Social Connections at Workplace and Productivity: Evidence from Indian Garment Manufacturing

Swati Sharma [a]

June 2018

Abstract

We look at the daily output of workers in the garment factories for 31 production days to study the effects of caste based social connections at workplace on firm's performance by studying worker productivity and production line productivity. We find that at the worker level, working with more workers of his/her caste increases worker's productivity. These results become stronger after controlling for the unobserved individual heterogeneity (such as ability) amongst the workers. An increase of 1 percentage point in the share of own caste of the worker in her production line, increases her performance by almost 10%. At the line level, we find that a line is more productive if it is more homogeneously organised in terms of caste distribution. Also, least performing workers increase their effort by 15 percentage points when their production lines become more homogeneous. These results hold for fixed wage framework where factories pay fixed daily wages to workers irrespective of their output and can be explained by pro social preferences.

JEL classification: Y40, Z13, J15, J24

Keywords: caste, team production, labor productivity, garment factory, India

---

[a] Indian Statistical Institute, Delhi

The Policy and Planning Research Unit (PPRU) at ISI, Delhi (Dr. Farzana Afridi) and the Centre for Competitive Advantage in the Global Economy (CAGE), Warwick University (Dr. Amrita Dhillon) provided financial support for this study. The usual disclaimers apply.

1. Introduction

In many industries production processes are organised into teams and final output is determined by the lowest level of individual effort in the team instead of the effort of single worker or addition of efforts of all workers. This type of production can be most commonly seen in assembly lines arrangements that are quite pervasive in labor intensive factories in the developing countries. It has been found that social incentives(such as performance/knowledge spillovers, social pressure, taste for discrimination, favouritism, prosocial preferences etc.) can help in explaining the mechanisms through which co-workers can affect each other's performance at the workplace (Falk and Achino (2006), Menzel (2014), Kato and Shu (2008), Mas and Moretti (2009), Lindquist et al. (2015), Hjort (2014), Bandiera et al.(2005, 2009)).

The role of ethnicity based social connections hasn't received much attention in the empirical analysis of team production[1]. This study is unique because it looks at Leontief production function in a real world setting in a uniform fixed daily wage framework. Majority of existing studies are based on individual performance with piece rates along with other financial incentives and additional variations in team incentives. However, in large labor intensive factories in the developing countries, workers are generally paid fixed daily wage because management feels that piece rate is not suitable for Leontief production process as leads to wastage of resources[2]. In this study focus is on garment factories where lowest performance amongst the workers in a production line determines that line's output[3].

Garment manufacturing contributes to 4% of India's GDP (APEC (2013-14)). India is world's second largest producer of textiles and clothing and is ranked at 6th place as an exporter of clothing. India has enjoyed advantage in manufacturing garments because of availability of huge labor supply at low cost and high installed capacity. However, in recent years, other South Asian economies like Bangladesh, Vietnam, Pakistan etc. have been giving fierce competition to India even though these economies are smaller in size with lesser capacity as compared to India (FICCI (2016)).

[1] Hjort (2014) looks at tribe based ethnicity of workers in a flower packing plant Kenya, Kato and Shu (2008) study rural and urban divide amongst workers in a Chinese textile factory, Bandiera et al. (2009) look at connections established among the workers and supervisors due to nationality on a fruit picking farm in UK.
[2] Also, piece rates involve lot of investment in monitoring processes which large firms are not willing to invest in.
[3] Each production line can looked as one team in a factory. Details of production process in the garment factory will be given in section 3.

Garment manufacturing is important for Indian economy not only because of its contribution to the exports and there by GDP, but, also because it generates employment for low skilled workers and creates forward and backward linkages within the broader textile sector. In India, this industry suffers from low and stagnant labor productivity (OECD Journal, 2009). This study is an attempt to understand worker behaviour for better organisation of workers to augment productivity without increasing labor costs. We exploit the possibility of social connections that can get formed at workplace among the workers and hence influence their work behaviour.

The studies providing micro econometric evidence on workers' behaviour, characteristics and productivity especially in Indian manufacturing sector have been very few (Bloom et al (2013)). This study tries to fill this gap by studying 1913 garment factory workers across two factories in the industrial hubs of National Capital Region (NCR), India.

We focus on caste based social connections as these are quite persistent in Indian context[4]. Social stratification based on caste is a unique and persistent feature of Indian society. Caste has always played an important role in the migration decisions from rural areas to urban factory hubs (Chandavarkar (1994), Holmstrom (1984)). In our sample almost 70% of the workers are migrants from the northern and eastern states of India.

The factories studied here are large scale organised firms and employ only trained workers for the stitching purposes. Visits to the residential hubs of these workers revealed that while it is important to have some social connections when they initially move to the city, ties formed at workplace amongst co-workers are also important at the later stages of their career. There can be many bases of these ties and we base our analysis on the possibility of workers feeling more connected to each other when to belong to same caste because social set up of Indian society is caste based.

In this study, focus is on all the workers employed in stitching department for 30 production days. Contribution of this study to the literature is that it looks at (1.) effects of social connections due to same caste that workers might exhibit at work place on worker productivity when workers of different caste work together given a fixed wage framework. These results can

[4] Sengupta and Sarkar (2012) find that the formation of social capital for each individual is higher among same (homogeneous) individuals along caste and religious dimensions in their social sphere. Munshi (2017) also notes that caste plays a role at every stage of an Indian's economic life.

be explained using the concept of pro-social preferences[5]; (2.) effects of caste composition of a production line on line's performance in absence of line output based incentives given Leontief production process which can be explained using social identity theory.

Data analysis suggests that the increase in proportion of workers belonging to own caste affects worker's own productivity positively and significantly. Controlling for individual heterogeneity among the workers within a line makes this result stronger. So, within a line, caste of the co-workers matter for an individual's performance. This result holds for line level analysis as well. Lines which are more concentrated in terms of caste composition of the workers are performing better validating positive effects of being in a homogenous group. For policy purposes it implies that putting together workers of similar background can incentivise workers to be more productive even in absence of monetary incentives. It should be emphasized that these estimates are upper bounds on impact of caste based preferences as this analysis is based on broader caste categories.

This paper is organized in the following manner. Section 2 discusses the related literature. Section 3 describes the context and the setting. Section 4 explains the theoretical framework. Section 5 describes and summarizes the data. Section 6 discusses methodology and reports regression results. Section 7 concludes.

2. Literature

There exists ample experimental literature that highlights that homogeneity in the teams leads to more efficient outcomes (Eckel & Grossman (2005), Charness, Rigotti & Rustichini (2007), Goette, Huffman & Meier (2006), Chen & Li (2009), Chen &Chen(2011)). The basis of these experiments come from the "social identity theory" that suggests that members of a team that is heterogeneous with respect to social categories may find it difficult to integrate their diverse backgrounds, values, and norms and work together (Jehn et al (1990)). Northcraft et al (1996) note, "…the discomfort or apprehension that individuals experience when interacting with members of a different social category is a natural consequence of social identification processes". In general, people feel more comfortable working with and are more likely to trust and cooperate with those whom they identify with, and they are more likely to identify with

[5] Pro social behaviour is described as altruistic behaviour whereby worker experiences disutility if she is acting non-cooperatively even if no one notices (Mas and Moretti, 2009)

members of their own characteristic group. Individuals do not deal with other individuals as individuals but as a member of some social group. However, the studies providing micro econometric evidence on workers' ethnicity, behaviour, and productivity have been very few especially in developing countries (Hjort (2014), Kato and Shu (2008), Menzel (2014)).

Eckel & Grossman(2005) show that strong team identity can augment team production by deterring shirking and free riding problem using a public good experiment framed as a team production problem. They further argue that if the maximum benefits are to be obtained from team production, it is imperative that distrust, lack of cooperation, and general unwillingness to work with others created by social category diversity be overcome. The more team members identify with one another, the more likely they are to believe that they hold similar goals, values, and norms, and the more willing they will be to cooperate and work together as a team. Chen and Chen (2011) find that making group identity salient leads to in-group coordination and to the efficient high-effort equilibrium in lab experiments.

Mas and Moretti (2009) look at workers working in supermarket chains in US and find strong positive spillover effects with the introduction of highly productive personnel into a shift. They conclude that worker effort is positively related to the productivity of the workers who see him but not workers who do not see him. Their results are driven by social pressure as workers care about how they are perceived by their fellow workers. However, social preferences can also drive these results when workers are working in a team.

Bandeira et al (2005) devise a model which shows that social preferences affect workers' effort choice under relative incentives but not under piece rates. They study 142 workers (10215 worker-field-days) in a UK based fruit farm for the season of 2002 who are paid relative incentive scheme for the first half of the season and piece rate for the other half. Under relative wage incentive worker's compensation depends on her productivity relative to the average productivity of her co-workers. Increasing her effort leads to increase in workers' payoff but causes a fall in payoff of other workers. Due to social preferences, workers partially internalizing the negative externality their effort imposes on others under relative incentives, especially when working alongside their friends. They find that the productivity of the average worker is at least 50 percent higher under piece rates than under relative incentives. They also show that workers

internalize the externality only when they can monitor others and be monitored. In this study social preferences are actually leading to fall in productivity. This is contrary to the Mas and Moretti(2009) result where productivity increased if worker knew that his efforts are observable but in their set up wages of workers were fixed.

