WARWICK

ECONOMICS

Do workers, managers, and stations matter for effective policing? A decomposition of productivity into three dimensions of unobserved heterogeneity

Amit Chaudhary

October 2021 No: 1377

Warwick Economics Research Papers

ISSN 2059-4283 (online)

ISSN 0083-7350 (print)

Do workers, managers, and stations matter for effective policing? A decomposition of productivity into three dimensions of unobserved heterogeneity.

Amit Chaudhary*

October 18, 2021

Abstract

Misallocation of resources in an economy makes firms less productive. I document the roles of heterogeneity, sorting, and complementarity in a framework where workers, managers, and firms interact to shape productivity. The approach I follow uses the movement of workers and managers across firms to identify the distribution of productivity. I webscraped novel microdata of crime reports from the Indian police department and combined them with the worker-level measurement of productivity. Using this data I show that the third source of heterogeneity in the form of manager ability is an important driver of differences in firm productivity. I empirically identify complementarities between workers, managers, and firms using my estimation methodology. Counterfactual results show that reallocating workers by applying a positive assortative sorting rule can increase police department productivity by 10%.

JEL Codes: C13, C23, D73, H11, J62, M50

*University of Warwick, Department of Economics (email: a.chaudhary.1@warwick.ac.uk website: `https://www.amit-chaudhary.com/`)
I am grateful to James Fenske, Thijs Van Rens and Roland Rathelot for supervision. I also thank Luis Candelaria, Camilla Roncoroni, Kenichi Nagasawa, Manuel Bagues, and Dan Dan Bernhardt as well as seminar participants at the Centre for Competitive Advantage in the Global Economics (CAGE) conference.

1 Introduction

A central question in economics is what makes some firms more productive than others. Past literature has shown that misallocation of factors of production can account for productivity differences across firms. Therefore, the aggregate productivity of the economy can be increased by reallocating the resources across firms (Banerjee and Duflo, 2005; Hsieh and Klenow, 2009). Misallocation studies have identified underlying sources of misallocation such as regulation, market imperfections, and even government corruption (Restuccia and Rogerson, 2017). However the studies often assumes homogeneous production function across firms, and so the importance of effects of heterogeneity in firms and workers for aggregate productivity remains unanswered. Worker and firm heterogeneity can shape the wage and productivity distribution (Bonhomme et al., 2019; Abowd et al., 1999), and uncovering heterogeneity can reveal the importance of sorting and complementarities for workers and firms. Another significant branch of literature (Lazear et al., 2015; Bloom and Van Reenen, 2007; Bloom et al., 2013) shows that managerial quality may partly explain the productivity gap across firms. In an economy, workers, managers, and firms interact simultaneously, so attempts to explain the productivity gap through two-sided heterogeneity between workers and firms or between managers and firms is inconclusive.

In this paper, I estimate the role of heterogeneity of workers, managers, and firms in police productivity. Most importantly, I document the complementarities and patterns of sorting among workers, managers, and firms. The empirical analysis relies on the quality and extent of the data I use. I webscrape novel crime reports data [1] from the Indian police department to create a matched database of employment histories of both workers (officers) and managers (station head officers). This data allows me to track the job movements of workers and managers across police stations. In a separate web scraping exercise, I match the outcomes of the half-million crime cases to construct a measure of productivity. I use the time taken to submit the final report or charge sheet in criminal cases as a productivity measure. Then,

[1] Half a million crime reports were web scraped for the Indian police department

I use an employee-manager-firm data set linked to this productivity measure to identify how workers and managers working across police stations contribute to productivity.

To empirically estimate workers', managers', and police stations' contributions to police department productivity, I model the production function without any parametric assumptions. I extend the standard model of Abowd et al. (1999, 2002) (henceforth AKM) and Bonhomme et al. (2019) where only workers and firms contribute to productivity by adding a third source of heterogeneity in the form of manager ability. Thus, my approach adds managers and the interaction among workers, managers, and establishments to the extant model. Therefore, in this paper's three-sided model of productivity, heterogeneity comes from workers (investigation officer), managers (station head officer), and firms (police station).

I also model the complementarities between workers, managers, and firms as unrestricted rather than additive, as described in the fixed-effects literature (AKM models). I assume that heterogeneity in the economy can be represented by discrete types of workers, managers, and firms. The identification of the individual contribution of workers and managers is a challenging problem even with microdata. However, since I have employee-manager-firm matched data, I use workers' job mobility and manager types across police stations to infer individual contributions. I represent the job transition of workers and managers using a first-order Markov chain process. The model is estimated using a two-step approach where, in the first step, I map managers and police stations to discrete classes representing quality. The first step is a dimension reduction technique, and I use the classification algorithm of k-means clustering to map the individual managers and firms to discrete types. The second step uses the estimated manager and firm classes from the first step as input to estimate the individual effect of workers, managers, and firms. The second step estimates the model parameters using a finite mixture model, where a specific distribution of productivity is realised based on workers moving between manager and firm classes in a short panel. Using a grid computation technique, I estimate the model using the conditional Expectation-

Maximixation (EM) algorithm to converge to the solution (Meng and Rubin, 1993).

