continuously with 3 production shifts per day. We obtain shift-wise data on the number of steel blooms rolled into rails in each shift, a list of all workers present during the shift (with their designations), and all delay episodes with their duration and a description of the cause of the delay. We combine these data on the production process with administrative data on worker characteristics and all types of training. Even though the overall number of workers remained relatively unchanged, the numbers and combination of workers on the floor changes from one shift to another, and we use this variation to estimate the determinants of output changes.

Our focus on the details of production processes can be seen as a part of a wave of work that relies on insider data to capture details of what is going on inside the firm. This could be detailed data on the floor that provides an objective view of what is happening there, more subjective data about management practices, or purely subjective data about something, such as job satisfaction. This "bottom-up" approach has been at the frontier of research in both economics and management. Our work is most closely related to that of Ichniowski et al. Ichniowski et al. (1997) use monthly data to look at the productivity effects of human resource management practices for 36 steel finishing lines across the United States. They find that workers in plants with traditional employment contracts and hierarchical supervisory structures are less productive than those in firms with innovative practices. A major source of this productivity gain is increased uptime. Also related is the work of Das and Sengupta (2004), who study blast furnaces making steel in India. They find that productivity increased by raising the quality of coal used, but that managers did not contribute to production unless they are trained. Another example of work that uses "insider" data is Bloom and

[2]Such work has been labelled "insider econometrics" as only insiders typically have access to such data, see Ichniowski and Shaw (2008).

Van Reenen (2007), who use survey data evaluating management practices combined with balance sheet data to estimate TFP in order to see if the two are correlated. This work can be contrasted to the usual approach that treats the firm as a single entity and uses aggregate factors of production like labor and capital without specifying explicitly what happens inside the firm.

Also related is literature that examines the response of TFP to competition. Galdon-Sanchez and Schmitz (2005) show that when competitive pressures mounted in the market for iron ore due to the collapse in the market for steel in the early 1980s, countries with mines that were close to becoming non-competitive increased efficiency, while others did not. Schmitz (2005) argues this efficiency increase came about from improved work practices. Loosening restrictive work practices increase labor productivity both because it allowed excess staff to be shed, but also because it allowed down time to be cut so that the mine machines can operate for longer parts of the day.

In our data, we also find that productivity increases occurred at a time when the plant was under severe competitive pressures due to their inability to produce rails to quality specifications. An interesting contrast with other studies is that we find increased uptime in the absence of any sizable changes in employment. Other work on competition and productivity includes Caves and Christensen (1980) , Rodriguez and Rodrik (1999), Tybout et al. (1991), Tybout and Westbrook (1995), Trefler (2004) and Nickell (1996).

While we do not have data on plants not threatened by closure to compare to those that were, we do have much finer data on a single plant during events that led to increased competitive pressure. Our data is far more detailed and extensive than that in any previous work in the area which allows us to look at what occurred on the shop floor and suggests where the apparent productivity improvements came from. We find that the lion's share came from reductions in preventable delays and increases in the rate of production. However, there were no changes in work rules that were constraining efficiency in our case. Instead, the main causes of improvements were the programs of employee training, better supply of quality raw materials and investments in new equipment.[3]

The plan of the paper is as follows. Section 2 provides some background of the BSP and the technology of production. Section 3 provides an overview

[3]These do not seem to have been very costly so the puzzle is why these were not undertaken previously!

of the data. Section 4 decomposes output growth into its component parts, that due to changes in rates of production, in delays and in the fraction defective and explores the patterns in each of these. Section 5 models production focusing on delays and fits it to the data. Section 6 develops some counterfactual experiments to identify the contribution of the various policies that might have led to the growth in TFP. Section 8 outlines the lessons learnt and concludes.

2 Background

The Bhilai Steel Plant covers about 17 square km. It provides work for roughly 55,000 regular workers and another 12,000 contract workers. Employing contract workers is a way around the restrictive labor laws that essentially prevent the firing of a worker. The jobs of regular workers are secure, with excellent fringe benefits including schooling, health care and housing, as well as travel benefits and ample leave (51 days a year).[4] As a result, these jobs were, and by all accounts remain, highly valued.

The description of the work environment is in some ways contradictory. On the one hand, we have the descriptions of life in the Bhilai Steel Mill from the social anthropology side as in the work of Parry (1999). His description is of a work environment where bursts of effort alternate with periods where little real work is required. Though shifts are 8 hours long, few workers stay for more than half of the shift and many for far less. Even when technically at work, much time is spent drinking tea and reading the newspaper, or just roaming around as there is far more labor on the books than needed. This is not surprising as there is little reason for workers to put in effort. There is little correlation in pay and performance, and much more between pay and seniority. Even overtime was abolished when it was determined that it encouraged workers to put off work so as to obtain overtime hours and earnings. Absenteeism is high, especially during peak demand times in agriculture. In fact, in Parry's mind, the real puzzle is why, despite little reason to do so other than a feeling of self respect and job satisfaction, a good share of them seem to work.

The other view is that which we obtained from field visits. Work on the floor of the mill is hard and intense and is often performed under adverse

[4]Much of the description below is based on Parry (1999), which is a must for the uninitiated. Not only is it informative, but it is just plain fun to read!

circumstances as the temperature on the floor can be quite high. While some stations have air conditioning, the floor does not and exposure to the heat from the furnace, especially during the hot summer months, creates extreme conditions. In addition, managers seem to put in long hours and pitch in where needed.

2.1 The Rail and Structural Mill

The Rail and Structural Mill is an integral part of the Bhilai Steel Plant (BSP). It was commissioned in 1960 with enough capacity to satisfy domestic demand at that time. Since then, it has been the sole supplier of rails for Indian Railways. The plant has had problems keeping up with orders from the Railways, although the stated objective of management was output maximization.

In 1998 and 1999, after a string of train accidents, which were supposedly caused by high hydrogen content in the rail steel, the Railways committed to using steel of higher quality. This quality could not be consistently provided by BSP at that time. Consequently, Indian Railways suspended orders from the BSP and switched temporarily to imports.[5]

At this time there was considerable questioning of the ability of the BSP to adequately provide the rails needed by the Railways. Imports from China and opening up the market to domestic entrants was considered. A question was even raised in Parliament on this issue. Combined with the liberalization on the domestic and import side that India has undertaken since the late 90's, this put pressure on the RSM to perform. Workers were suddenly aware their plum jobs were at risk. If competition forced the mill to close, they could all be legally fired.

3 Data.

In order to model production, we need to understand how the rail mill operates. The details of this are contained in Appendix 1. In a nutshell, the inputs (*"blooms"*) enter the rail mill from the *"bloom yard"* and pass through different sections in a sequential process. In this process, they are transformed

[5] According to the UN Comtrade data, in 1999 India imported 92,000 tons of rails. This is almost three times as much as it used to import in 1998 and more than eight times as much as it imported in 2007.

into not defective rails, or what is termed "blooms rolled". The rails are kept in the *"cooling bed"* before being going to the *"finishing bay"* to be cut and stacked. Defective blooms (*"blooms cobbled"*) are set aside. The production cycle at the Rail and Structural Mill runs 24 hours a day, 7 days a week, with very rare shutdowns for service and repairs. A typical day consists of three 8-hour shifts. For each of these shifts the available data includes total input of steel, the total output, the share of defective output, the date, time, the identities of workers who were present on the floor, their characteristics, such as age, years worked, caste, education, training, as well as the reason for and length of each delay in production that occurred during the shift. More on the data is below.

3.1 Shifts and Brigades

There are 3558 total shifts in the dataset, covering the dates between January 1, 2000 and March 31, 2003. We dropped year 1999: at that time, the problem of high hydrogen content has not been solved and the mill seems to have been intentionally operated below full capacity.[6]

Shifts are operated by brigades – groups of workers relatively stable over time. There are more people in a brigade than typically work in a shift, allowing for days off on the part of workers. The brigades are rotated weekly across shifts: if a brigade works the morning shift on week 1, it is switched to work the afternoon shift on week 2 and so on. Overall, labor on the floor per shift increased only by a bit less than 10%, while output rose by more than 35%.

3.2 Workers

We have detailed information on employees. First, there is an attendance sheet for each worker, which lists the shifts he spent on the floor, as well as his designation in the production process. We can track people as they are transferred across brigades, get hired, fired or promoted. The data allows us to control for the composition of labor force, which is not possible in more aggregate studies.

Second, we have information on the social background of the personnel, including caste affiliation and home state. As informally observed by Parry

[6] This can be clearly seen in Figure 7; the rolling rate was remarkably low throughout most of 1999.