There is another line of literature that looks at networks within teams and productivity. These networks can be based on being peers at workplace (Lindquest et al (2015)) or ethnicity (Hjort (2014), Kato and Shu (2008)). This study adds to literature on social connections in teams as it exploits exogenous caste variation in production line due to placement of workers by the management to meet the production process demand and tackle absenteeism. Lindquest et al (2015) define two workers as co-workers if they both come from the same team (team composition changes every week) and their working hours in a particular week overlap. Using number of hours worked together for a call centre employees, they create a matrix that describes links amongst co-workers for all the workers in the sample. They look at 1.) aggregate efforts of a worker's co-workers and describe this as local aggregate network effects; 2.) deviation from the average of group effort and describe this as local average network effects that reflects moral cost of not confirming to the social norm. They conclude that worker's current productivity is affected by her co-workers' current productivity through her desire to confirm to the social work norm (local average network effect) rather than strategic complementarities (local aggregate network effects). A 10% increase in current productivity of a worker's co-worker network leads to a 1.7% increase in own current productivity attributing it to the conformist behaviour.

Kato and Shu(2008) focus only on performance spillovers due to networks at workplace in a cloth manufacturing firm in China by looking at defective rates of the weavers. Using estimated individual weaver fixed effects they divide workers into high ability and low ability type. They argue that due to *hu-kou* system(urban housing registration requirement for migrant workers) divide between urban workers and rural migrants is stark and rural workers tend to identify themselves more with other rural workers and that serves as the source of knowledge transfer and affect the performance of workers. They find that there are strong performance spillovers from high ability workers to low ability workers but not vice versa. Also, the process of performance spillover/knowledge sharing takes place within the confines of social networks.

Similarly, Hjort (2014) looks at team production and social divides amongst different ethnic groups at a flower processing plant in Kenya where workers from two rival clans work in teams. A team has three workers, one supplier and two processors who can belong to same clan or different (vertically or horizontally). He finds that ethnic homogeneity can lead to higher team output as compared to heterogeneous teams because discriminating against members of other social group gives a worker higher utility even if it entails forgoing monetary benefits as individual payment depends on team output. Additionally, output gap between homogeneous and diverse teams doubles in a national situation of intense political conflict. Take away message from Kato and Shu (2008) and Hjort (2014) is that individual productivity is higher in ethnically homogeneous teams given the piece rate structure. We attempt to show the effect of team composition in a fixed wage structure.

To the best of our knowledge, this study is first attempt in India to look at caste based connections in a fixed wage framework and worker's productivity when they are working in a team. Even though analysis is based on garment factory production lines, it is applicable to situations where production process is organised into teams with fixed monthly salaries.

3. Context and setting

3.1. Context

Apparel or garment manufacturing is the part of the textile industries which is one of the oldest industries in the Indian economy. Even today textile sector is one of the largest contributors to India's exports with approximately 13 per cent of total exports in 2015-16. Readymade garments had a share of 47.7 per cent in these exports. Textile and garment sector is the second largest employment provider in the country employing nearly 51 million people directly and 68 million people indirectly in 2015-16. The textiles industry is also labour intensive and is one of the largest employers. The textile industry has two broad segments. First, the unorganised sector consists of handloom, handicrafts and sericulture, which are operated on a small scale through traditional tools and methods. The second is the organised sector consisting of spinning, apparel and garments segment which apply modern machinery and techniques such as economies of scale (IBEF, 2017).

This study focuses on modern factories in organised sector which produce garments for big well known brands and follow global methods of manufacturing garments. A garment factory

has multiple departments, e.g. training, cutting, sampling, production, packaging and finishing. Stitching of garments falls under production department and it is organised into production lines. Each stitching floor has multiple lines and in each line stitching machines are placed one after the other(refer fig A.1). Average size in the sampled production lines is 40 workers and they can be classified into operators (those who sit on sewing machine and are responsible for stitching); helpers (those who fold, cut and match different parts of garments so that operators can perform their operations smoothly); pressman (who are mainly helpers but perform specific operation of ironing pieces of garments). Every line has a supervisor who is responsible for smooth functioning of overall line and is held responsible for line level productivity by the management. His or her job is to discipline the workers, ensure that all the machines are working properly and there is enough stock of cut pieces of garments[6] with the workers. Line incharge/floor incharge lies above supervisor and is responsible for any issues across the lines on his floor. Floor incharge reports to the production mangers and the factory head. Factory head is held responsible for the overall functioning of the factory. Heads of different department also report to the factory head. Factories also have a Human Resource(HR) department which maintains attendance, salary records of the workers and organize activities amongst the workers.

Each operator is allotted a stitching machine and is responsible for performing at least one operation. Helpers and Pressman are allotted tables that are generally at the ends and middle of the production lines where they fold, mark, invert and press different parts of garments. Multiple workers can be doing similar operations on different bundles of garment pieces. Henceforth, the term worker will be used to denote operators, helpers and pressmen (i.e. those who contribute in stitching of the garment). To explain production we take an example of manufacturing of a shirt. Figure A.2 summarizes the general production process. There are two broad processes in which production is organized in the factories in the sample as described below.

1. Workers receive cut pieces of garments in bundles (one bundle consists of 20 pieces in the sampled factories) from the cutting department at the beginning of every hour. Helpers at the end of line mark points for stitching or do matching of patterns so that operators can readily stitch pieces. Production process begins from the end of the line and as one moves along the line different operations are performed on different parts of the garment. At the front section of the line, shirt is finished by performing operations that assemble different parts of the shirt. In the line, those responsible for making collars do different operations on the cuttings of collars (like

[6] Also known as WIP(Work in Progress) in garment factories.

collar run stitch, collar hemming), those responsible for sleeves do different operations on sleeves (like cuff attach, sleeve placket attach, armhole top stitch) and so on. On a typical production day there is enough "WIP" so that workers are not sitting idle in the line. Thus, each worker in a line contributes to the manufacturing of a shirt by performing assigned operation at some point of time in the process of production.

2. Generally in small scale factories, production of parts of garments is assigned to different lines. For example, one line will be responsible for producing only collars, other for only sleeves, another for only cuffs, and so on. These ready pieces are then sent to assembly lines, where final product is manufactured. In assembly lines, starting from the rear of the line different parts are attached and shirt is finished as it reaches front. There is enough WIP so that there is not an issue of undersupply of intermediate pieces of garments. Although management pointed out that initially when a new style is introduced in production lines it takes a few days for style to settle down.

It is to be noted that operators working on similar parts of garments tend to sit together and can observe each other's output. They might interact and be influenced by each other's productivity and attitude.

In the sample, workers and supervisors are employed on fixed monthly salary basis. So it doesn't make a difference to the workers' pay if he is able to achieve 50% of the target or 100% of the target. If performance is too abysmal, let's say worker is producing only 20% of the target, repeatedly; the supervisor tries to resolve the issue by understanding the nature of the problem faced by the worker and suggests the solution. However, if performance doesn't improve then the supervisor reports the incidents of slackness to the production manager and worker is given a warning. If worker doesn't improve even after 2-3 warnings, he/she is fired. Supervisor has no final power in hiring or firing the worker.

Workers are granted unpaid leaves if planned. Workers can also take sick leaves and casual leaves but for limited number of days. Their monthly fixed salaries can be looked upon as show up payoff. If they are absent then their daily salary is cut(monthly salary /30). Management has policies in place for minimizing absenteeism.

Given high demand of workers due to complementarities in the production process and high attrition rates in garment factory hubs, it is not difficult for workers to find a new job. Also, since fixed wage system doesn't reward high individual effort, workers have no incentive to be more productive than a minimum level which ensures continuation in the firm. In the sample

studied there is no bonus scheme in place. In such scenario, norms among co-workers stemming out of social identity might be important for worker's decisions regarding efforts to put in and thus for the factory's performance.

A line can be more homogeneous or heterogeneous in terms of caste composition as compared to other lines. As discussed earlier, social identity theory suggests that workers belonging to same caste as co-workers feel more connected and this tend to affect their performance. They can co-ordinate at the lower level of effort or high level depending upon the norm of the group. In this study an attempt is made to look at the caste composition of the lines and its effect on the productivity of the workers and lines.