The estimation methodology adopted in the paper has numerous advantages. Firstly, the literature on managerial quality frequently uses the fixed-effects model popularised by Abowd et al. (1999) and finds that the best managers are allocated to the least productive workplaces or i.e. there is negative assortative matching of managers with firms. Due to this, evidence on the presence of complementarity are inconclusive as Becker (1973) shows that sign of sorting should be positive if complementarities are present and more recently Shimer and Smith (2000) and Eeckhout and Kircher (2011) emphasise the importance of sorting of workers in the efficient production of output . Thus, the AKM model is restrictive and often gives results that are not reconcilable with theoretical models. In addition to the above issue, results obtained from the additive model of worker, manager, and firm productivity can produce erroneous results in the counterfactual analysis when the researcher aims to determine which sorting pattern of workers can maximise the aggregate productivity. In this study, I keep the interaction unrestricted in a three-sided model of workers, managers, and firms. So, I can estimate the match-specific contribution or the complementarities arising from the worker, manager, and firm heterogeneity on the productivity gap.

Secondly, the model's assumption of finite classes of managers and police stations reduces the problem of limited mobility bias. In the AKM class of models, the correlations between the worker and firm effects are negatively biased due to a small number of workers moving across individual managers and police stations (Andrews et al., 2008). Rather than considering the job moves of individual workers across managers and firms, I map the managers and firms to a small number of classes. Hence, the number of job movers across manager and police station classes is large enough, and this dimension reduction technique solves the problem of limited mobility bias. The manager and police station classes are the inputs in the second step of the estimation, where I recover the model parameters leveraging job mobility as the source of identification.

Thirdly, using the structural estimates of the model, I empirically identify the heterogene-

ity and the degree of complementarities between the workers, managers, and firms. After structurally modelling the interaction among workers, managers, and firms in the production function, whether total productivity in the economy can be increased by reallocating the workers can be answered. This is possible if worker-manager-firm match specific complementarities enable gains from matching different types of workers with different types of managers and firms. For example, reallocating a low-productivity worker to a better manager would produce a larger increase in productivity than moving a high-productivity worker to a lower quality manager. I then use these estimates in the second part of the paper, where I perform counterfactual simulations by varying the sorting patterns between workers and manager-police station types. I use these simulations to find the sorting pattern that generates the police department's maximum aggregate productivity.

I do these simulated experiments by using the estimated productivity distributions of workers and the worker-manager-firm complementarities obtained by estimating the three-sided model.In other words, I answer the question; do we increase the police department's aggregate productivity by reallocating workers across police stations?

Measurement of productivity in government-controlled public services is a well-known issue in the literature (Ostrom, 1973; Cook, 1979; Mastrobuoni, 2020) because the government does not maximise profits. Due to the difficulty in productivity measurement, the productivity of public services is not documented extensively like it is for the private sector. This gap in research becomes prominent in the case of performance measurement of public services like police departments due to the lack of data in developing countries. Apart from measuring productivity, there is limited research that shows that managers matter in the public sector, and specifically in the law enforcement department of the government. Understating the sources of the productivity gap in the police department will help identify the drivers of efficiency in civil services like the police.

The results of this study are three fold. Firstly, the results delineate the individual contributions of workers and managers to the productivity of the police departments. Using the

variance decomposition exercise (Abowd et al., 1999; Card et al., 2013; Andrews et al., 2008), I find that managers individually account for 6% of the explained variation in productivity. This result on manager "fixed effects" is comparable with the literature (Fenizia, 2019) that estimates how much of the contribution of productivity in firms is explained by managerial talent. I also find that the police station effect (the firm level fixed effect) is 6% which is significantly smaller than the individual effect due to workers (57%). This significant difference in police station effect (6%) and worker effect (57%) is reminiscent of the wage dispersion literature (Abowd et al., 1999; Bonhomme et al., 2020), in which a large variation in earnings across firms is found to come mostly from workers.

Secondly, using the three-sided estimator, I find that worker, firm, and manager heterogeneity is present and is an essential factor in determining productivity. There are substantial complementarities between workers, managers, and establishments/police stations. My estimates are based on the number of manager classes M=2 and police station classes K=2. Low-type workers are 57% more productive when matched with high-type managers and more productive police stations rather than low-type managers and low-performing police stations. High-type workers are 87% more productive when matched with high-type managers and more productive police stations rather than low-type managers and less-productive police stations. Thus the productivity of workers depends on which types of managers and firms they are matched with. The fact that the gains from matching high-type workers (87%) are higher than those of low-type workers is utilised to find the sorting rule that increases the aggregate productivity in the police department. In the variance decomposition, the part of the variation in productivity explained by the covariance of worker type with the manager and the police station type is 10.3% and 10.6% respectively. This shows the prevalence of positive worker sorting in the police department productivity – workers in police departments sort moderately towards high-productivity managers. Similar positive worker sorting is reported in the literature (Bonhomme et al., 2020) on wage determination, where high-wage workers tend to sort to firms that offer high wages (Bonhomme et al., 2019). This

result reaffirms the presence of complementarity in organisations.