Table 1: Categories of training

Category	Days spent	Days, %
Motivational	511	27
Productivity	479	26
Environmental	184	10
Quality Control	165	9
Cost Reduction	135	7
Safety	131	7
Computer Skills (IT)	61	3
Job Instruction	57	3
Other	151	8
Total	1874	100

(1999), the local population of Bhilai experienced tensions with the newcomers, who had often moved from other states of India to work at the plant. Potentially, communal conflicts like this may be strong enough to impair cooperation at the workplace and decrease productivity. We use both caste and home state data to account for this possibility.

Finally, we obtained very detailed data on training programs administered to the employees. Each program is given a brief description, a list of trainees, a starting and an ending date. The programs can be roughly split into nine categories by the purpose and the targeted skill: motivational, productivity, environmental management, quality control, cost reduction, safety, computer skills, job instruction, other training (see table 1).

On average, the recipients were less experienced than their peers.[7] Some training was conducted because it helped in obtaining International Organization for Standardization (ISO) certification. Most programs did not seem to target any particular designation; only few of them focused on a narrow

[7]This is especially noticeable for the computer skills, cost reduction, safety and motivational training types.

workplace-specific skill. With minor exceptions, training was administered to big groups of workers rather than individuals. Below is the list of the four biggest programs; they account for two thirds of total training time.

1. Acceptance of rails program – accounts for 22% of total training time, administered in June–July 2001, productivity category;

2. ISO-9000 workshop – 9% of total training, two episodes: May 2001 and March 2002, quality control category;

3. ISO-14001 workshop – 10% of total training, January 2002 and July 2002, environmental management category;

4. Success through empowerment of people – 24% of total training, October 2002 – January 2003, motivational category.

3.3 Delays

We have comprehensive information on delays, their durations and causes. Delays are of four kinds. First, there are "outside delays"; these are denoted by D_o, and usually occur due to events outside the control of the managers, such as shortages of steel, electricity or gas. These may be unanticipated, as in the case with gas shortages, or anticipated but unavoidable as in the case of electricity blackouts or load shedding.

Second, there are finishing delays, D_f. These are delays caused by the finishing bed being full and unable to accept more rails. This is a downstream constraint that can shut down or slow down production in the mill.

Third, there are planned delays, D_p, which are used for scheduled maintenance or adjustments of equipment.

A fourth kind of delay is the most important one for our analysis; it consists of unplanned delays avoidable for the brigade and is denoted by D_a. These are the delays caused by workers making mistakes. We will argue that reductions in D_a made possible notable productivity improvements at the RSM during the time period of interest.

4 Decomposition of Output Growth.

In this section we define an analog of the production function. We formulate it in terms of variables that are either directly available in the data or can be easily calculated. This way, we utilize the information on the internal structure of the process instead of relying on the totals of raw inputs and the final output as most of plant-level studies usually do.

Let X_s denote the total number of steel blooms used by the brigade on duty during shift s. Some part, p_s, of these blooms are successfully rolled into rails, while the remaining ones get "cobbled" and removed from the line as defective. The final output is

$$Y_s = p_s X_s.$$

The number of blooms the brigade is able to process is the product of uptime T_s and rolling rate R_s[8]:

$$X_s = R_s T_s.$$

Since we know the uptime and the input of steel for each shift, we can infer the processing rate as X_s/T_s. Uptime makes 480 minutes less time spent in delays of all kinds:

$$T_s = 480 - D_{o,s} - D_{p,s} - D_{f,s} - D_{a,s}.$$

Combining these three equations we obtain:

$$Y_s = p_s R_s (480 - D_{o,s} - D_{p,s} - D_{f,s} - D_{a,s}). \tag{1}$$

As mentioned above, we know precisely what type of product was rolled in each shift.[9] To be sure, our statistics are not contaminated by movements in the product mix, we only keep "rail" shifts in the sample.[10]

[8] Das and Sengupta (2004) refer to R and T, as *the rate of output* and *the rate of utilization.*

[9] From discussions with the management we gathered that the product mix is not driven by price/cost margins (in any case prices are administerd and not market) but is demand driven since this is a public sector undertaking. As the mill is the sole supplier for Indian Railways, they first produce what the Railways need. The buyers of structural steel are big public sector firms such as BHEL and NTPC with whom they have long term relationships. Production decisions do not depend on prices, but rather on excess capacity after rail demand is satisfied. In coming years they expect proportion of structural output to go up as other firms enter the rail market.

[10] The reason why we chose not to analyse structural shifts at all is because their output is very heterogenuous (beams, angles, channels, crane rails, steel sleepers) and they account for too few observations in 2003 to produce reliable growth statistics.

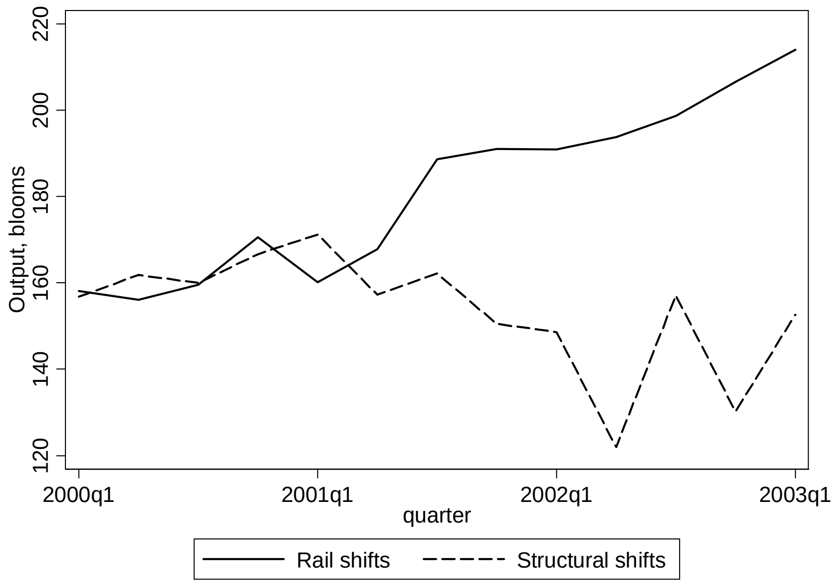


Figure 1: Output per shift

Average output per shift grew from 158 blooms in the first quarter of 2000 to 214 blooms in the first quarter of 2003, a 35% increase (see Figure 1).[11]

As equation 1 suggests, there are six parts of the production function that may be responsible for this growth: p_s, R_s, and four downtime values $D_{x,s}$.[12] In the remainder of this section, we study each of them in more detail. However, before proceeding further, it is useful to compare the importance of these variables by looking at their contributions to output growth. By definition, the percentage change in blooms rolled equals the sum of the percentage change in p_s, R_s, and uptime. The change in uptime can be further decomposed into its component parts as done in Table 2. For example, the contribution of outside delays to growth in total uptime is $\frac{(46.8-31.9)}{183-129} = .276 \approx .28$.

[11]It is interesting to draw a comparison to a case from the similar industry studied by other authors. Ichniowski et al. (1997) found that the introduction of new human resource management practices accounted for 2-8% growth in output of the U.S. steel finishing lines.

[12]This is slightly richer than that in the literature. Ichniowski et al (1997) for example, focus only on the increase in uptime as the major source of productivity improvements.

Table 2: Components of the production function and their marginal contributions to output growth

	Q1 2000	Q1 2003	Growth
p	0.987	0.995	3%
R	0.54	0.614	42%
$(480 - D)$	297	351	55%
D_o	46.8	31.9	28%
D_p	90.4	72.5	33%
D_f	2.41	2.22	0.35%
D_a	43.6	22.7	39%
D	183	129	100%
Y	158	214	100%

According to table 2, finishing downtime and fraction non-defective do not seem to be very important. The other two, D_o and D_p, contribute greatly, but represent the delays that are unavoidable to the workers, and thus are exogenous to a certain degree. The remaining variables, R and D_a, are determined on the mill floor and make a substantial contribution to the increase in output, and hence warrant most of our attention.

4.1 Fraction Non-defective

The growth in the non-defective fraction is about a percentage point (Figure 2). This variable contributes very little to productivity growth per se. However, the same figure also shows that the percentage of defective rails fell from 1.3% to 0.5%, a fall of more than a half. The fraction defective is significantly higher in structurals than in rails (1.7% vs. 0.6%), so the overall reduction of defects is largely attributed to the change in the composition of output towards rails.

Like avoidable delays, defectives are associated with mistakes on the part of workers. This is consistent with the data: the correlation between share defective and avoidable downtime is positive and significant at the 1% level.

Figure 2: Fraction non-defective

These two variables appear to be closely related to each other. Hence, their effect on output will be more apparent in changes in avoidable delays than directly via a reduction in defectives: hence we do not look any more at the direct effect of defectives.

4.2 Delays

A regular shift rarely runs without delays in production. Delays make a considerable part of a work day; during fiscal years 2001–2003 they accounted for 30% of an average shift time.