Caste is a unique feature of Indian society and caste categorisation is enforced upon individuals at birth due to its hierarchical feature. Historically, caste system divided Hindu society into four *varnas- Brahmins* (priests and scholars) at the top, followed by *Kshatriyas* (warriors and rulers), *vaishyas* (merchant class), *shudras(engaged in land related occupations)*, respectively. There also existed another group-*ati-shudras* who were not assigned any varna and lived in the peripheries of the villages. These individuals were engaged in menial tasks of scavenging, sweeping and were considered untouchable. Over time caste system became hierarchal and stratification on the basis of caste became stricter. Each *varna* had sub groups known as *jati* (sub-caste) and lifestyles, skills, beliefs, customs, social networks and livelihood choices were shaped by the *jati* an individual was born in. Stratification led to discrimination across *varna* with individuals belonging to same *varna* identifying with each other strongly. This social inequality translated into economic inequality with *shudras* and *ati shudra* becoming the most oppressed section of society. Originally, the caste system was a feature of Hindu society, however, overtime it encompassed other religions that came up in India. (Sinha (1993))

After independence, policies aiming at the upliftment of the weaker sections were adopted (also known as affirmative action) and Indian government mapped sub-castes into four broad categories namely, Schedule caste, Schedule Tribes, Other backward classes and Unreserved. Schedule caste covers sub-castes that were considered untouchable/*ati-shudra*, schedule tribes are tribes from remote regions, other backward classes consist of sub-castes that came under *shudra*. Unreserved or general categories consist of all other sub castes. On the basis of this we divide the sample into three broad categories-Low caste=L (SC/ST), middle caste=M (sub-castes belonging to OBC categories) and High castes=H (sub-castes belonging to

Unreserved categories also known as Forward Castes). Caste system still continues to dominate an individual's choices and it can have implications when workers from different caste come together and work.

3.2. Quasi-random allocation of workers to lines

Typically, a garment style is fixed for a line and runs till its order is completed. Management explained that for a style in a line operators have fixed position and thereby operations, till a particular style runs in a particular line.

Each operation has a target associated to it. Supervisors try to discipline the workers by making them work according to the targets. Data suggest that targets are set at a very high level and often workers are unable to achieve it. Management justified high target as a step taken to push workers to work harder. Management is more concerned about the line level productivity and they want lines to achieve at least 50% of the set target. Since manufacturing a garment (shirt in this instance) requires all assigned operations to be performed on it, supervisor has to ensure that each operation is performed at least once. He shuffles workers within his line in case of absenteeism so that all the necessary operations are being performed. For example, suppose three operators are doing collar profiling but no one is doing sleeve attach then supervisor can shift one or two worker(s) from collar to sleeve position. Movement across lines happens only if worker performing that particular operation is absent and supervisor is not able to find anyone else from his line for performing that operation. However, supervisor does not control movement of workers across lines. Shuffling across the lines is controlled by floor in-charge or line in-charge. In case, supervisor is absent substitute supervisor is decided by the production manager. Workers' and supervisors' lines are typically fixed and are determined by the management when workers are hired depending upon the production requirements. So, variation in our data comes from unplanned absenteeism and attrition[7]. Lines differ in caste composition on the daily basis as shown by Figure A.3. This variation will help in identifying the effect of caste composition of the line on the worker's performance. Discussions with management revealed that assignment of the worker to the lines purely depend on supply and demand of labor force plus skill requirement for a particular style. According to the management while assigning workers to the production lines, caste of the workers is not a determining factor.

[7] Average weekly absenteeism is about 10% in the sample. Refer to Fig B.1

3.3. Exogeneity of caste and assignment to the lines

In this section we attempt to validate the claim made by the management about the independence of the caste and line assignment of the operators which will form the basis of the empirical analysis. We use Pearsons' chi square test with null hypothesis being that probability of a worker belonging to a caste and probability of a worker being assigned to a line are independent. From the data we have caste composition of each line on a day denoted by $P(Cw \cap Lw)$ where $P(Cw)$ is probability of a worker w belonging to caste category C, and $P(Lw)$ is probability of a worker w getting assigned to a line L. For exogeneity, $P(Cw \cap Lw)=P(Cw)*P(Lw)$ must hold for all lines on each production day. For the period of productivity data days, share of H is 0.43, M is 0.33 and L is 0.23 in the production lines[8]. We perform test for a random day in each week for both the factories. TableB.2 in Appendix B gives a snapshot of the caste distribution of workers in production lines on a random productive day for the exporting factor. Using this test we fail to reject the null hypothesis at 5% level of significance and thus conclude that the claims made by the management are justified. This implies that the caste composition of lines is exogenous.

4. Theoretical Framework

5. Data and Summary Statistics

5.1. Data

This study consists of two factories located in industrial hubs of the cities of Faridabad and Gurugram in the state of Haryana, India. The one located in Faridabad caters to exports (larger of the two) and the one located in Gurugram manufactures garments for domestic market. 90% of the sampled workers belong to the exporting firm. Final data set consists of data from multiple sources – survey of the workers on the stitching floor, worker attendance and demographic data from the HR department, productivity data from the production department. Details of these data sources are as follows.

5.1.1 Survey data

Timeline for the census survey for collecting data on demographics and worker's characteristics was from 5th August 2015 to 15th October 2015, covering a total of 1916 workers and 74

[8] In export factory 20.67% of workers belong to L category, 28.71 % belong to M category and 47.98% belong to H category. About 2.65% workers report that they don't know their caste. For domestic factory, 25.04% of the workers belong to L, 46.98% belong to M, 24.80% belong to H and 3.18 don't know their caste. Production lines are of different sizes with average size for export factory being 43 workers and domestic factory being 15 workers.

supervisors across both the factories. Since in such factories, workers join and leave at different point of times, surveys were carried out throughout the study period. All the workers on the stitching floors were interviewed[9]. The time period of the study falls under normal production time i.e. there is no shortage of labor as it is neither the harvesting season nor the festivity time in India. Also, it's not a peak demand season which is typically before the spring or the fall that causes rise in the demand of the labor. Personal interviews using questionnaires lasted for approximately 20 minutes. Workers' questionnaire had five sections. First section had questions on worker characteristics like age, education level, native place, caste, sub-caste etc., second section asked questions on worker's experience in garment industry, third section had questions on the process of obtaining current job and asked for details of referee if job was obtained through referrals, fourth section had questions on worker-supervisor relationship and finally, fifth section asked questions on relations with the co-workers.

Supervisors were also interviewed using a supervisor questionnaire. Data on their characteristics like age, caste, marital status, education etc., work experience, process of obtaining current job etc. was collected.

5.1.2 Factory records

Attendance data

Data on worker and supervisor daily attendance was obtained from the Human Resource Department of the factories from 1st August 2015 to 15th October 2015[10] that came about to be 61 working days. HRD is responsible for keeping a track of worker attendance data, sanctioning of leaves, emergency health issues related leaves etc. In the Faridabad factory, card punching system is used for attendance recording. In the Gurugram factory, workers are required to submit their cards with the HR representative who then enters these rolls in the computer system. Workers can leave only after collecting these cards at the gate enabling HRD to keep a track of half day leaves taken as well.

HR attendance records also contained information on types of leaves taken, half days and additional information on joining dates, date of leaving, designated line number, designation and

[9] For the questionnaire used for survey refer to appendix D.
[10] Attendance data was cleaned and coded as 0=absent, 1=present, 0.5=half day, dot/missing=not on roll.

contact number[11]. Contact numbers from this data set were used for recovering information on workers who left the factory during the sample period via telephonic interviews.

5.1.3. Productivity data

Data on hourly worker output, line level output and line composition that mapped worker to the operation within each line were obtained from the production department of the factories for the from 8th September to 15th October. Due to absenteeism and attrition total number of workers covered daily varied. Collectively this study covers 37 production lines, out of which 31 are assembly lines. Line working days vary from 18 to 31 for production lines. Throughout the sample period, a line catered to multiple styles across days. On any given day, different styles run in different production lines and new styles are introduced at different point of times in different lines leading lines to go into "setting" (i.e. getting adjusted to a new style). Since different lines were in setting at different point of time, it was not feasible to obtain a balanced sample for all the lines and we observe 1043 production line days in the study. Also, if some day feeder (responsible for collecting worker data) was absent then data was not recorded.

In the sampled factories, management uses efficiency as a measure of worker and line productivity which is described by the following formula.

Efficiency per worker =Daily output/Daily target per worker

Using the same concept for measuring worker's performance, I calculate daily efficiency of the worker and use it as the dependent variable in the regression analysis[12]. Since workers perform different types of operations which vary in terms of difficulty levels and thus time requirements, it is necessary to make output comparable across different operations. Every style-operation combination has a specific daily target associated with it. This target is set by the industrial engineer of the factory and is available in the data entry sheets used by the feeder and line level

[11] Workers' reported unique card numbers were cross checked using the HR data.