Thirdly, the results show evidence of the magnitude of misallocation of resources in the police department. The previous result shows the presence of heterogeneity as well as complementarities between workers, managers, and firms. The allocation of workers to managers and police stations can increase police productivity depending upon the nature of complementarities. I simulate various matching rules such as matching high type workers with high type managers (positive assortative matching) or low type managers (negative assortative matching). The counterfactual exercise allows me to conclude that if the current sorting level is raised using the optimal matching rule (positive assortative matching), then there is an 10 % increase in the aggregate productivity of the police department. Hence, social planner can maximize aggregate productivity of the police department by following the optimal worker reallocation strategy.

This paper contributes to three strands of literature.

First, it contributes to the literature that studies the worker and firm-specific effects on wage or productivity using employee-employer matched data (Abowd et al., 1999; Bonhomme et al., 2019; Card et al., 2013; Goldschmidt and Schmieder, 2017). I add managers as an additional source of heterogeneity and provide a computationally tractable model to estimate the productivity distributions. I also add to the literature that studies the matching process of workers and firms (Jackson, 2013; Finkelstein et al., 2016). In a three-sided model, I study the match-specific complementarities that arise due to the interaction of workers with different managers and firms (police station types).

Second, this paper contributes to the literature on how managers and management practice impact firm-related outcomes (Bloom et al., 2013; Bloom and Van Reenen, 2007; Lazear et al., 2015). Past research has shown that high-wage workers sort themselves to the firms that offer higher wages (Abowd et al., 1999). However, there is limited research on match-specific complementarities and the sorting patterns of worker-manager and managers-firms. Recently, Fenizia (2019) and Adhvaryu et al. (2020) have shown that manager ability ex-

plains some of the productivity gaps across heterogeneous establishments, but their results are inconclusive on manager-firm complementarity. I add the non-linear match-specific effects of the manager on police performance and show that the manager indeed contributes significantly to productivity. Managers contribution is significant in providing match-specific complementarities with the workers.

Third, my work is also related to research that evaluates the performance of public services (Best et al., 2017; Janke et al., 2019). The output measures for the government-controlled public sector have been scarce, which has limited research into the productivity of the government sector. This gap in research is large in developing countries, and I fill this gap by providing a worker-level productivity measure in the Indian police department. I use the time to clearance as a productivity measure to calculate police effectiveness (Council, 2004). My work also contributes to research that has studied the impact of civil servants on the performance of public institutions (Bertrand and Schoar, 2003; Best et al., 2017; Rasul and Rogger, 2017; Finan et al., 2017). I extend this literature by showing that the police officers' and their managers' effects on the police's performance is substantial, and present results on match-specific interactions between worker, manager, and police station.

The measure of productivity I use in the paper is documented to be better than other productivity measures in the literature like the clearance rate, crime incidence, and survey-based police perception or performance indicators (Ostrom, 1973; Cook, 1979; Eeckhout et al., 2010). My measure of time to charge sheet in Indian law enforcement agencies is used to provide the empirical measure of productivity and also identify the sources of variation of productivity across police stations.

The remainder of the paper is organised as follows. Section 2 provides relevant information regarding the institutional context. Section 3 details the theoretical model which establishes a framework for the empirical analysis. Section 4 describes the identification of the model. Section 6 documents the estimation of the parameters of my three-sided model. Section 7 describes the data and Section 8 presents the estimates of the model. Section 9

shows the results of counterfactual simulations done using worker reallocation. Section 10 presents the conclusion.

2 Background

2.1 Police structure in India

Police in India come under the state government's purview, and each of the 28 states has its own police force. The central government also has a small specialised unit primarily used to assist the state police in investigating major events and to help state governments with tasks such as intelligence gathering and research. The police force is responsible for maintaining law and order by preventing and investigating crimes. Every state is divided into various field units: zones, ranges, districts, sub-divisions or circles, police stations, and outposts for effective policing (Mitra and Gupta, 2008). For instance, a state will comprise two or more zones; each zone will comprise two or more ranges, and ranges will be sub-divided into the other field units similarly. The critical field unit in this setup is the police station within a district (Verma, 2010).

A police station is generally engaged with (i) registration of crimes, (ii) local patrolling,(iii) investigations, (iv) handling of various law and order situations (e.g., demonstrations, strikes), (v) intelligence collection, and (vi) ensuring safety and security in its jurisdiction (reference) (Mitra and Gupta, 2008; Das and Verma, 1998). A police station is headed by a Station Head Officer (SHO), generally of the rank of Inspector and occasionally of Sub-Inspector. In the hierarchy of police, the manager of the police station is the SHO, and other police workers assist him in the functioning of the police station. Junior police officers are of the rank of sub-inspector, assistant sub-inspector, and constable. When a police officer investigates a crime, he or she is called the Investigation Officer (IO) in the official documents. I treat IOs as workers.

2.2 Crime reporting and investigation

The main responsibiliy of the police is to investigate crimes. Crime reporting and investigation in India are well-established by the statutory, administrative, and judicial frameworks. Victims of an offence or anyone on the victims' behalf, including police officers, can file a complaint. Generally, a First Information Report (FIR) is registered with the police station under whose jurisdiction the geographic location of crime falls. The crimes covered under the Criminal Procedure Code (CrPc) are documented [2] in a First Information Report (FIR) (Bayley, 2015; Kumar and Kumar, 2015). An FIR is a crucial document as it sets the process of criminal justice in motion. It is only after the FIR is registered in the police station that the police can investigate the case. The FIR gets assigned to an Investigation Officer (IO) who takes up the investigation and is supervised by the Station head Officer (SHO).