4.2.1 Avoidable delays

The avoidable downtime decreased by almost half, from 43 minutes per shift in the first quarter of 2000 to 22 minutes in the first quarter of 2003 (Figure 3). The outlier in Q2 2001 coincided in time with an unusually long shutdown in production. According to our information, the time when the mill stood idle was used for training and equipment replacements. A training episode to

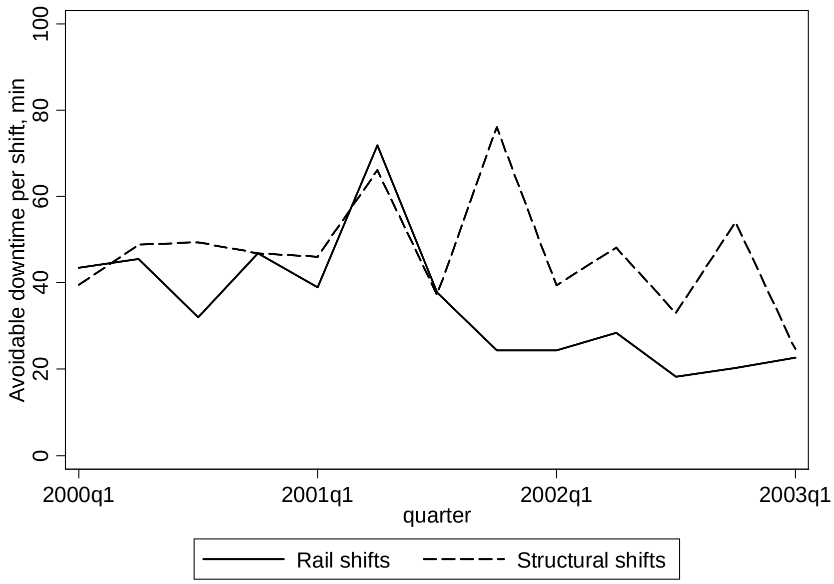


Figure 3: Avoidable downtime

raise productivity in rails termed the "acceptance of rails program" occurs at this time.[13] The decline in delays that followed may thus have been caused by either training or equipment replacement. However, it is reasonable to expect the better equipment is likely to get broken less frequently in both rail and structural shifts, which is not observed in the data (for structurals, the avoidable downtime becomes even higher in Q4 2001). The training program explicitly focused on raising the output of rails, which is consistent with the observed decrease downtimes during rail shifts, but not structural ones.

Although delay descriptions are available in the data, it is rather problematic to locate the source of each delay on the process chart in Figure 10. In each case, there is no reference to any person or group who were at fault, so we can only rely on management's ability to properly classify delays as avoidable.

The descriptions look very heterogeneous. Most entries contain one of these keywords: "not working", "tripped", "fallen", "broken", "jammed", "grinding", "adjustment", "crane down".

[13] As we argue below, this training episode is the only one that looks like it actually worked.

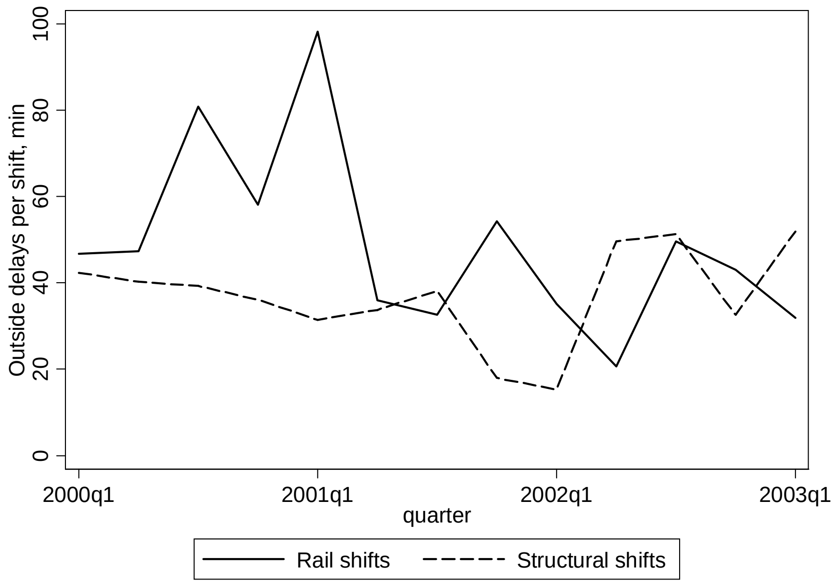


Figure 4: Outside downtime

4.2.2 Outside Delays

The patterns of outside delays were very different for rail and structural shifts (Figure 4). This is especially noticeable for year 2000, when RSM faced problems with the supply of low-hydrogen steel. Since hydrogen content is not so important for the steel used in heavy structurals, as it is for rails, there was less outside downtime during structural shifts.

More than 60% of outside delays were associated with insufficient supply of inputs (keywords “shortage”, “voltage”, “restriction”) or their bad quality (keywords “lengthy”, “short”, “bad metal”, “asymmetry”). Some keywords from the “avoidable delay list” are also frequently used for outside delays (e.g. “adjustment”, “broken” and “jamming”). However, they account for less than 20% of observations. Even if some part of this 20% was misclassified by the management, this is unlikely to make a dramatic impact on our statistics.

During 2001–2003, outside delays stayed at approximately the same level, but were very volatile at the same time. For this reason, they do not noticeably affect long-run growth of output, but they do greatly affect quarter-to-

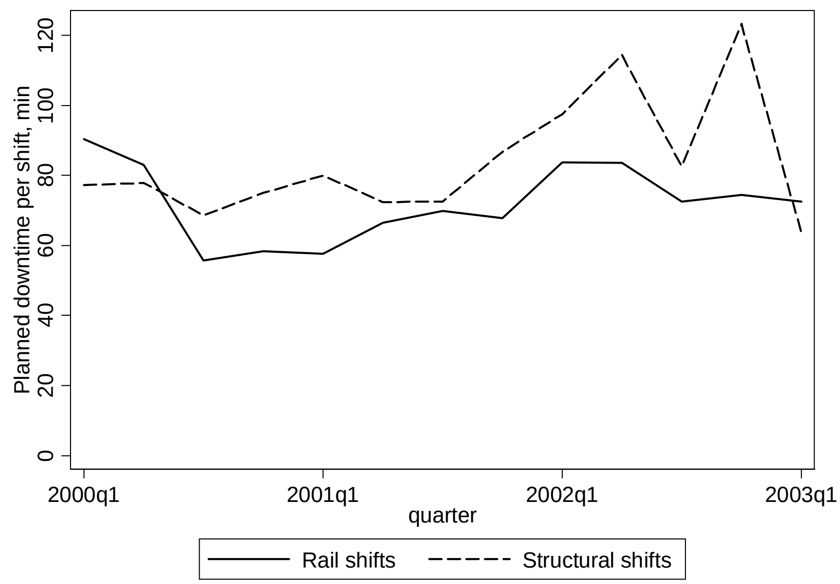


Figure 5: Planned downtime

quarter fluctuations.

To summarize, there was a striking fall in outside delays in the first and second quarter of 2001, while after that, the average remained roughly the same, though there was a lot of variance. Since the fall in 2001 was outside the control of the mill, we say no more about this.

4.2.3 Planned Delays

After an initial drop in 2000, planned downtime has been slowly increasing until mid-2002 (Figure 5). We interpret this increase as a natural consequence of higher capital utilization. As output per shift grows over time, the equipment requires more frequent service. We did not observe any qualitative difference between rail and structural shifts which is consistent with this interpretation.

The descriptions of delay causes suggest that planned delays were primarily used for regular maintenance. More than 90% of planned delays were associated with "checking", "adjustment" and "changing" of "section", "stand" or "hot saw disc".