[12] The daily output and mapping of worker to the operation was hand written by the feeder of the line and this led to multiple errors. One was of writing wrong worker's card number (for example interchanging of digits or missing a digit while writing) or writing wrong card number against a worker's name (usually happened in vertical sequence such that names and card numbers interchanged). Other major issue was of not writing individual output in case multiple workers are sitting on an operation. These errors accounted for 8.18% of the productivity data. After getting line composition and output entered, these issues were addressed in the following manner. Using survey data as reference, card numbers were corrected for unique names. Assumption here is based on the observation that feeder tends to know the name of the workers correctly rather than the card number. For non unique names we looked at the productivity data of that operation for the sample period in that particular line. The mode card number for that worker name and operation in that particular line was then used. For the cases where individual break-up of the output was not given, we extrapolated by looking at the contribution of each worker in case of joint output for that operation whenever break up was given in the sample. Using this method we were able to recover 95.54% of the errors. (We were unable to correct card numbers for 0.45% of the productivity data.)

efficiency data files provided by the factory[13]. Summing output across designated working hours (i.e. 8 hours if worker doesn't take off at any point of the day) gave daily output per worker, and, dividing it by the specified style-operation target we get daily efficiency of the worker. Thus, efficiency equals to 1 if the worker produces according to the target[14].

I measure line level performance in two ways. One, by calculating the average efficiency of all the workers sitting in a line on a day. Two, by looking at least performing worker as lowest effort determines final output in assembly lines.

Production lines are organised across floors which function like mini-sub-factories. Floors can be an important factor in explaining differences in efficiency because lines on the floor with the HR department or Factory Head office may behave differently as compared to lines which are located on the floor far from the scrutiny of factory head or with weaker management.

After combining worker survey data, attendance data from HR records and normalization of efficiency data we obtain an unbalanced panel for 1916 workers. It is to be noted that there is a timeline difference in the collection of data from different sources. As collection of production data started a month later, our final worker analysis is based on an unbalanced panel 1796 workers with 35614 person days[15].

5.1.4. Social connections

Sub-caste or *jati* reported by the workers during personal interviews of the workers have been used to measure caste based social connections. *Jatis* belong to different *varna*s and thus

[13] I calculate the manpower sitting on an operation by counting the number of workers sitting on a particular style operation in a line for that day.

[14] After normalization it was observed that about 1.2% of person days had efficiency>1 (mapping into 149 workers). *t*-test shows that these 149 workers have significantly higher efficiency on other working days as well. A comparison of their characteristics, using a probit model, shows that these workers are more likely to be younger, less educated and more experienced. So we keep these observation in our analysis and approximate efficiency>1 to 1.

[15] For 1916 workers with 45948 person days, 11.18% are absent person days, 0.48% are half days, 80.96% are present person days and 7.37% are dots. Productivity data was recorded for only 47% of the half days, simply depending upon the discretion of the line feeder, therefore, we drop person days with data for half days and in this process we lose 2 workers. Out of remaining 1914 workers, 112 surveyed workers left before we started collecting the productivity data. 6 workers don't appear in the productivity data but appear in HR attendance records for 0.0003% of person days. Using a probit model for comparison of characteristics of these 118 attrited workers, we find that attrited workers are no different than the workers who were on roll during collection of the productivity data.

determine social status of a person in caste hierarchy. In democratic India, *varna* categorisation is not relevant for official purposes but administrative caste categories defined by the government are. Workers were asked about their sub-castes and caste categories. The issues with reported caste categories were twofold. One was that of missing information. About 34.3% of the workers didn't know their administrative caste categories and about 1.56% of sampled worker didn't know their sub caste. Other issue was of misreporting of caste categories. Both the issues were resolved by making use of the reported native state and district. Every state has list of sub-castes recognised as OBC and SC categories along with central list for ST categories. After standardizing and cleaning the spellings of the sub-castes, we used native state lists to map this variable into 4 caste categories-SC, ST, OBC and UR(unreserved). In the analysis, L category refers to workers belonging to SC/ST categories(21.49%), M refers to the worker belonging to OBC category(30.16%) and H refers to worker belonging to UR/Forward castes category(45.22%)[16]. Using this strategy, caste category of the 96.87% workers was recovered. Similar approach was followed for rectifying caste variable for the supervisors. Caste categories for those workers who didn't know their sub-caste or native states could not be recovered. Cleaned caste categories were then used for generating variables of interest which reflect the magnitude of caste based social connections and are as described below.

1**. Workers' own caste proportion in his/her line on a given day**: It is defined as number of workers belonging to workers' caste group as a proportion of total strength of the line worker belongs to on a day. Increase in own caste proportion means possibility of having more caste based social connections[17].

2. **Caste concentration index:** The concept is borrowed from HHI. It is sum of squares of caste proportions of the three caste categories in a line on a day. Higher caste concentration index for a line implies higher homogeneity in that line.

5.2. Summary Statistics

Table 1 summarizes the characteristics of the 1796 workers who appear in our final analysis. Similar table for the entire sample covered (1916) is given in Appendix B.1. Females form the major proportion of the workforce (84%), most likely to be a married Hindu with an

[16] Caste Composition of Indian population is 19.7% SCs, 8.5% STs, 41.1% OBCs, 30.8% Forward Castes. (Census 2011, GOI)

[17] Another way of putting it is that degree of caste based social connections increases with increase in own caste proportion

average age of 30 years. About 60% of the workers have studied till secondary level. About 40% of the workers are migrants from U.P. Other major migrant state is Bihar.

Table 2 summarizes worker characteristics by their caste categories. Workers belonging to high caste are older, more experienced and have higher level of education.

Table 3 summarizes worker level efficiency. Panel I shows average efficiency of a worker on a given day in a given line. Workers are achieving around 31% of their targets on an average. This number is quite low and can be because factories set targets at very high level and fixed wage framework doesn't incentivise workers to attain high efficiency levels. Panel II shows average efficiency of a worker across the sampled period. Worker efficiency is not statistically significantly different across caste categories and can be seen from figure 1 as well.

Table 4 summarizes line level efficiency. In export factory production process has no forward or backward linkages with other lines and every stitching operation related to a particular style is performed in the line itself. We used worker's data to calculate average performance of a line as explained in earlier section. Mean efficiency at line level is around 30%. Factory also provided records of daily line efficiency. Interestingly figures given by factory are much inflated(~47%).

Figure 2 gives proportions of caste categories over the sampled period. In domestic factory, average proportion of workers belonging to H category in production line is 0.27 over 30 production days, for M figure is 0.535 and for L it is 0.196. The two factories differ in caste composition with 0.49 being the average proportion of H type in production lines in export factory. Average share is 0.293 and 0.213 for M and L type, respectively. As discussed earlier, variation in lines comes due to absenteeism. Average weekly absenteeism at worker level is at 10% as shown in Table B.1. It also gives breakup of absenteeism by caste. Workers belonging to L category have highest rate of absenteeism but these differences are not statistically significant across caste[18]. Absenteeism data obtained from the factories is for 61 working days (August 2015 to October 2015) and typically a worker is on-roll in factory records for almost 53 days.

6. Empirical strategy and Results

6.1. Estimating equation

[18] Along with Pearson's chi square test for proving exogeneity we conduct regression analysis for absenteeism data using weekly absenteeism rate as dependent variable to see relationship amongst caste and absenteeism. Results are in appendix C.2.

Baseline specification exploits panel structure of worker productivity data and is given by equation (1).

$$Y_{itl} = \alpha_i + \beta Social\ Connection_{itl} + \gamma X_i + \epsilon_{it} \qquad (1)$$

where, $Y_{i\,tl}$ is measure of labor productivity (i.e. efficiency of i-th worker on t-th day in l-th line), $Social\ Connection_{itl}$ is proportion of same caste workers in the line "*l*"of *i*-th worker's on *t*-th day. It reflects the magnitude of possible caste based social connection a worker can have in a line on a given day. β is the main coefficient of interest which will help in testing the null hypothesis that caste based connections have no effect on worker's productivity. X_i are controls for worker characteristics (age, marital status, religion, native state, experience, education, number of reported friends etc.). Standard errors are clustered at line level. We subsequently add floor fixed effects/line fixed effects (to control for floor level/ line level unobservables) and time fixed effects(month/week fixed effects) while presenting the results.

Since this specification ignores unobserved heterogeneity amongst workers (such as ability) we use stricter specifications with individual worker fixed effects in equation 1.

For analyzing line level productivity, we use equation (2):

$$Y_{tl} = \alpha_l + \beta Social\ Connection_{tl} + \gamma X_l + \epsilon_{it} \qquad (2)$$

where, Y_{tl} is a measure of line level productivity (i.e. average efficiency of l-th line on t-th day), $Social\ Connection_{tl}$ is a measure of caste homogeneity in l-th line on t-th day making β as the main coefficient of interest, X_l are controls for line level characteristics (e.g mean age, proportion of married worker, proportion of female labor force, proportion of hindus, proportion of migrants from bihar, mean experience of the workers etc.). Due to absenteeism line composition changes and thus these mean characteristics vary daily for each line. Standard errors are clustered at factory line level. Similar to the worker analysis, we add line fixed effects (to control for line level unobservables) and time fixed effects (month/week fixed effects) while presenting the results.