The investigation of crime has many possible steps including collecting evidence, identifying suspects, recording statements of the accused, statements of witnesses, arrests, forensic analysis, and gathering expert opinion if required (Mitra and Gupta, 2008; Bayley, 2015). Criminal investigation requires skills, training, and other resources such as adequate forensic capabilities and infrastructure. The ability of the police workers plays a crucial role in criminal investigation. The quality of police officers may vary with their training, expertise, and legal knowledge in the department. The manager of the police station, or SHO, plays a vital role in supervising the police officers, as their input and direction can help speed up the investigation (Lambert et al., 2015). High-quality managers can efficiently allocate resources within a police station across multiple simultaneous investigations (Raghavan, 2003). On completion of the investigation, the police submit the final report or charge sheet to a magistrate. The submission of the charge sheet is another important step in a criminal investigation that leads to the start of a legal trial. Unsolved cases where police cannot identify suspects are closed after the magistrate's approval, and details are submitted as the final

[2] Non-serious crimes such as forgery, cheating, and defamation, which are categorised as non-cognisable in the Indian criminal codes, require prior authorisation by a magistrate before police can start investigating them

report.

2.3 Time to submit final report/charge sheet as productivity measure

There is a long-standing debate in the literature about how to measure the productivity and performance of individual police officers and police institutions (Ostrom, 1973; Verma and Gavirneni, 2006; Cook, 1979). To measure police productivity, I use the time to clear the crime as a productivity measure which is calculated as the difference between the final report/charge-sheet submission date and crime registration (FIR) date. This time to clear the crime measure is associated with police productivity as the probability of clearance of the case falls over time. The "cold case" phenomenon in criminal investigations is widely seen as an indication of poor police performance; therefore, time taken to complete the investigation directly relates to the police performance (Regoeczi and Hubbard, 2018; Addington, 2008). The time to file a charge sheet to the judicial magistrate also reflects the quality of the investigation carried out by the Indian police (Iyer et al., 2012; Amaral et al., 2019). The longer the police take to complete the investigation, the more time the accused has to manipulate the evidence and even abscond from the law. The larger times to submit a charge sheet are generally due to a longer time taken by police to record witness statements (Law, 2015). The delay in recording statements can affect witnesses' recollections of events related to crime and the identities of the accused (Read and Connolly, 2017).

Another reason I use the time to submit the charge sheet as a productivity measure is that a delay in charge sheet filing has consequences for criminal justice outcomes. The Law Commission of India (2015) survey states that 55% of pending cases in courts are delayed at the investigation stage due to the inordinate delays in filing of the charge sheets by the police. The survey [3] reports that the time gap in charge sheet filing is the most prominent reason for the delay in a criminal convictions. Low conviction rates in criminal cases are indicative of

[3] Law Commission of India (2015) survey; random sample size of 1630 responses

poor performance of law enforcement agencies. India follows the adversarial system of legal justice, where the onus of proof is generally on the state (prosecution) to prove a case against the accused. Unless the allegation against the accused is proven beyond a reasonable doubt, the accused is presumed to be innocent. Therefore, a delay in the investigations also leads to more acquittals because the accused are more likely to get bail in such cases (Krishnan and Kumar, 2010). Judges in India were asked the following question in the survey: *"Does delay in filing charge sheets adversely affect the prosecution of cases?"*. 100% of the randomly sampled judges answered yes (N=50) [4].

This measure of police productivity is related to the clearance rate, which has been widely used in the literature to measure police effectiveness (Cook, 1979; Mastrobuoni, 2020). The clearance rate is generally measured at the year-end cut-offs when crime statistics are published. These clearance rates are frequently adjusted in the upcoming reports as crimes occurring towards the year-end pose survival bias. This data adjustment causes the clearance rate statistics to be unstable because the past clearance rate improves as time passes. I use the time to charge sheet or time to solve the case as a stable measure of individual productivity at the intensive margin, whereas the clearance rate is a time censored variable. My measure of police effectiveness is also validated by Blanes Vidal and Kirchmaier (2018), who show that police response time directly affects the crime clearance rate and time to clearance.

Individual criminal cases can have characteristics, observed or not, that may determine the difficulty of solving the case itself. In my analysis, I do not control for the individual characteristics of crime. However, this may not bias the result because I control the location or police station fixed effects. For example, some police stations would encounter certain types of crime more with different difficulty levels in solving these crimes. By controlling the location fixed effects, I control the composition of crime at the police station level. However, this relies on the assumption that crime composition does not change at the police station

[4] Report of Bureau of Police Research and Development (BPRD) on increasing acquittals in India, 2013

level, which is likely to be satisfied because I rely on a short panel to estimate my results.

3 Model

I model the production in an economy where heterogeneity is three-sided and comes from workers, managers, and firms. I assume that discrete classes can represent heterogeneity (Bonhomme et al., 2019; Bonhomme and Manresa, 2015). The discrete nature of heterogeneity means that there are finite types or classes of workers, managers, and firms in the economy.