4.2.4 Finishing Delays

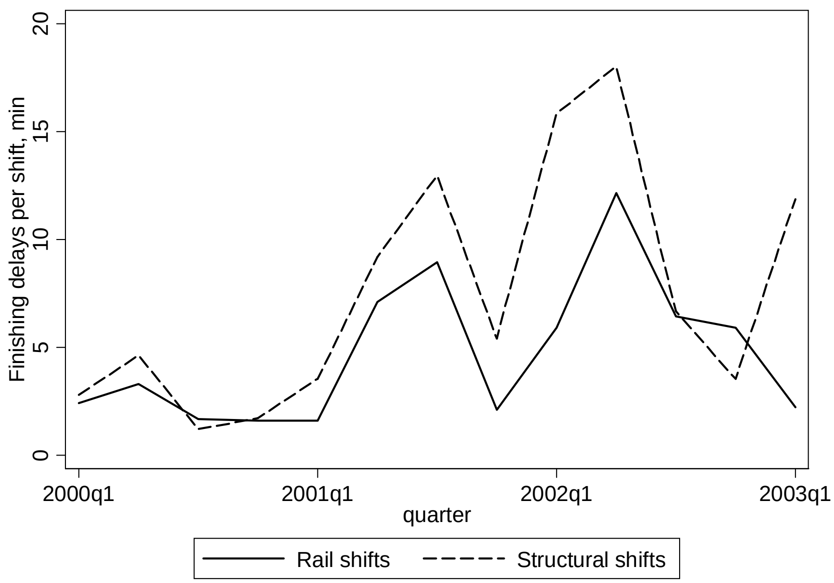


Figure 6: Finishing delays

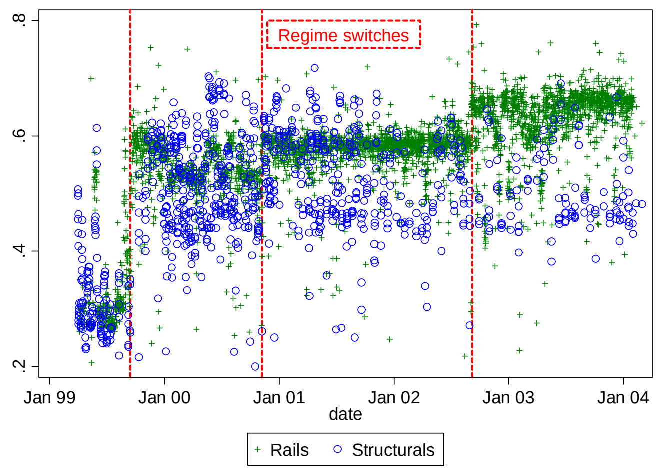
Finishing delays fluctuated around 5–6 minutes per shift in 2000–2003 (Figure 6). This is a very low level compared to all other delay categories. Finishing downtime change does not contribute much to output growth per se. However, it serves as a source of information about downstream bottlenecks that may restrict the productivity of the mill.

Finishing delays occur at the final phase of production – when the rails are coming from the Hot Saw section to the cooling bed. There is only one cause listed for all finishing delays: "cooling bed full". If there is no sufficient space on the cooling bed, the operations at the Rail Mill are halted until the space becomes available.[14]

[14] To check for consistency of this story, we ran a regression of finishing delay on the performance measures of the upstream sections: processing rate, fraction non-defective, outside, avoidable and planned delays. If it is true that the worse Rail Mill personnel perform, the easier it is for downstream workers to keep up and not cause finishing delays.

According to the estimates, delays at the rail mill lead to lower finishing downtime. The processing rate coefficient is negative, but not significant. This may be explained by

Figure 7: Rolling rates



4.3 Rolling Rates

Figure 7 plots the dynamics of the rolling rate. It includes both heavy structurals and rails.

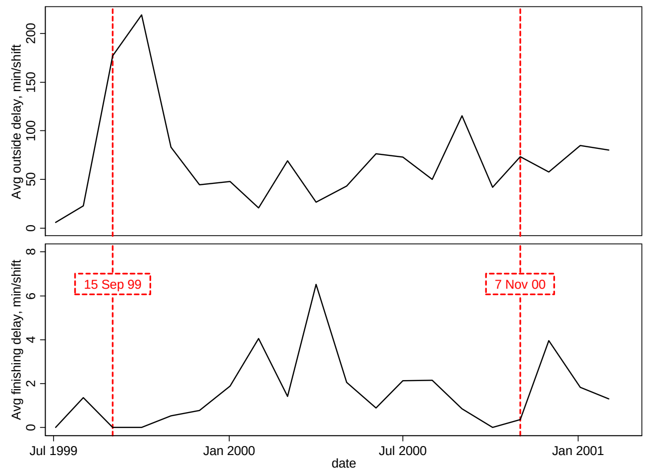
It is evident from this figure that the rates seem to switch between discrete regimes. The switching clearly occurs at least three times: on September 15th 1999, November 7th 2000 and September 4th 2002. Within each regime, the rates are dispersed around some average level that is stable overtime. This makes sense as dispersion will naturally arise in day to day operations.

Before September 1999, the mill had few orders as it was deemed incapable of producing the required quality. As a result, it was operating far below capacity. It is recorded that one furnace out of four was running between Sept. 1999 and November 1999, consistent with the low average rolling rate in this period.

After this first switch, between September 1999 and November 2000, there is a period where rolling rates fluctuated from one level to another. In this period, the mill had limited access to low hydrogen steel from outside. There are two ways of reducing the hydrogen content in the rails. One is to use a degasser to make better steel. The other is to accept steel with a high

possible endogeneity: the workers at the Rail Mill may slow down when the cooling bed is nearly full. Overall, the data seems to support our understanding of how finishing delays occur.

Figure 8: Finishing and outside delays



hydrogen content but to cool the rail slowly, allowing hydrogen to escape (see Rai and Agarwal (2007)). The mill installed a degasser in early 2000.[15] It took six months or so to get consistent operation out of this unit, and until October of 2000, it was not fully effective. Note the high and variable levels of outside delays and finishing delays as depicted in Figure 8 (consistent with using the finishing bed to slowly cool the rails to reduce the hydrogen content) in this period.

After November 2000, the degasser was running consistently, and this is reflected in the higher, more stable, rate pattern. Finally, the regime switch that took place on September 4th 2002 is explained by the installation of some new equipment. We identify this using delay cause descriptions. On September 4th, a new delay cause started appearing in the data; it is listed as "jamming at new descaling unit". This delay occurred nine times in the first three days following the regime switching. Gradually, its frequency declined to five occurrences per quarter. Since the increase in the rolling rate occurred simultaneously with the installation of the new equipment, we conclude that the former was likely to be caused by the latter.

Overall, it seems fair to say the long run dynamics of the rolling rate seem

[15] It is recorded that the degasser was put in for hot trials in March 2000 (Hindu Business Line Newspaper, June 9th 2000.) The Degasser was effective October 1st 2000, as recorded in the controller general report 2003.

to be determined by technological considerations and the outside constraints operating.

5 A Semi Structural Model of Production

Now we turn to estimating the production model. We could do so at various levels of reality. One way would be to model the process on the computer, including all interactions that occur at various stages of the production process. For example, in practice, whenever any upstream delay occurs, it makes the downstream finishing delay less likely, as rails can exit the cooling bed during this time. Then, we could choose some arbitrary values of the parameters of the generating process, generate data using these parameters, then choose the parameters to best match the relevant moments of the generated data to the actual data.[16] This SMM (simulated method of moments) approach is widely used in structural estimation.

We choose a more simple route to begin with that does not allow feedback loops to be present. While such loops are undoubtedly part of reality, they result in enormous complications for the estimation, but we hope to graduate to incorporating them in future work.

We build the following stylized model of the Rail Mill. Each bloom that is sent to the Rail and Structural Mill triggers the following sequence of events:

1. A steel bloom is fed into the furnace area for reheating.

2. An event may occur at this point, call it an "outside event". When this outside event occurs, it triggers an outside delay. We model the occurrence of the event by drawing a binary variable M_o from the distribution $\Pr\{M_o = 1\}$. We assume that the probability of an outside delay is constant within a calendar quarter q: $\Pr\{M_o = 1\} = P_o(q)$. This will be a rough approximation of the true probability distribution. As the factors that create delays outside the RSM are not covered in the data we choose to model them in the simplest possible way.

[16] For a concrete example, consider an ice cream stall. Customers arrive at random, say according to a Poisson process. Each customer takes t time to serve. As a result of random arrival, customers may have to wait in line. This generates a distribution of waiting times. A similar distribution of waiting times could be generated by the computer for given parameters of the Poisson arrival process. These parameters could thus be estimated by making the empirical and simulated distributions of waiting times as close to each other as possible.

(a) If $M_o = 1$, then a delay is triggered. This takes D_o minutes, where D_o is drawn from the distribution $F_o(D_o|Z)$. We assume $F(.)$ depends on Z, the characteristics of the brigade on the floor. Some delays involve intervention by the Rail Mill personnel; this is why we condition $F_o(D_o)$ on Z.[17]

(b) If $M_o = 0$, there is no outside delay.

3. The workers make an avoidable mistake ($M_a = 1$) with probability $\Pr\{M_a = 1|Z\}$.

   (a) If $M_a = 1$, it takes D_a minutes to fix the mistake. D_a is sampled from $F_a(D_a|Z)$.

   (b) Otherwise, there is no avoidable delay.

4. The bloom is rolled into the final product. This takes time t where $t = 1/R$ where R is the rate of production. This rate is set to be the quarterly rate in the data and so is denoted by $R(q)$, where q denotes the quarter. As we demonstrated in the previous section, the processing rate seems to be driven largely by outside factors and is switched rather infrequently. Since we cannot control for these factors, we approximate R by quarterly averages.

5. The cooling bed gets full ($M_f = 1$) with probability $\Pr\{M_f = 1\} = P_f(q)$. In other words, we set the probability quarter by quarter to equal the number of finishing delays relative to the blooms rolled. We do not think we are missing much by doing so as changes in finishing downtime per se make a very small contribution to output growth.