6.1. Results

6.1.1. Social connections and worker's productivity

The results of the analysis using equation (1) are presented in Table 5. 1% increase in own caste proportion increases a worker's efficiency by 6.9%. This estimate is huge if translated into number of pieces an individual produces. Every day a worker is given 40-50 bundles of 20 pieces each and thus 6.9% implies 44-55 pieces more daily. Table 5 also shows that older

workers have higher efficiency. However, more experienced ones have lower efficiency. Workers with more educational qualifications have lower efficiency. We introduce different fixed effects from col(2)-(5). While results on worker characteristics are unaltered, the coefficient on own caste proportion becomes insignificant with line and week fixed effects[19]. It seems like even though caste composition is important across lines and days, it is not relevant within a line-day. Before moving to the line level analysis we look at equation (1) with worker fixed effects to control for unobserved heterogeneity amongst the workers(such as ability). Results are presented in table 6. Controlling for unobserved heterogeneity amongst worker we find that coefficient for own caste proportion increases by almost 3 percentage points. These results hold with line fixed effects as well as shown by column 5 of table 6.

6.2. Homogeneity in lines and line's performance

Before conducting regression analysis we plot mean line level efficiency[20](in line days) against sum of squares of caste categories shares(caste concentration index[21]) in Figure 3 which shows that more homogeneous lines have higher efficiency on an average.

Using equation (2) we get table 7 that gives marginal effects of increasing share of H and M on average line efficiency with respect to L(base category). Increasing shares of H vis-à-vis L category increases mean line level efficiency. Given the fact that H forms majority of workforce in a line throughout our sample, it implies that as a line becomes more homogeneous average its performance improves. These results disappear with line fixed effects and time fixed effects(given by month and week fixed effects) for M category. Interestingly, increase in proportion of females in a line affects mean line efficiency negatively[22][23].

Using caste concentration index for measuring caste networks in a line instead of just the shares of caste categories in equation (2), we get results shown in Table 8. In column (1), coefficient against caste concentration index (0.255) implies increase in caste homogeneity increases average performance of the line. However, this result disappears with line fixed effects. Results for gender and experience are similar to that of Table 7.

[19] Similar regressions are run using mean line level characteristics as well. Results remain unaffected.
[20] Mean line level efficiency is mean of worker's efficiency for a day in a line(=sum of efficiency of all the workers efficiency in a line on a day/strength of the line on that day)
[21] Caste concentration index= $\sum c_c^2$ i.e. sum of square of share of each caste category in a line on a day(in line days)
[22] Regression were run with many mean line level characteristics as discussed in section 5.2 but only few are reported in the tables.
[23] Using line efficiency reported by factory instead of the calculated efficiency doesn't affect the results.

Since factories care about finished garments and given the Leontief production function, we look at minimum efficiency(amongst workers) in a line on a day as an outcome variable for measuring line's productivity in Table 9. Column (1) shows that minimum efficiency (least performing worker) is not affected by caste composition of the line. However, when we add line fixed effects then caste composition of a line matters. So, even though lowest link is giving similar performance across lines irrespective of caste composition of the line, within a line across days caste composition matter(as shown by Col2-4).

7. Conclusion

Our analysis validates the implications of the social identity theory. Working with members of same caste increases worker's and line level productivity. In absence of performance based incentives, more homogeneous production lines perform better. This study used caste as defining characteristic for the possibility of social connections amongst the workers. Factories can build this feeling of connectedness through training programs, small courses or events thus increase their performance.

References:

APEC report. *" http://www.aepcindia.com/news/aepc-annual-report-2016"*

Bandiera, O., I. Barankay, and I. Rasul. 2005. "Social Preferences and the Response to Incentives: Evidence from Personnel Data." *Quarterly Journal of Economics* 120: 917-62.

Bandiera, Oriana,Iwan Barankay and Imran Rasul. 2009. "Social Connections and Incentives in the Workplace: Evidence from Personnel Data." Econometrica, Vol. 77, Issue No. 4, pp: 1047–1094.

Bloom, Nicholas, Benn Eifert, Aprajit Mahajan, David McKenzie and John Roberts. 2013. "Does Management Matter? Evidence from India." *The Quarterly Journal of Economics,* Vol. 128, Issue No. 1, pp: 1-51

Chandavarkar, Rajnarayan. 1994. "The origins of industrial capitalism in India: Business strategies and the working classes in Bombay 1900-1940." Cambridge University Press.

Charness, Gary, Luca Rigotti and Aldo Rustichini. 2007. “Individual behavior and group membership.” American Economic Review, Vol. 97, Issue No. 4, pp: 1340–1352.

Chen, Yan and Roy Chen.2011. “The potential of social identity for equilibrium selection.” American Economic Review, Vol. 101, Issue No.6, pp: 2562–2589.

Chen, Yan and Sherry Xin Li. 2009. “Group identity and social preferences,” American Economic Review, Vol. 99, Issue No.1, pp: 431-457.

Dustman, Christian, Thomas Cornelissen and Uta Schönberg. 2013. “Effects in the Workplace.” IZA DP No. 7617

Eckel, Catherine and Philip Grossman. 2005. “Managing diversity by creating team identity.” Journal of Economic Behavior & Organization, Vol. 58, Issue No.3, pp: 371–392.

FICCI. 2016. “Knowledge Paper on Global Shifts in Textile Industry & India’s Position.” TAG 2016, Mumbai, India.

Goette, Huffman & Meier(2006) Goette, Lorenz, David Huffman and Stephan Meier. 2006. “The Impact of Group Membership on Cooperation and Norm Enforcement: Evidence Using Random Assignment to Real Social Groups,” American Economic Review,Vol.96, Issue No.2, pp: 212-216.

Gupta, Dipankar. 1992. “Social Stratification.” Oxford India Paperbacks, Oxford University Press.

Hjort ,Jonas. 2014. “Ethnic Divisions and Production in firms.” The Quarterly Journal of Economics, Vol. 129, Issue No. 4, pp: 1899–1946

Holmstrom, Mark. 1984. “Industry and Inequality: The Social Anthropology of Indian Labour.” by Cambridge University Press

IBEF. 2017. *“https://www.ibef.org/industry/textiles.aspx”*

Jehn, Karen A., Gregory B. Northcraft and Margaret A. Neale. 1999. “Why differences make a difference: a field study of diversity, conflict, and performance in workgroups.” Administrative Science Quarterly , Vol. 44, Issue No. 4, pp:741–763.

Sengupta, Jhumur and Debnarayan Sarkar .2012 "Caste and religious diversity on formation of social capital: A field survey in India and measurement issues", Equality, Diversity and Inclusion: An International Journal, Vol. 31 Issue: 2, pp.158-175.

Kato, Takao and Shu Pian. 2008. “Performance Spillovers and Social Network in the Workplace: Evidence from Rural and Urban Weavers in a Chinese Textile Firm”. IZA Working Paper No. 3340.

Lindquist, Matthew, Jan Sauermann and Yves Zenou. 2015. “ Network effects on worker productivity.” Working paper.

Mas, A. and E. Moretti. 2009. "Peers at Work." American Economic Review, Vol. 99, Issue No. 1, pp: 112-45.

Menzel, Andreas. 2016. “Social Networks and Productivity Spill-over within Firms: Evidence from Bangladeshi Garment Factories.” Working paper.

Sinha, Surajit Chandra. 1993. “Anthropology of Weaker Sections.” Concept Publishing Company, New Delhi.

Table 1: Worker characteristics

Variable	N	Mean	Std error	Min	Max
Age	1796	29.56	0.162	17	53
Female	1796	0.845	0.008	0	1
Hindu	1796	0.928	0.006	0	1
Married	1796	0.751	0.010	0	1
Education					
Upto Primary	1795	0.227	0.010	0	1
Upto Secondary	1795	0.604	0.012	0	1
Upto Senior secondary	1795	0.135	0.008	0	1
College and above	1795	0.035	0.004	0	1
Migrant Status					
From U.P	1792	0.403	0.012	0	1
From Bihar	1792	0.259	0.010	0	1
Caste Distribution					
L(SC/ST)	1744	0.220	0.001	0	1
M(OBC)	1744	0.311	0.011	0	1
H(Forward Castes)	1744	0.468	0.012	0	1
Workers' network					
Experience in garment manufacturing (in years)	1796	3.568	0.091	0	24
Received informal information on job opening	1796	0.746	0.010	0	1
Obtained this job through referral[@]	444	0.419	0.023	0	1
No. of friends	1796	1.748	0.033	0	7
Mean Work Days(in days)	1796	19.83	0.182	1	31

Note1: Couldn't map sub-caste of 2.9% workers into administrative categories. Note2: @ is conditional on referee being still employed in the factory.
Source: Factory worker's survey data, aug 2015-oct 2015