Let us assume that there are N workers, H managers, and J firms in the economy. N workers are indexed by i or $i \in \{1, ..., N\}$. H managers are indexed by h or $h \in \{1, ..., H\}$. J firms or establishments are indexed by j or $j \in \{1, ..., J\}$. I assume workers are of L different types and this type index of the i'th worker is represented by α_i where $\alpha_i \in \{1, ..., L\}$. I represent h_{it} as the identifier of the manager with whome worker i is employed at time t. I partition managers into M classes, which represents the heterogeneity across managers. I denote $m()$ as the mapping function which maps individual managers h_{it} to their classes m_{it} or $m_{it} = m(h_{it}) \in \{1, ..., M\}$. The heterogeneity across firms or establishments is described by the finite number of K partitions or classes. j_{it} is the identifier of the establishment where worker i is employed at time t. Individual firms are mapped to their classes k_{it} using the function $k()$ which takes firm identity j_{it} as the input or $k_{it} = k(j_{it}) \in \{1, ..., K\}$.

The *latent* classes m_{it} of managers and the latent classes k_{it} of firms are to be estimated, and in section 6.1 I describe this dimension reduction method. Nonetheless, the model allows the number of individual managers and firms to be equal to the number of classes or $M = H$ and $K = J$. The implication of allowing this is similar to equating the manager and firm identifiers to the class membership indicators as $m_{it} = j_{it}$ and $k_{it} = j_{it}$.

There are two time periods in the model. In time period t, the worker draws log productivity from a distribution, that depends on the worker type α_i, the worker's manager

class m_{it} and the firm class k_{it} where the worker is employed. The conditional cumulative distribution function of log productivity can be represented as below.

$$Pr[Y_{it} \leq y | m_{it} = m, k_{it} = k, \alpha_i = \alpha] = F_{mk\alpha}(y)$$

Workers at the end of the time period t who remain with the same manager and firm class is indicated by $s_{it} = 0$, these are "stayers". Workers can change either manager classes, firm classes, or both. These are "movers", end of period moves are represented by $s_{it} = 1$ and log-productivity in the next time period $t + 1$ is drawn from a distribution which depends on the worker type α_i, manager type $m_{i,t+1}$ and firm type $k_{i,t+1}$. Here, $Y_{i,t+1}$ is drawn from a distribution that depends on the parameters denoting worker state $(\alpha_i, m_{i,t+1}, k_{i,t+1})$ and t time period productivity Y_{it}. The probability that type α worker moves is not restricted in the model, and I use assumptions 1 and 2 below to simplify its dependence on specific worker states.

The following two assumptions are used in the model.

Assumption 1

A worker's probability of moving (s_{it}) and subsequent match with manager and firm (m_{it+1}, k_{it+1}) are both independent of workers current period productivity Y_{it} conditional on the worker type (α_i), current manager class (m_{it}), firm class (k_{it}) and previous moves (s_{it-1})

$$s_{it}, m_{it+1}, k_{it+1} \perp\!\!\!\perp Y_{it} | m_{it}, k_{it}, \alpha_{it}, s_{it-1}$$

Assumption 2

This assumption relates to the serial independence of productivity conditional on current state. In time period $t + 1$ worker draws productivity Y_{it+1} that depends only on α_i, m_{it+1} and k_{it+1} but not on its past productivity Y_{it}, past worker states $(m_{it},\ k_{it})$ and previous worker moves s_{it-1}

$$Y_{it+1} \perp\!\!\!\perp Y_{it}, m_{it}, k_{it}, s_{it-1} | m_{it+1}, k_{it+1}, \alpha_{it}$$

I now discuss these assumptions, their prevalence in the literature, and their implications for the model. The assumptions are related to models where the next period wage is determined only by the current state (static model of Bonhomme et al. (2019); Shimer (2005)). This means that there is no historical dependence on worker productivity beyond their current and t-1 period matches with managers and firms. Thus the model is also compatible with the class of the models where the state variables are (α, m_t, k_t) (Delacroix and Shi, 2006). The productivity drawing process is similar to the first-order Markov chain process where the current worker, manager, and firm matches break with some finite probability, and the next state is reached through a stochastic process. The model assumes that there is no human capital accumulation or on-the-job learning/training in a short period of time. This no human capital accumulation is evident in the model as workers do not change types in a short panel. My model adds managers as a *third source* of variation in the outcome (productivity), and can be seen as an extension of the two-sided labour market models where wages are the outcomes of the match between worker types and firm classes (Card et al., 2013; Alvarez et al., 2018; Bonhomme et al., 2019; Lentz et al., 2020; Abowd et al., 1999).

In my institutional context, the assumption states that Investigation Officers (worker) mobility is random, conditional on Station Head Officer (manager), police station, and time fixed effects. The assumptions allow workers to sort themselves based on manager and police station match specific productivity realizations. Thus sorting of workers does not violate the identification assumption. There are particular scenarios where these exogenous mobility assumptions will be violated. For example, workers whose productivity declines over time are reallocated to managers or police stations who have not been performing well in the past. There is less likelihood of this assumption being violated in my scenario because I work with a short period that leaves less scope of Human-capital depletion or reduction in ability in few years.