   (a) If $M_f = 1$, the workers at the Rail Mill stand by for D_f minutes until the cooling bed is cleared. The delay duration D_f is drawn from $F_f(D_f)$. It does not depend on Z, since Rail Mill personnel are not involved into clearing the delay.

   (b) Otherwise, no finishing downtime is registered.

[17] For example, flooding in the rainy season requires darainage of the affected area before production can be resumed.

6. With probability p the final product is non-defective, and this is set at the share of non-defective blooms quarter by quarter and so is denoted by $p(q)$. Again, we lose little by doing so since its contribution to output dynamics is small.

7. The equipment requires maintenance ($M_p = 1$) with probability $\Pr\{M_p = 1\} = P_p(q)$. Again, this is proxied for quarter by quarter by the number of planned delays relative to blooms rolled. Though it can be argued that planned delays are neither low, nor exogenously determined, approximation by quarterly averages should be precise enough, as long as the equipment requires uniform servicing per unit of input.

   (a) If $M_p = 1$, D_p minutes are spent in a planned delay, where D_p is drawn from $F_p(D_p|Z)$.

   (b) If $M_p = 0$, no maintenance is scheduled.

8. The process is repeated starting from step 1.

We assume the processes that generate delays of various kinds have a logit form. For avoidable delays the probability of a delay depends on brigade characteristics, so

$$\Pr\{M_a = 1|Z\} = (1 + \exp(-\theta_a Z))^{-1}.$$

For other delays this probability is allowed to vary by quarter only, thus

$$\Pr\{M_x = 1|q\} = (1 + \exp(-\theta_x(q)))^{-1}, \; x = o, p, f.$$

Turning to the length of the delays, the distributions $F_x(D_x|Z)$ for $x \in \{a, o, p\}$ are assumed to take the gamma form[18] where the shape parameter, β, is allowed to depend on Z while the scale parameter, λ, is not. Thus:

$$f(D_x|M_x = 1, Z) = D_x^{\beta_x Z - 1} \frac{e^{-D_x/\lambda_x}}{\lambda_x^{\beta_x Z}\Gamma(\beta_x Z)}, \; x = a, o, p.$$

We chose the Gamma distribution for its flexibility and as it fits the data quite well. In Appendix 2 we plot the actual data and show that the

[18] In the data, the shape of downtime distribution seems to be well approximated by the exponential family, which is a special case of the gamma family. Exponential distribution is commonly used in queuing theory to model service times.

fitted gamma distributions do a good job approximating the data. The other reason we choose this form is that it allows for a simple interpretation of the estimates. Assume the observed delay durations come from the sum of delays caused by each individual on the floor and that these individually generated delays are independently generated. Then if the delays of a single worker come from the gamma distribution $\Gamma(\alpha_i, \lambda)$ where α_i is the individual characteristic of worker i, by the summation property of the Gamma distribution, the total delay is distributed as $\Gamma(\Sigma_i \alpha_i, \lambda)$. This property makes the estimates easy to interpret.

We assume that $F_f(D_f)$ takes a similar gamma form but with the additional restriction that the shape parameter is not dependent on Z, though it is allowed to vary by quarter. Thus

$$f(D_f|M_f = 1, q) = D_f^{\beta_f(q)-1} \frac{e^{-D_f/\lambda_f}}{\lambda_f^{\beta_f(q)} \Gamma(\beta_f(q))}.$$

Recall that the mean of the gamma distribution is given by $\lambda\beta Z$, while the variance is $\lambda^2\beta Z$. Thus, our parametrization (for all but the finishing delay) allows both the mean and the variance of the delays to depend on who is on the floor. For finishing delays, it allows the mean and variance to vary by quarter only.

We could estimate the model we describe by using SMM techniques. If there are correlations between the random processes in the model we would have no choice but to do so or to use some other such technique. To stay simple, we further assume that all random processes in the model (M_x, D_x) are jointly independent conditional on Z. This means the only correlation between outside, avoidable and planned delays goes through the brigade on the floor. This assumption allows avoidable, outside and planned delays to be estimated independently of each other.

We estimate β_x and λ_x by applying the method of maximum likelihood to the sub-sample of blooms with positive D_x. We do this independently for $x = a, o, p, f$ (avoidable, outside, planned and finishing delays). The unknown parameters are straightforwardly estimated by using maximum likelihood. It is well known that this likelihood function is concave and so has a unique maximum. See for example, Choi and Wette (1969). For completeness, we provide a sketch of the proof in Appendix 2.

We estimate the probability of delays of various forms by applying the logit model to the sample of all rolled blooms.

6 Estimation

We begin by describing the variables we use in our estimation.

6.1 Brigade Characteristics

The vector of brigade characteristics Z includes labor by designation ($LABOR\_X$), diversity indices ($LOCAL\_MIX$ and $CASTE\_MIX$), and training by category ($SSTOCK\_X$).

6.1.1 Labor

The technological process is organized around ten groups of workers. Seven of them are assigned to a particular section of the line (see Figure 10) in Appendix 1. The other three groups are Executives, Crane operators and Technicians, who may appear at any stage of the process. Since different groups perform different tasks, we treat them as separate types of labor. We construct ten total labor variables for each shift: number of workers in the Services group ($LABOR\_SERV$), in the CM (control men) group ($LABOR\_CM$), and so on. The variables are mnemonically defined to correspond to the labor groups in Appendix 1.

6.1.2 Diversity

Efficiency of labor depends greatly on the level of cooperation within working groups. Communal and social tensions could hinder cooperation, which would make labor productivity decline.

In his study of Bhilai working communities, Parry (1999) observed the local population was increasingly anxious about the inflow of newcomers. At the same time, he rejected the claim that people of different castes are uncooperative at the workplace.

To allow for a possible link between communal tensions and productivity, we look at two dimensions of diversity: home state and caste affiliation. We construct an index for both of them:

$$LOCAL\_MIX = \min(S_{local}, 1 - S_{local}),$$
$$CASTE\_MIX = \min(S_{scst}, 1 - S_{scst}),$$

where S_{local} is the share of locals in a brigade, and S_{scst} is the share of workers from scheduled castes (backward castes) and tribes.

6.1.3 Training

Brigade characteristics also include training obtained by the workers. For each individual in the sample we construct nine training stocks – one per category of training as defined in Table 1:

$$
\begin{aligned}
ISTOCK_X(w,s) &= \sum_{t<s} TRG_X(w,t) \\
X = & MOTIV, PROD, ENVIR, QC, SAFETY, \\
& COST, IT, JOBINST, OTHER.
\end{aligned}
$$

That is, an individual stock of worker w on day s equals the total amount of training administered to him by that date. To obtain shift-level values, we add up the individual stocks of all workers on the attendance sheet:

$$
SSTOCK_X(s) = \sum_{w \in attend_s} ISTOCK_X(w,s).
$$

The model is estimated on the sample of "rail" shifts that covers the period of January 1, 2000 – March 31, 2003. Table 3 presents the results.

The estimates suggest that more labor does not necessarily leads to higher output. First, note that (column 1 of Table 3) whenever the coefficient on the number of workers on the floor of a given type in explaining the probability of an avoidable mistake is significant, it is positive. Second, there is no clear evidence that the mistakes are fixed faster by larger brigades (column 2-4 of Table 3). Hence, an increase in labor input is unlikely to reduce downtime and raise output per shift. These results are supported by the anecdotal evidence on overstaffing at the BSP given in Parry (1999).

Productivity training seems to lower the probability of avoidable mistakes. The effects of ISO 14001 workshop and motivational training are also significantly negative. However, as we are about to show, they do not survive a change in model specification.

The estimation results do not support the hypothesis that diversity impairs productivity; on average, caste heterogeneous brigades seem to make fewer mistakes.