Table 2: Characteristics by Caste of the workers

Variable	Low Caste	Middle Caste	High Caste
N (1744)	**384(21.38%)**	**543(30.23%)**	**817(45.49%)**
Age (in years)	28.13	29.5***	30.43***
	(0.336)	(0.305)	(0.871)
Female	0.813	0.823	0.885
	(0.020)	(0.016)	(0.011)
Hindu	0.982	0.890***	0.935***
	(0.007)	(0.135)	(0.009)
Married	0.695	0.757**	0.785***
	(0.024)	(0.018)	(0.014)
Educated Secondary & above	0.151	0.158	0.186
	(0.018)	(0.016)	(0.014)
Migrant Status			
From U.P	0.982	0.890***	0.935
	(0.007)	(0.013)	(0.009)
From Bihar	0.156	0.322***	0.277***
	(0.019)	(0.020)	(0.016)
Workers' network			
Experience in garment manufacturing	3.09	3.497	3.854***
(in years)	(0.178)	(0.170)	(0.137)
Received information on job opening	0.794	0.753	0.717**
	(0.021)	(0.019)	(0.016)
Obtained this job through referral	0.347	0.451	0.435
	(0.049)	(0.042)	(0.036)
No. of friends	1.818	1.772	1.714
	(0.073)	(0.062)	(0.048)
Mean Work Days(in days)	19.802	20.116	19.723
	(0.405)	(0.322)	(0.270)

Note: L is the benchmark category for t-test. Significant at *10%,**5% and ***1%.
Standard Error in parenthesis. Caste category couldn't be recovered for 2.9% (52 workers) of the workers with efficiency data. *t*-tests for difference in efficiency and working days for these 52 workers against other give insignificant results. Though, probit model shows that these workers are more likely to be Hindus, from Bihar and older. Source: Factory worker survey data, Sep-Oct 2015.

Table 5: Own caste proportion and Workers' efficiency(in person days)

Variables	Efficiency	Efficiency	Efficiency	Efficiency	Efficiency
	(1)	(2)	(3)	(4)	(5)
Own caste proportion	0.069**	0.056**	0.054**	0.022	0.018
	(0.029)	(0.024)	(0.024)	(0.023)	(0.023)
Age	0.002**	0.002**	0.002**	0.002**	0.002**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Married	-0.013	-0.006	-0.006	0.000	0.000
	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)
Female	0.008	-0.0125	-0.0126	-0.0103	-0.0102
	(0.024)	(0.021)	(0.021)	(0.019)	(0.019)
Education	-0.015**	-0.014**	-0.014**	-0.013**	-0.013*
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
From Bihar	-0.003	0.012	0.013	0.016	0.017
	(0.013)	(0.012)	(0.012)	(0.012)	(0.012)
Hindu	0.000	0.002	0.002	-0.001	-0.002
	(0.016)	(0.016)	(0.016)	(0.017)	(0.017)
Experience	-0.006***	-0.009***	-0.008***	-0.009***	-0.009***
(in years)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
Received job information	-0.010	-0.009	-0.009	-0.011	-0.011
(through informal source)	(0.015)	(0.014)	(0.014)	(0.013)	(0.013)
No. of co-workers	0.002	-0.000	-0.001	0.001	0.001
Reported as friends	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)
Constant	0.283***	0.305***	0.281***	0.310***	0.278***
	(0.038)	(0.036)	(0.037)	(0.032)	(0.033)
Floor fixed effects	No	Yes	Yes	No	No
Month fixed effects	No	No	Yes	No	Yes
Line fixed effects	No	No	No	No	Yes
N	34255	34255	34255	34255	34255
R^2	0.012	0.042	0.056	0.080	0.094

Note: Own caste proportion=number of workers belonging to the caste category of the workers to the strength of the line. Standard errors in parentheses clustered at factory line level. Significant at *10%,**5% and ***1%.

Table 6: Own caste proportion and Workers' efficiency (person days)

Variables	**Efficiency**	**Efficiency**	**Efficiency**	**Efficiency**	**Efficiency**
	(1)	(2)	(3)	(4)	(5)
Own caste proportion	0.0996**	0.0997**	0.0803**	0.0909**	0.0724*
	(0.0443)	(0.0431)	(0.0375)	(0.0425)	(0.0374)
Constant	0.269***	0.191**	0.168*	0.244***	0.173**
	(0.0174)	(0.0839)	(0.0893)	(0.0738)	(0.0810)
Floor fixed effects	No	Yes	Yes	No	No
Month fixed effects	No	No	Yes	No	Yes
Line fixed effects	No	No	No	No	Yes
N	34255	34255	34255	34255	34255
R-square	0.545	0.546	0.558	0.550	0.563

Note: Regression with individual fixed effects. Own caste proportion=number of workers belonging to the caste category of the workers to the strength of the line. Standard errors in parentheses clustered at factory line level. Significant at *10%,**5% and ***1%.

Table 7: Caste proportion and Line efficiency (line days)

Variables	**Line Efficiency**	**Line Efficiency**	**Line Efficiency**	**Line Efficiency**
	(1)	(2)	(3)	(4)
Share of H	0.212*	0.143	0.218*	0.209*
	(0.106)	(0.130)	(0.123)	(0.118)
Share of M	0.185*	0.062	0.081	0.074
	(0.099)	(0.115)	(0.102)	(0.098)
Share of females	-0.131*	-0.324***	-0.363***	-0.376***
	(0.069)	(0.112)	(0.101)	(0.101)
Mean experience	0.0082	-0.027**	-0.028**	-0.027**
	(0.009)	(0.012)	(0.011)	(0.011)
Received Job	-0.139**	-0.159	-0.101	-0.0749
Information informally	(0.0540)	(0.120)	(0.104)	(0.103)
Constant	-0.065	0.365	0.257	0.262
	(0.174)	(0.293)	(0.258)	(0.252)
Line fixed effects	No	Yes	Yes	Yes
Month fixed effects	No	No	Yes	No
Week fixed effects	No	No	No	Yes
N	1043	1043	1043	1043
R-square	0.210	0.442	0.509	0.520

Note: Regressions are run with additional line level observables such as mean age, proportion of hindus, proportion of migrants from bihar, proportion of married workers, mean number of reported friends etc. Share of H and M are with respect to the share of L. Standard errors in parentheses clustered at factory line level. Significant at *10%,**5% and ***1%.

Table 8: Caste concentration index and Line efficiency (line days)

Variables	**Line Efficiency**	**Line Efficiency**	**Line Efficiency**	**Line Efficiency**
	(1)	(2)	(3)	(4)
Caste Concentration Index	0.255***	0.130	0.065	0.053
	(0.063)	(0.145)	(0.126)	(0.125)
Share of females	-0.108	-0.314***	-0.333***	-0.347***
	(0.071)	(0.110)	(0.099)	(0.099)
Mean experience	0.011	-0.024*	-0.026**	-0.025*
	(0.009)	(0.013)	(0.013)	(0.013)
Received Job Information Informally	-0.126**	-0.180	-0.110	-0.0827
	(0.0550)	(0.115)	(0.103)	(0.102)
Constant	0.026	0.421	0.335	0.337
	(0.159)	(0.296)	(0.262)	(0.255)
Line fixed effects	No	Yes	Yes	Yes
Month fixed effects	No	No	Yes	No
Week fixed effects	No	No	No	Yes
N	1043	1043	1043	1043
R-square	0.221	0.442	0.502	0.514