Using the above model, I will recover the distributions of the joint production function for different worker types and their match with a different manager and firm classes. I would also focus on recovering the proportions of different worker types employed within all manager and firm classes. These productivity distributions will be essential to identify the complementarities in the production functions where heterogeneity is three-sided. Apart from complementarities, the sorting patterns of workers will be recovered from data using the worker proportions. The measure of complementarities and sorting will be used to run counterfactual simulations to observe the role of heterogeneity in maximizing the total productivity in the economy.

4 Identification

In this section, I use the model described in the previous section and apply assumptions 1 and 2 to show formal identification using the observable data. There are two time periods in the model. In time period 2, worker of type α draws log productivity y_1 from cumulative distribution function $F_{mk\alpha}(y_1)$ working with manager m and firm k. Similarly the cumulative distribution function of log-productivity in period 2 is defined as $F'_{m'k'\alpha}(y_2)$. In the case of job mobility ($s_{it} = 1$), either the manager, firm class is different ($m \neq m'$ or $k \neq k'$), or the worker changes both manager and firm class ($m \neq m'$ and $k \neq k'$). I define $p_{mm',kk'}(\alpha)$ as the probability distribution of the job movers of α types between different manager and firm classes. So for moves between classes (m, k) to (m', k'), the sum of probability across worker types $\sum_{\alpha=1}^{L} p_{mm',kk'}(\alpha)$ is equal to 1. $\pi_{mk}(\alpha)$ is the distribution of α type workers in manager class m and firm class k. I can write the distribution of job movers as

$$Pr[Y_{i1} \leq y_1, Y_{i2} \leq y_2 | m_{i1} = m, m_{i2} = m', k_{i1} = k, k_{i2} = k'] = \sum_{\alpha=1}^{L} p_{mm',kk'}(\alpha) F_{mk\alpha}(y_1) F'_{m'k'\alpha}(y_2) \quad (1)$$

Equation 1 for job movers is derived after applying the assumptions 1 and 2. Assumption 1 states that the first period log productivity Y_{1t} does not depend on the next period manager (m_{i2}) and firm (k_{i2}) classes for *movers* ($s_{it} = 1$). The independence of Y_{i1} is conditional on the current match specific state of type α_i worker, manager (m_{i1}) and firm k_{i1} class. Assumption 2, which relates to serial independence, makes productivity Y_{i2} at time period 2 independent of the previous period productivity (Y_{i1}) and worker's state (m_{i1}, k_{i1}) in the previous time period. This independence assumption is conditional on the worker's match state (m_{i2}, k_{i2}) after a job move $s_{i1} = 1$. In contrast to additive model of productivity with fixed effects (Card et al., 2013; Fenizia, 2019), equation 1 allows that job mobility of workers to be endogenous in nature. Workers change job not only according to their own type, manager classes and firm types but also due to complementarity associated with the match specific realizations with managers and firms in current and future period.

I can also define the log-productivity in period 1 as below:

$$Pr[Y_{i1} \leq y_1, Y_{i2} | m_{i1} = m, k_{i1} = k] = \sum_{\alpha=1}^{L} \pi_{mk}(\alpha) F_{mk\alpha}(y_1) \quad (2)$$

I want to identify the following parameters in the model: Productivity distributions in both time periods: $F_{mk\alpha}(y_1)$, $F'_{m'k'\alpha}(y_2)$. Transition probabilities: $p_{mm',kk'}(\alpha)$. Worker proportions (sorting patters): $\pi_{mk}(\alpha)$. I rely on Theorem 1 of (Bonhomme et al., 2019) that shows the identification for a two-sided model of workers and firms. The identification of the three-sided model follows the argument that managers and firms are defined as discrete classes so the manager-firm class can be taken as a cartesian product. The replacement

of firm classes with manager×firm classes only increase the dimensionality in the BLM's identification setup, and I perform the simulations in section 6 to test if I can recover the three-sided model parameter estimates using my model.

5 Manager and firm classes identification

Identification in the previous section assumes that there are M classes of managers and K classes of firms. This dimension reduction removes the limited mobility bias (Andrews et al., 2008; Bonhomme et al., 2020), since by using finite numbers of classes, mobility of workers is from one manager or firm class to another. Thus making the number of job movers within the classes sufficiently large to avoid the small sample bias caused when job mobility within the individual managers and firms are considered. In the model illustrated in section 4, distribution of log-productivity of manager id h and firm id j does not depend on its identity beyond its manager class m and firm class k. First period log productivity shown in equation 2 can be rewritten as equation 3 for manager h and firm j. In equation 3, the left hand side depends only on the manager and firm classes which are obtained from the mapping functions $m = m(h)$ and $k = k(j)$.

$$Pr[Y_{i1} \leq y_1 | h_{i1} = h, j_{i1} = j] = \sum_{\alpha=1}^{L} \pi_{mk}(\alpha) F_{mk\alpha}(y_1) \tag{3}$$