Total labor variables alone might not be able to capture all the relevant dynamics in workforce composition. By looking at total numbers of employees, we implicitly assume that workers' characteristics are homogeneous within groups. If this is not the case, our estimates may be subject

Table 3: Estimates of downtime components

Dependent variable	M_a, logit	D_a, ML	D_o, ML	D_p, ML
LABOR_CM	$2.75^{\dagger}$	-0.19	-0.97	0.48
LABOR_COGG	-1.24	-0.25	1.24	0.40
LABOR_CRANE	-0.26	0.25	3.56^{*}	-0.23
LABOR_EXEC	-5.24	-0.48	0.92	-1.28
LABOR_FURN	-0.46	-0.13	-2.10	-0.16
LABOR_GRST	$1.78^{\dagger}$	0.33	0.58	-0.03
LABOR_SCM	2.54^{*}	$-0.43^{\dagger}$	-1.13	-0.11
LABOR_SERV	$3.14^{\dagger}$	0.02	-1.00	-0.10
LABOR_SS	0.65	-0.11	-0.92	0.02
LABOR_TECH	8.18^{**}	-0.21	0.83	0.81
SSTOCK_COST	4.66	-0.03	-3.14	0.19
SSTOCK_ENVIR	-7.52^{**}	0.06	-0.77	-0.72
SSTOCK_IT	2.53	-1.20	3.03	0.29
SSTOCK_JOBINST	-4.04	1.30	0.72	1.75
SSTOCK_MOTIV	-1.76^{*}	0.17	0.71	-0.12
SSTOCK_OTHER	-4.71	-1.08	-0.13	0.79
SSTOCK_PROD	-4.79^{**}	-0.13	$-1.08^{\dagger}$	-0.08
SSTOCK_QC	-1.09	-0.16	-1.97	-0.17
SSTOCK_SAFETY	-0.60	0.86	11.90^{**}	0.89
CASTE_MIX	-2.11^{**}	-10.8	-26.0	-18.5
LOCAL_MIX	0.34	-7.47	-36.5	15.6
Observations	418,819	3,366	1,788	5,219

Significance levels : $\dagger$: 10% $*$: 5% $**$: 1%

Column 1: SSTOCK coefficients are scaled up by 1000, LABOR coefficients are scaled up by 100

Columns 2-4: SSTOCK coefficients are scaled up by $10\lambda_x$, all other coefficients are scaled up by λ_x

Unlisted control variables: brigade dummies

Table 4: Estimates of downtime components – robustness check with worker dummies

Dependent variable	M_a, logit	D_a, ML	D_o, ML	D_p, ML
LABOR_CM	1.40	-0.30	-8.36	1.28
LABOR_COGG	0.11	-0.36	1.71	1.23
LABOR_CRANE	-3.14	5.22	5.93	1.22
LABOR_EXEC	-0.17	0.46	-1.26	-2.26
LABOR_FURN	2.49	-2.30	-19.1	3.85
LABOR_GRST	2.23	1.85	-12.1	1.46
LABOR_SCM	1.22	-0.18	-12.7	0.86
LABOR_SERV	3.82	2.74	-9.55	2.17
LABOR_SS	2.03	2.54	-13.0	0.33
LABOR_TECH	-1.24	-1.62	20.0	-4.64
SSTOCK_COST	-10.6	2.24	4.14	-2.80
SSTOCK_ENVIR	-0.24	-0.77	-0.85	-0.95
SSTOCK_IT	-4.73	0.27	21.5	2.61
SSTOCK_JOBINST	68.8**	4.61	-30.0	5.63
SSTOCK_MOTIV	0.99	-0.21	0.48	-0.60**
SSTOCK_OTHER	-4.87	0.45	-3.88	1.06
SSTOCK_PROD	-5.96**	-0.43	-1.88	-0.10
SSTOCK_QC	-3.18	-1.00	-1.38	-0.46
SSTOCK_SAFETY	4.79	2.88	9.96	0.30
CASTE_MIX	-6.18	-92.3	-262	3.67
LOCAL_MIX	0.16	-5.11	46.8	11.8

Significance levels : † : 10% $*$: 5% $**$: 1%

Column 1: SSTOCK coefficients are scaled up by 1000, LABOR coefficients are scaled up by 10

Columns 2-4: SSTOCK coefficients are scaled up by $10\lambda_x$, all other coefficients are scaled up by λ_x

Unlisted control variables: worker dummies to the omitted variable bias. To address this issue, we perform a robustness check. We augment Z by individual worker dummies and reestimate these equations.

The estimates presented in table 4 confirm our main result: productivity training significantly reduces downtime. In addition, job instruction training becomes associated with more frequent avoidable mistakes, while motivational training seems to reduce the average time spent in planned delays. Caste mix is no longer significant. Thus, the only robust result is that productivity training helps.

7 Counterfactual Experiments

The estimates reported in tables 3 and 4 apply to the separate elements of the production function. In this section, we put these elements together to study the impact of counter factual changes in labor, diversity and the stock of training on the overall output. For example, could run the model with all the elements that existed at the start of the period. Then we add the effect of productivity training and see how much of the variation over time this explains.

To implement this we simulate production bloom by bloom, following the same multistep procedure outlined in Section 5. Each bloom that enters the mill takes time 1/R. to be processed if no mistakes or delays occur. If the model generates the event that a draw from a delay distribution is warranted, then the delay drawn is added to this time. Many delays may occur and these are additively incorporated. At the end, there is a probability that the bloom may be deemed defective, i.e., cobbled, in which case the simulation will trow this bloom out. This goes on till the 480 minutes of the day are used up. At the end of each day, the total blooms rolled are generated. We then take the monthly output generated by the simulation and label this to be the simulated output. The simulation are run using the parameters estimated in the previous section. To avoid omitted variable bias, we choose the version with worker dummies.

In *each* of our counterfactual experiments, the set of brigade characteristics is split in two parts: $Z = [Z_1, Z_2]$. This split varies simulation by simulation. The first part contains variables that we freeze at the level of quarter 1, 2000 in the simulation. The second consists of characteristics that are allowed to change over time as in the data. This way, we predict the time path of output that would occur, had the management chosen not to adjust the variables in Z_1.

We start with simulating a full model, i.e., where Z_2 contains all the covariates:

$$Z_2 = [CASTE\_MIX, LOCAL\_MIX, LABOR\_j, WD, SSTOCK\_k]$$

where j takes all the values for the different kinds of labor and k takes all the values for the different kinds of training. Z_1 is therefore empty. Then, we incrementally shrink the list of variables in Z_2, and observe how simulated output changes in response.

The results of these experiments are depicted in Figure 9. Panel (a) shows that the full model does a good job of fitting monthly output data. In the interest of not over parametrizing the model, we only allow for quarterly changes in the probabilities of outside, planned and finishing delays. As a result, if outside delays, for example, are high in a particular week or month, the model will not pick it up. This is why the model does not track the data spike by spike, but it does a good job of tracking it on average.

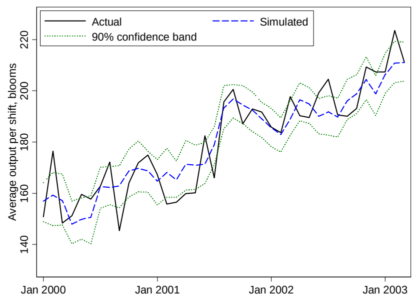
In panel (b), we assume that the diversity indices are kept at the level of Q1 2000. This does not change output predictions much (compare panels (a) and (b)). In other words, the effect of changing diversity is very small in its absolute magnitude.

In the next panel, we impose restrictions on an additional set of variables: total labor and worker dummies. Now the management does not control the composition of workforce at all. The fluctuation in output are now driven by training and the outside factors only. This causes a slight decrease in predicted output starting in the beginning of 2002 onwards. This suggest that the adjustments in labor employment and the composition of the working brigades are unlikely to have been responsible for output growth. This contrasts the findings by Schmitz (2005), who argues that the productivity improvements at the US iron ore mines were primarily driven by the reductions in the excessive workforce. The example of BSP demonstrates that the inability to fire unnecessary workers need not curtail productivity improvements.

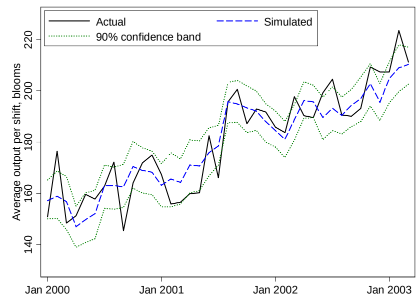
Panel (d) shows what output would be produced if no changes in diversity or labor composition were allowed and only productivity type of training were administered to the workers. This way, we shut down the effects of all training that does not belong to the productivity category. Although the

Figure 9: Diversity, labor and training, and their overall effects on output

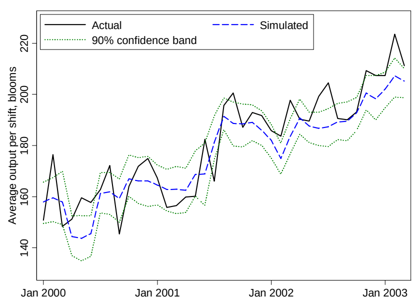
(a) Full model: $Z_2 = Z$



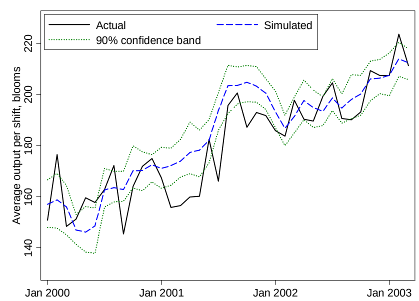
(b) Freeze diversity indices at the level of Q2 2001: $Z_2 = [LABOR\_X, WD, SSTOCK\_X]$



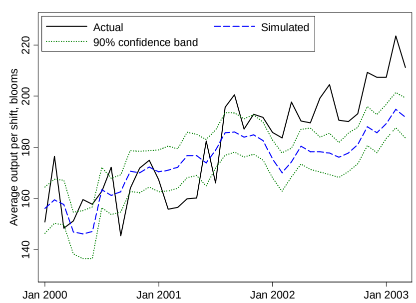
(c) Freeze diversity, total labor and worker dummies: $Z_2 = [SSTOCK\_X]$



(d) Freeze all covariates, except for productivity training: $Z_2 = SSTOCK\_PROD$



(e) Freeze all covariates: $Z_2 = []$



Confidence bands are based on simulation draws only, conditional on the model parameters. They do not account for the randomness in the parameter estimates.

latter is the only covariate not frozen in time, the model still fits the data well.