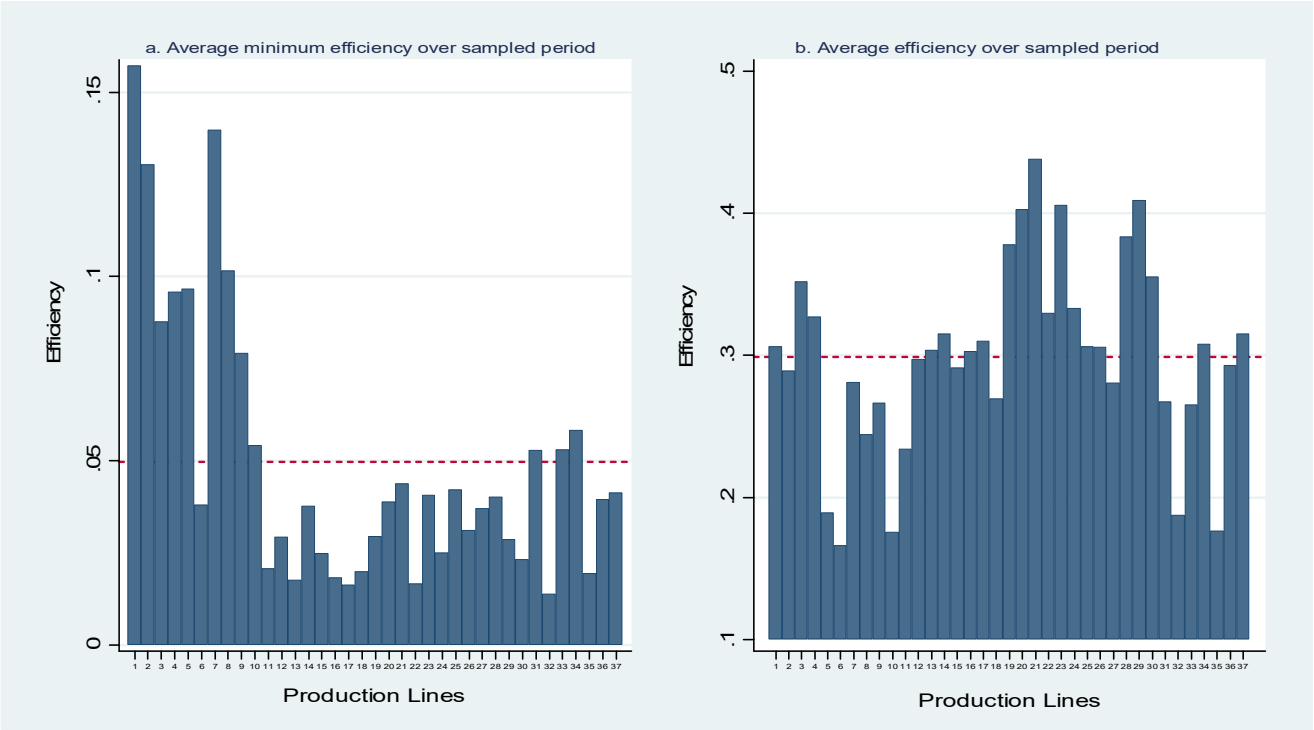
Note: Regressions are run with additional line level observables such as mean age, proportion of hindus, proportion of migrants from bihar, proportion of married workers,mean number of reported friends etc. Caste concentration Index== $\sum c_c^2$ i.e. sum of square of share of each caste category in a line on a day(in line days). Standard errors in parentheses clustered at factory line level. Significant at *10%,**5% and ***1%.

Table 9: Caste concentration index and Line performance

Variables	**Minimum Efficiency**	**Minimum Efficiency**	**Minimum Efficiency**	**Minimum Efficiency**
	(1)	(2)	(3)	(4)
Caste Concentration Index	0.026	0.087**	0.070**	0.065**
	(0.038)	(0.033)	(0.031)	(0.031)
Share of females	-0.019	-0.002	-0.002	0.001
	(0.045)	(0.061)	(0.059)	(0.058)
Mean experience	-0.003	-0.004	-0.004	-0.004
	(0.003)	(0.004)	(0.004)	(0.004)
Received Job Information Informally	-0.071**	-0.129**	-0.113**	-0.103**
	(0.033)	(0.048)	(0.043)	(0.042)
Constant	0.107	0.228**	0.202**	0.189**
	(0.071)	(0.087)	(0.076)	(0.074)
Line fixed effects	No	Yes	Yes	Yes
Month fixed effects	No	No	Yes	No
Week fixed effects	No	No	No	Yes
N	1043	1043	1043	1043
R-square	0.439	0.625	0.644	0.648

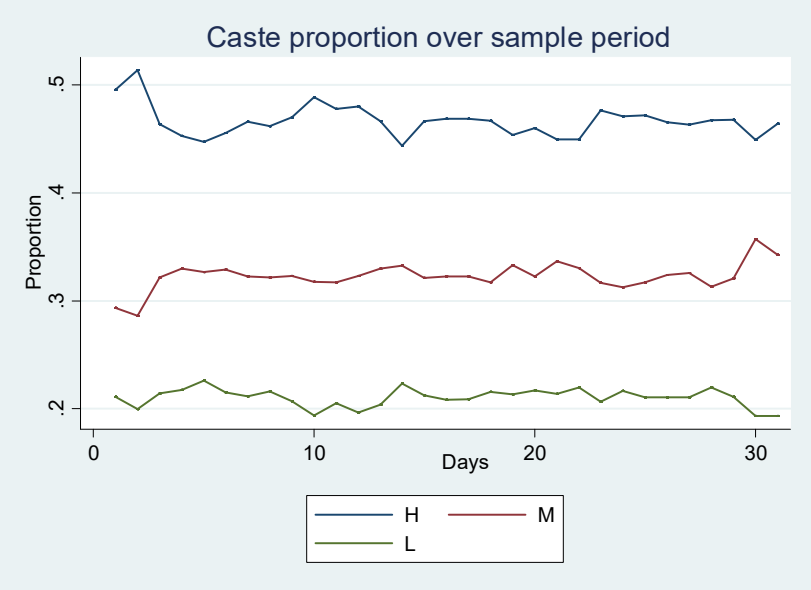
Notes:as elucidated above

Figure 1: Variation in average line performance over sampled days:



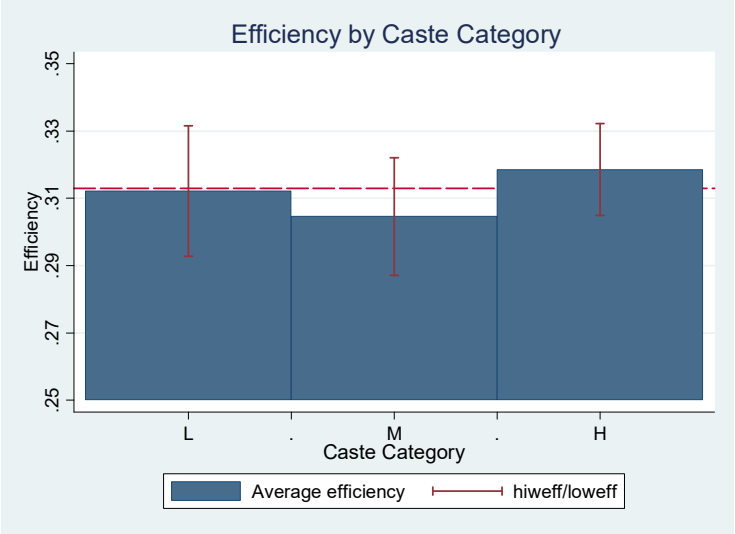
Note: Number of production lines=37. Graph (a) shows mean of daily minimum worker efficiency of the lines over sampled days. Average minimum line efficiency over sample period is 0.0497(given by dashed red line). Graph (b) shows mean of daily average worker efficiency of the lines over sampled days. Average line efficiency over sample period is 0.299(given by dashed red line). Working days for lines vary from 18 to 31.

Figure 2: Variation in Caste proportion over sampled days



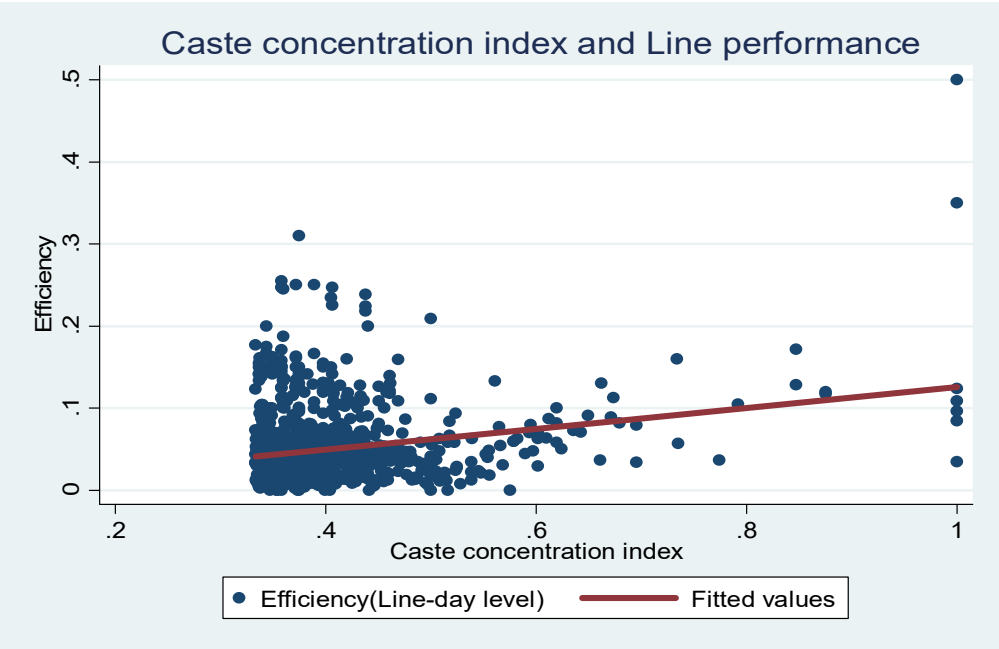
Note: Variation is coming from absenteeism data. Average daily absenteeism is approximately around 10%. Source: Factory production data from Sep-oct 2015.