The aim is to identify the class membership of managers and firms from theirs individual identifiers and productivity data. I first start with the intuition of identification to recover the manager classes then using a similar identification strategy firm classes can be recovered. For illustration , I assume that number of manager classes (M) = 2, the firm classes (K) = 2 and number of worker types (L) is also 2. This simplifies the graphical representation of the distribution of productivity. Figure A.1 below shows the tree diagram of the distribution represented in equation 3. There are different combinations of manager and worker classes that a worker can work with. In the example where M=2 and K=2, these combinations

are (m_1, k_1), (m_1, k_2), (m_2, k_1) and (m_2, k_2). Within each of these combinations, there are 2 types of workers $\alpha = 1$ and $\alpha = 2$ which draw productivity from match specific distribution $f_{mk\alpha}$. Next, I combine worker and firm classes together. In my example, within manager class 1 (m_1), now there are four types of worker-firm classes namely $k_1\alpha_1$, $k_1\alpha_2$, $k_2\alpha_1$ and $k_2\alpha_2$. In figure A.1, $\pi_m(k_\alpha)$ is the combined worker proportion in the manager class where $k_\alpha \in \{k_1\alpha_1, k_1\alpha_2, k_2\alpha_1\, k_2\alpha_2\}$. Thus equation 3 can be re-written as in equation 4 after combining the firm and worker classes to $K \times L$ discrete classes.

$$Pr[Y_{i1} \leq y_1 | h_{i1} = h] = \sum_{\alpha_k=1}^{K\times L} \pi_m(\alpha_k) F_{mk\alpha}(y_1) \tag{4}$$

It follows from equation 4 that the first period distribution of workers matched with manager id h is identical to the distribution of its manager class m. Thus the recovery of manager classes is essentially a classification problem where we classify the managers having similar productivity distribution to the same class. My identification strategy is similar to the Bonhomme et al. (2019) and Bonhomme and Manresa (2015), where firms classes are recovered in a two sided model. The approach to recover the firm classes follows a similar methodology, where I combine manager and worker types together. In our example above, within firm class 1 (k_1), there are four types of combined worker-manager types: $m_1\alpha_1$, $m_1\alpha_2$, $m_2\alpha_1$ and $m_2\alpha_2$. I can rewrite equation 3 conditional only on the firm id j in equation 5 below, combining the manager and worker classes to $M \times L$ classes. Equation 5 shows that firms who are of same type have identical productivity distribution, Thus recovering the firm classes is also a classification problem.

$$Pr[Y_{i1} \leq y_1 | j_{i1} = j] = \sum_{\alpha_k=1}^{M\times L} \pi_k(\alpha_m) F_{mk\alpha}(y_1) \tag{5}$$

6 Estimation

The estimation methodology of the model is divided into two steps in this section. Since the productivity distribution in Section 4 is identified using the finite manager and firm classes, I first estimate the class membership of managers and firms. The first step in Section 6.1 describes this dimension reduction methodology. Once I have classified the manager and firms into distinct classes, then I estimate the model parameters in step 2, shown in Section 6.2.

6.1 Estimating manager and firm classes

I recover the manager and firm class using a clustering algorithm. I partition H managers into M classes and J firm into K classes. This estimation strategy follows directly from equation 4 and 5. I start with recovering the managers class by solving the following equation: the three-sided counterpart of the Bonhomme et al. (2019) two-sided classification model for firms.

$$\min_{m(1),\ldots,m(H),H_1,\ldots,H_M} \sum_{h=1}^{H} n_h \sum_{d=1}^{D} \left(\widehat{F}_h(y_d) - H_{m(h)}(y_d)\right)^2 \tag{6}$$

In the above equation, $\hat{F}_h$ is the empirical CDF of log-productivity of manager h having finite support and discretized into D grids. $H_{m(h)}$ are CDFs of the manager classes. n_h is the number of workers employed under manager h. I partition the managers into M classes having cdfs $H_1, ...H_k$ so that sum of the squared error within the cluster is minimized. I weight this least square minimization problem with the number of workers in each cluster. I minimize the equation using large number of partitions in an iterative algorithm following Steinley (2006) and Bonhomme et al. (2019). The weighted k-means clustering algorithm is widely used in literature (Bonhomme et al., 2020; Zhang et al., 2019). The manager classes computed using the above k-means clustering have workers working with different firms, which is consistent with the equation 4 where I combine the worker and firm classes

to $(K \times L)$.

I now recover the firm classes identified in equation 5 by combining the worker and manager type employed within a firm to $M \times L$ classes. I use a similar clustering algorithm to partition the firms by solving the equation below.

$$\min_{k(1),\ldots,k(J),H_1,\ldots,H_K} \sum_{j=1}^{J} n_j \sum_{d=1}^{D} \left(\widehat{F}_j(y_d) - H_{k(j)}(y_d)\right)^2 \tag{7}$$

where $\hat{F}_j$ is the empirical CDF of log productivity of firm h. n_j is the number of workers employed in a firm j. I minimize the within-cluster sum of squared error to partition the firms to K classes having $H_1, ...H_K$ cdfs. This methodology results in firm class clusters having $M \times L$ types of latent worker classes according to equation 5.

Using equations 6 and 7, I estimate the firm and manager classes in the framework of Bonhomme et al. (2019) but applied to the three-sided model having a manager and firm classes. The methodology can be treated like a nested approach where I combine worker-firm types to recover manager classes and worker-manager types to estimate the firm classes. One of the advantages of this methodology is that the estimated firm and manager classes behave like Bonhomme and Manresa (2015). When the number of firms and firm size both increase to a sufficiently large number, then the estimated firm classes converge to population classes. This result directly applies to equation 6 and 7, when number of managers and firms grows or $H \to \infty$ and $J \to \infty$. Additionally, when worker per manager class is large $n_h \to \infty$ and firm size is large $n_j \to \infty$, then model estimation done in the next section is not affected by the error due to the classification step (recovery of manager and firm classes).