There is a period in the first and second quarter of 2001 when the model in panel (d) over-predicts the output. In the panel (c) specification, this over-prediction is canceled out by something. This something may be the negative effects of job instruction and safety training. Intuitively, job instruction is likely to be administered when the workers are hired or promoted to the positions they never held. In our data, most promotions and hires take place in the beginning of a calendar year. Thus, the inflow of inexperienced workers around January 2001 could coincide in time with the job instruction training, so that this training is associated negatively with productivity. Safety training also contributes to the downtime; this type of training is likely to occur in the periods when the management is highly concerned about safety. Since there is always a trade-off between safety and the speed of production, there is no surprise that safety training is associated with lower output.

Finally, in panel (e) we fix all covariates at the level of Q1 2000. The predicted time path of output is driven by the outside factors only, such as outside mistakes and variations in the processing rate. This experiment yields a large discrepancy between prediction and the actual data.

The last two panels suggest that productivity training was crucial in increasing output. Had management done nothing to train the employees, the growth in output would have been much more modest.

8 Conclusion

In this paper we study a proprietary dataset that documents floor-level operations at Bhilai Rail and Structural Mill, a unit of Steel Authority of India, during a time when output increased by about a third in response to external pressures. We provide a decomposition of output changes into six components: due to changes in the processing rate, the share of defectives and the four types of downtime. We find that changes in the rate and delays account for a almost all of the growth. Delays account for about 55% of the growth. Avoidable, planned and outside delays account for 39, 33, and 28% respectively of the contribution of all delays.

We then present and estimate a simple model of production that goes beyond the traditional production function approach and exploits the structure of the technological process.[19] Our estimated model allows us to turn on and off various channels through which production could have increased. By conducting such counterfactual experiments, we show, for example, that most of the growth in production that came from reductions in avoidable delays occurred due to a single training episode. We also show that there were significant changes in rates across quarters that we attribute to the installation of new equipment.

Our results suggest that training and other improvements have the capability to increase output significantly. Yet these were not undertaken by the firm until the threat of closure in the face of non-performance became dire. In this way, our work suggests that the old fashioned "x-inefficiency" a la Liebenstein (1966) exists today, at least in public sector undertaking like the Bhilai RSM.

References

Bartel, Ann. 1991. "Productivity Gains from the Implementation of Employee Training Programs." NBER Working Paper 3893.

Bloom, Nicholas, and John Van Reenen. 2007. "Measuring and Explaining Management Practices Across Firms and Countries." *Quarterly Journal of Economics*, 122(4): 1351–1408.

Caves, Douglas W., and Laurits R. Christensen. 1980. "The Relative Efficiency of Public and Private Firms in a Competitive Environment: The Case of Canadian Railroads." *Journal of Political Economy*, 88(5): 958–976.

Choi, S. C., and R. Wette. 1969. "Maximum Likelihood Estimation of the Parameters of the Gamma Distribution and Their Bias." Technometrics, 11(2): 683–690.

Das, Sanghamitra, and Ramprasad Sengupta. 2004. "Projection Pursuit Regression and Disaggregate Productivity Effects: The Case of the Indian Blast Furnaces." *Journal of Applied Econometrics*, 19(3): 397–418.

[19] The model that we propose is not specific to the steel industry. It may be applied to any production unit involved in a processing task of an arbitrary nature. Since many manufacturing firms are organized around tasks in established technological chains, our approach is most likely to be useful in the manufacturing sector.

Emerson, Michael, Michel Aujean, Michel Catinat, Philippe Goybet, and Alexis Jacquemin. 1988. *The Economics of 1992: The E.C. commission's assessment of the economic effects of completing the internal market.* New York: Oxford University Press.

Galdon-Sanchez, Jose E., and James A. Schmitz, Jr. 2002. "Competitive Pressure and Labor Productivity: World Iron-Ore Markets in the 1980's." *American Economic Review*, 92(4): 1222–1235.

Hall, Robert and Charles Jones. 1999. "Why Do Some Countries Produce So Much More Output Per Worker Than Others?" Quarterly Journal of Economics, Vol. 114, No. 1, February, Pages 83-116.

Ichniowski, Casey and Kathryn Shaw. 2008. "Insider Econometrics: A Roadmap for Estimating Empirical Models of Organizational Design and Performance." Mimeo.

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

Klenow, Peter and Chang-Tai Hsieh. (2008) "Misallocation and Manufacturing TFP in China and India" forthcoming in the Quarterly Journal of Economics.

Leibenstein, Harvey. 1966. "Allocative Efficiency and X-Efficiency", *The American Economic Review*, 56, pp. 392-415.

Nickell, Stephen J. 1996. "Competition and Corporate Performance." *Journal of Political Economy*, 104(4): 724–746.

Parry, Jonathan P. 1999. "Lords of Labour: Working and Shirking in Bhilai." *Contributions to Indian Sociology*, 33(1-2): 107–140.

Schmitz, James A., Jr. 2005. "What Determines Productivity? Lessons From the Dramatic Recovery of the U.S. and Canadian Iron Ore Industries Following Their Early 1980s Crisis." Research Department Staff Report 286 Federal Reserve Bank of Minneapolis.

Rai, Abhai Kumar, and Atul Agarwal. 2007. *Rail Steel*, Indian Railways Institute of Civil Engineering, Pune.

Rodriguez, Francisco, and Dani Rodrik. 1999. "Trade Policy and Economic Growth: A Skeptic's Guide to the Cross-National Evidence." NBER Working Paper 7081.

Rothwell, Geoffrey, and John Rust. 1997. "On the Optimal Lifetime of Nuclear Power Plants." *Journal of Business and Economic Statistics*, 15(2): 195–208.

Rust, John. 1987. "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher." *Econometrica*, 55(5): 999–1033.

Trefler, Daniel. 2004. "The Long and Short of the Canada–U.S. Free Trade Agreement." *American Economic Review*, 94(4): 870–895.

Tybout, James, Jamie de Melo, and Vittorio Corbo. 1991. "The Effects of Trade Reforms on Scale and Technical Efficiency: New Evidence From Chile." *Journal of International Economics*, 31(3–4): 231–250.

Tybout, James R., and M. Daniel Westbrook. 1995. "Trade Liberalization and the Dimensions of Efficiency Change in Mexican Manufacturing Industries." *Journal of International Economics*, 39(1–2): 53–78.

9 Appendix 1: Production in the Mill

The layout plan of the RSM can be pictured as an assembly line production process, being made up of different sections, where employees of different designations work, either individually or in groups (see Figure 10). The inputs ("blooms") enter the rail mill from one end and pass through these different sections in a sequential process, coming out as rails or "blooms rolled" from the other end or discarded blooms ("blooms cobbled"). Thus, the final output of the rail mill can be modelled as a combined effort or output of these different sections with a vertical structure (as production is sequential) coming from the assembly line feature of the process.

The layout of the rail mill can be divided into 3 sequential parts: the furnace, the mill area and the finishing area that has a "hot saw area" and a "cooling bed".The steel slabs are first reheated in the furnace and then enter the mill area where they are rolled. Ultimately the final output comes out in forms of standardized rails or "blooms rolled" from the hot saw area and goes on to be cooled on the cooling bed.