Figure 3: Worker efficiency by caste category



Source: No differences in efficiency by caste. Worker's mean efficiency is 0.31. Worker productivity data, sep 2015-oct 2015 (30 working days)

Figure 4: Caste composition and Line level productivity



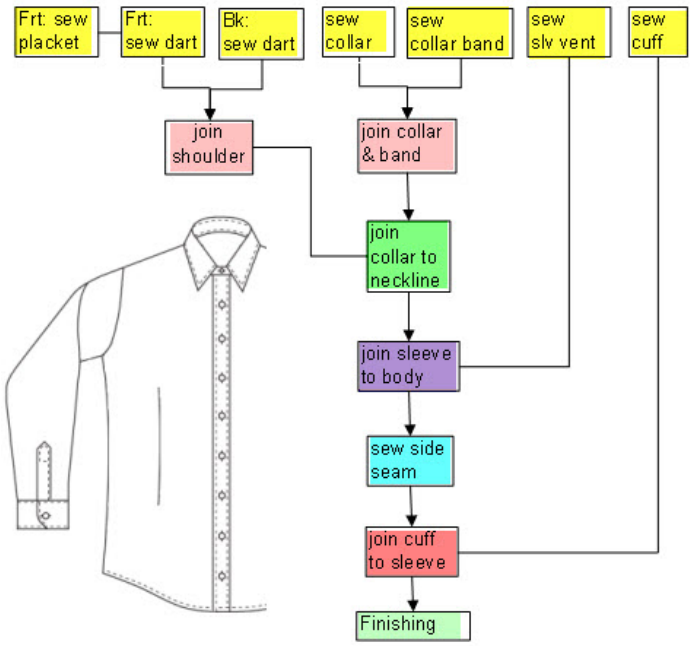
Note: No. of lines=37, line days=1043. Line days vary from 18 to 31 per line. Caste concentration index=$\sum c_i^2$ i.e. sum of square of share of each caste group in an assembly line on a day. Line Efficiency is the minimum efficiency among workers on a line-day.Source: Factory productivity data for efficiency for September-october 2015.

Appendix A

Figure A.1: Factory floor organisation

Location:Faridabad

Figure A 2: Manufacturing process of a shirt



Source:https://www.pinterest.co.uk/neelamparveen78/garment-production-manufacturing

Appendix B

Table B.1: Worker characteristics

Variable	N	Mean	Std error	Min	Max
Age	1916	29.44	0.157	17	53
Female	1916	0.848	0.008	0	1
Hindu	1916	0.928	0.006	0	1
Married	1916	0.749	0.010	0	1
Education					
Upto Primary	1915	0.225	0.010	0	1
Upto Secondary	1915	0.606	0.011	0	1
Upto Senior secondary	1915	0.135	0.008	0	1
College and above	1915	0.034	0.004	0	1
Migrant Status					
From U.P	1912	0.403	0.011	0	1
From Bihar	1912	0.259	0.010	0	1
Caste Network					
Low caste	1857	0.222	0.010	0	1
Middle caste	1857	0.308	0.010	0	1
High caste	1857	0.470	0.012	0	1
Workers' network					
Experience in garment manufacturing (in years)	1916	3.498	0.087	0	24
Received informal information on job opening	1916	0.743	0.010	0	1
Obtained this job through referral[@]	469	0.422	0.023	0	1
No. of friends	1916	1.735	0.032	0	7

Note1: Couldn't map sub-caste of 3.08% workers into administrative categories. Note2: @ is conditional on referee being still employed in the factory.
Source: Factory worker's survey data, aug 2015-oct 2015

Table B.2: Pearsons' chi square test

Caste	L1	L2	L3	L4	L5	L6	L7	L8	L9	L10	L11	L12	L13	L14	L15	L16	L17	L18	L19	L20	L21	L22	L23	L24	L25	Total
L	4	10	6	4	13	7	3	6	12	9	10	10	13	4	9	11	18	11	7	12	10	10	5	7	15	226
	4.6	10	9.6	8.6	7.6	7.8	8.2	9.2	9.4	11	9	10.4	11.6	4.8	13	11.8	11.6	9.6	5	9.2	9.8	9.8	9.6	4.6	9.4	226
	0.1	0	1.4	2.5	3.8	0.1	3.3	1.1	0.7	0.4	0.1	0	0.2	0.1	1.3	0.1	3.5	0.2	0.8	0.8	0	0	2.2	1.2	3.3	27.1
M	7	12	13	10	10	8	10	12	12	17	16	15	14	10	21	14	14	14	5	16	9	14	17	9	13	312
	6.4	13.9	13.3	11.9	10.5	10.8	11.4	12.7	13	15.2	12.5	14.4	16.1	6.7	18	16.3	16.1	13.3	6.9	12.7	13.6	13.6	13.3	6.4	13	312
	0.1	0.2	0	0.3	0	0.7	0.2	0	0.1	0.2	1	0	0.3	1.7	0.5	0.3	0.3	0	0.5	0.8	1.5	0	1	1.1	0	11
H	11	26	29	28	15	21	27	27	22	28	19	26	30	10	34	29	26	23	10	17	29	24	25	7	19	562
	11.5	25	24	21.5	19	19.5	20.5	23	23.5	27.5	22.5	26	28.9	12	32.4	29.4	28.9	24	12.5	23	24.5	24.5	24	11.5	23.5	562
	0	0	1.1	2	0.8	0.1	2.1	0.7	0.1	0	0.5	0	0	0.3	0.1	0	0.3	0	0.5	1.5	0.8	0	0	1.7	0.8	13.8
DK	1	2	0	1	0	3	1	1	1	1	0	1	1	0	1	5	0	0	3	1	1	1	1	0	0	26
	0.5	1.2	1.1	1	0.9	0.9	0.9	1.1	1.1	1.3	1	1.2	1.3	0.6	1.5	1.4	1.3	1.1	0.6	1.1	1.1	1.1	1.1	0.5	1.1	26
	0.4	0.6	1.1	0	0.9	4.9	0	0	0	0.1	1	0	0.1	0.6	0.2	9.7	1.3	1.1	10.2	0	0	0	0	0.5	1.1	33.9
Total	23	50	48	43	38	39	41	46	47	55	45	52	58	24	65	59	58	48	25	46	49	49	48	23	47	1,126
	23	50	48	43	38	39	41	46	47	55	45	52	58	24	65	59	58	48	25	46	49	49	48	23	47	1,126.00
	0.6	0.9	3.5	4.8	5.5	5.8	5.6	1.9	0.9	0.6	2.7	0.1	0.5	2.7	2	10.1	5.4	1.4	12	3.2	2.4	0	3.3	4.6	5.2	85.8

Note1: Production line number L1- L25 in 1st row. Caste categories in 1st column. 1st row against L shows actual frequencies of L across different lines. 2nd row shows expected frequencies under the null hypothesis of independence of probability of caste assignment and line assignment. 3rd row shows contribution of Pearson's chi square in each of the cell. Pearson's Chi square statistics is 85.7973 with 72 degrees of freedom with p value= 0.127. We can't reject the null hypothesis.
Note 2: Snapshot of a Friday in the month of September for the exporting factory.

Appendix C

Table C.1: Worker's average weekly absenteeism

	N	mean	se(mean)	min	max
Full sample	1910	0.101	0.003	0	0.842
L	403	0.105	0.006	0	0.761
M	563	0.097	0.005	0	0.842
H	854	0.099	0.004	0	0.770

Note 1: About 3.5% workers didn't know their sub-caste/native states
Note 2: t-test shows that there is no significant difference in average weekly absenteeism rate by caste categories.

Table C.2. Regression using absenteeism data

	Absenteeism	Absenteeism
Variables	(1)	(2)
Same caste	0.00144	-0.00188
as the supervisor	(0.00686)	(0.00717)
High Caste(H)	-0.00245	-0.00170
	(0.00583)	(0.00639)
Low Caste(L)	-0.00792	-0.00483
	(0.00684)	(0.00685)
Age	-0.00262***	-0.00257***
	(0.000533)	(0.000533)
Married	0.0255***	0.0254***
	(0.00675)	(0.00727)
Female	-0.00604	-0.00678
	(0.00824)	(0.00876)
Education	-0.00410	-0.00431
	(0.00391)	(0.00402)
Bihar	-0.0124*	-0.0107
	(0.00616)	(0.00679)
Hindu	-0.0295**	-0.0280**
	(0.0120)	(0.0125)
Experience	0.00118	0.00129
	(0.000825)	(0.000911)
Informally received	-0.00847	-0.00678
job information	(0.00892)	(0.00959)
No. of co-workers	-0.000862	-0.00194
reported as friends	(0.00205)	(0.00187)
Constant	0.213***	0.210***
	(0.0178)	(0.0151)

Line fixed effects	No	Yes
N	1835	1835
R-Square	0.0227	0.0405

Note: Absenteeism=average weekly absenteeism of worker i. “Same caste” is a dummy variable that takes value 1 if caste of the line supervisor is same as the worker given by(modal supervisor has been used). Standard errors in parentheses clustered at reported factory line level. Significant at *10%,**5% and ***1%.