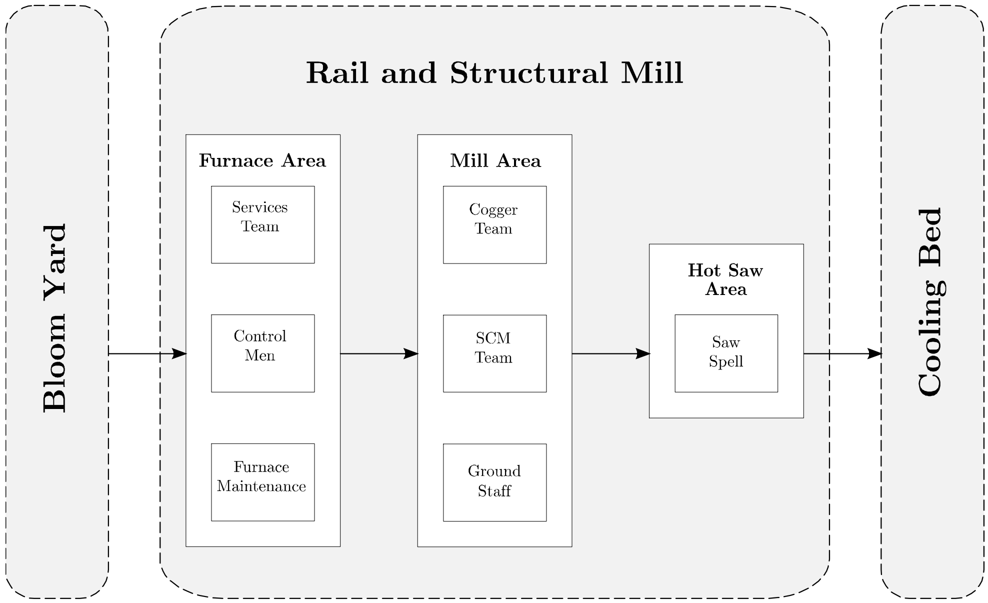


Figure 10: Production process

The production process has a sequential as well as a hierarchical structure of employment. At the top of the operational hierarchy are the executives (designations: Roller and Foreman), who interact with the heads of non-executives. The furnace area is headed by Master Operative (reheating) and in his absence by Senior Operative (reheating). The mill area is headed by the Assistant Roller.

Let us first describe the furnace area of the rail mill. The first three pulpits are used for recording of the inputs that come through. The Services group (designations: Operative(services) and Senior Operative (services)) do the recording job and pass on the information to the Master Operative (reheating) before the inputs enter the main production process.

The next three pulpits in the furnace area are manned by the Control Men, where each of them works individually. Within a shift, there maybe more than three Control Men, but at any given point of time only one of them mans each pulpit. They are responsible for pushing the blooms into the furnace and also for pushing out the intermediate output into the mill area.

In the furnace area, there also works a group of people of different designations, who are responsible for the proper working of the furnaces – they see to it that there is no jamming or jumping in the furnace during the production process. We call this group the Furnace Maintenance group, which is made up of the Operative (Furnace), Senior Operative (Furnace) and Senior Operative (Reheating).

We next come to the mill area. The next three pulpits are the most important sections of the rail mill in the sense that the chance of having "blooms cobbled" occurs here. Pulpit 6 is manned by a group of Coggers, Manipulators and main Motor Operators designated as the Cogger group. Pulpits 7 and 8 are manned by a group of Control Men (Mill), Senior Control Men and Main Motor Operators designated as the SCM group.

Next we come to the biggest section of the rail mill, the ground stand and roll building area. We call this group the Ground Staff, which is comprised of the Roller (executive), the Assistant Rollers (head of non-executives in the mill), the Senior Operatives (Roll Building), the Operatives/Senior Operatives (Mill) and the Plant Attendant.

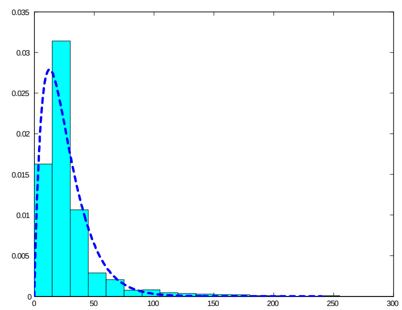
Finally we come to the Hot Saw area and the stamping machine. Here we have the Operatives (Saw Spell) and the Senior Operatives (Saw Spell), stamping the blooms rolled and cutting them to standard lengths. We call this group the Saw Spell. This is the point at which the blooms are recorded as having been rolled in our data. After this, the rails pass on to the Cooling Bed

There is also a shift manager, or the Foreman, who takes charge of the entire production process of the rail mill. Sometimes there are two Foremen in a shift. In case there is no Foreman on a given shift, the Roller (shift in-charge) takes over. We call this group the Executive group; it supervises the entire production process. In addition, we have Crane Operators who remove cobbled blooms and Technicians who appear mostly in the maintenance groups. This is, in a nutshell, the assembly line production process of the rail mill.

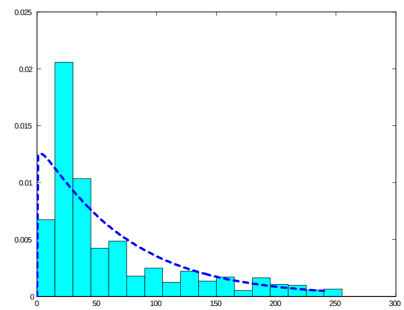
10 Appendix 2: Use of the Gamma Distribution

10.1 The Fit to the Data

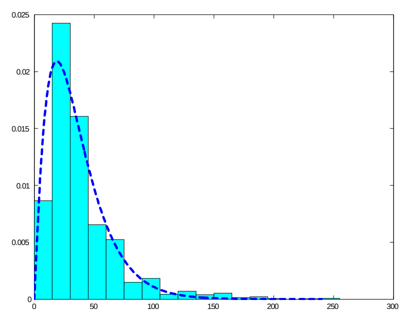
We chose the Gamma distribution for its flexibility and a fairly good fit with the actual data. Below the histograms represent the actual distributions, while the dashed lines are fitted gamma pdf's where the parameters are chosen to best fit the actual data. As is evident, the fitted distributions are every close to the data.



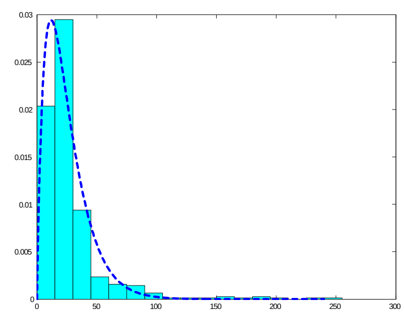
Avoidable delays



Outside delays



Planned delays



Finishing delays

10.2 Estimation

We assume that α is a linear function of brigade characteristics: $\alpha = \beta' X$. The values of β and λ are not restricted to be the same across delay classes; this gives four separate unknown pairs (β, λ), one pair per class. For each delay class the data provides us with a sample of delay durations D_s, indexed by shift, as well as the brigade characteristics X_s. The unknown parameters are straightforwardly estimated by using maximum likelihood. Loglikelihood function takes the following form:

$$\ln L = \sum_s \left[(\beta' X_s - 1) \ln D_s + \beta' X_s \ln k - \ln \Gamma(\beta' X_s) - k D_s \right], \text{ where } k = 1/\lambda$$

It is easy to demonstrate that $\ln L$ is a concave function of β and k. Let S be a matrix of its second derivatives, $y = \begin{bmatrix} y_1 & y'_{-1} \end{bmatrix}'$ be an arbitrary vector partitioned into the scalar y_1 and the remainder y'_{-1}. Let ψ_1 denote a trigamma function: $\psi_1(z) = \frac{d^2}{dz^2}(\ln \Gamma(z))$.

$$
\begin{aligned}
y'Sy &= \begin{bmatrix} y_1 & y'_{-1} \end{bmatrix} \begin{bmatrix} -k^{-2}\beta' \sum X_s & k^{-1} \sum X'_s \\ k^{-1} \sum X_s & -\sum \psi_1(\beta' X_s) X_s X'_s \end{bmatrix} \begin{bmatrix} y_1 \\ y_{-1} \end{bmatrix} = \\
&= \sum_s \left(-\frac{\beta' X_s}{k^2} y_1^2 + \frac{2X'_s}{k} y_{-1} y_1 - \psi_1(\beta' X_s) y'_{-1} X_s X'_s y_{-1} \right) = \\
&= \sum_s \left(-\left(\sqrt{\frac{\beta' X_s}{k^2} y_1^2} - \sqrt{\frac{y'_{-1} X_s X'_s y_{-1}}{\beta' X_s}} \right)^2 - \left(\psi_1(\beta' X_s) - \frac{1}{\beta' X_s} \right) y'_{-1} X_s X'_s y_{-1} \right)
\end{aligned}
$$

As the trigamma function $\psi_1(z) = \sum\limits_{n=0}^{\infty} \frac{1}{(z+n)^2}$, and since $\frac{1}{(z+n)^2}$ decreases in n,

$$
\psi_1(z) = \sum_{n=0}^{\infty} \frac{1}{(z+n)^2} > \int_0^{\infty} \frac{dt}{(z+t)^2} = \frac{1}{z}.
$$

That means, $y'Sy \leqslant 0$, i.e., S is negative semi-definite. It is guaranteed to be negative definite if matrix $\sum\limits_s X_s X'_s$ has full rank, in which case the objective likelihood function is strictly concave. Therefore, the first order conditions for maximum likelihood have a unique solution that can be easily found numerically given arbitrary starting values of β and k.