6.2 Estimation of model parameters

In the previous section, I estimated the manager and firm class membership. In other words, I estimated $\widehat{m}(h)$ and $\widehat{k}(j)$, and can obtain the class $\widehat{m}_{it}$ and $\widehat{k}_{it}$ for each worker. I now use these manager and firm classes in the second step to estimate the model parameters. I

assume that there ar L types of workers in the model. These are L latent types of worker, capturing unobserved heterogeneity from the worker side. In the model specification, I now define the parametric vectors. Let $f_{mk\alpha}(y_1;\theta_f)$ be the first period earnings distribution for worker type α employed with manager class m and firm class k. θ_f is the parameter vector of the distribution. For example, if the distribution $f_{mk\alpha}$ is Gaussian, then the parameter θ_f contains the mean and standard deviation (μ_f, σ_f). In the estimation, I will assume that the distribution of log productivity is Gaussian. When matched with the manager and firm class, every worker type has a different distribution, which implies that the Gaussian distribution of productivity differs along the lines of parameter θ_f. $f'_{m'k'\alpha}(y_2;\theta_{f'})$ is the productivity distribution in the second period. For the job movers of type α who change their manager and firm classes from (m,k) to (m',k'), the worker-type proportion is $p_{mm',kk'}(\alpha;\theta_p)$. θ_p is the parameter vector in the probability distribution having the length equal to the number of worker types L. Worker type proportions in the manager class m and firm class k are $\pi_{mk}(\alpha;\theta_\pi)$. θ_π is again the parameter vector whose length is equal to L types of workers. The inclusion of the two different distributions f and f' at consecutive time periods gives the model flexibility of time interaction even in the short panel. Thus the productivity distribution can vary across time and incorporates the "time fixed effects" from a Card et al. (2013) type model. I use equation 1 to write the log-likelihood function of productivity for job movers shown in equation 8 below. As previously explained in section 4, the distribution of log productivity in both time periods are independent of each other conditional on the match state of worker types with the manager and firm classes in both time periods. The log-likelihood equation is

$$\sum_{i=1}^{N}\sum_{m=1}^{M}\sum_{m'=1}^{M}\sum_{k=1}^{K}\sum_{k'=1}^{K}\mathbf{1}\{\widehat{m}_{i1}=m\}\mathbf{1}\{\widehat{m}_{i2}=m^{'}\}\mathbf{1}\{\widehat{k}_{i1}=k\}\mathbf{1}\{\widehat{k}_{i2}=k^{'}\}\times$$
$$\ln\left(\sum_{\alpha=1}^{L}p_{mm',kk'}(\alpha;\theta_p)f_{mk\alpha}(y_1;\theta_f)f'_{m'k'\alpha}(y_2;\theta_{f'})\right) \quad (8)$$

In equation 8 above, N is the number of job movers. I estimate $\widehat{\theta}_p, \widehat{\theta}_f, \widehat{\theta}_{f'}$ by maximising equation 8, and is equivalent to a mixture model representation where I do not observe the latent class of the worker. Job moves from one state (m, k) to another state (m', k') happens for al types (L) of workers. I use a modified Expectation Maximisation algorithm (Dempster et al., 1977) to estimate the parameters. One of the drawbacks of the Expectation Algorithm is that it has a slow convergence rate towards an optimal solution. I increase the convergence rate by using the Conditional Expectation-Maximization (CEM) algorithm, which maximizes conditional likelihood (Meng and Rubin, 1993; Lentz et al., 2020).

I estimate the proportion of workers in each manager class m and firm class k represented by $\pi_{mk}(\alpha; \theta_\pi)$. θ_π is the parameter vector whose length equals the number of worker types (L). I use equation 2 to write the maximum likelihood function of the worker's productivity as

$$\sum_{i=1}^{N} \sum_{m=1}^{M} \sum_{k=1}^{K} \mathbf{1}\{\widehat{m}_{i1} = m\} \mathbf{1}\{\widehat{k}_{i1} = k\} \times \ln \left(\sum_{\alpha=1}^{L} \pi_{mk}(\alpha; \theta_\pi) f_{mk\alpha}(y_1; \theta_f) \right) \tag{9}$$

I maximize the likelihood in equation 9 to estimate the parameter vector θ_π for every manager and firm class. Since I have already estimated the log productivity distribution $f_{mk\alpha}(y)$ in equation 8, the maximisation problem in equation 9 is solved by linear programming.

I now summarise the two-step estimation of the three-sided model of productivity. In Step **1**, I estimate the manager and firm classes using the classification algorithm. In Step 2, I use the manager and firm classes to estimate the parameter values for the productivity distributions and worker proportions, and the transition matrix for the different types of workers who change states. The two-step estimation approach described above is computationally tractable and combines the approach of the recent two-sided heterogeneity literature (Bonhomme et al., 2019; Lentz et al., 2020), while adding a third layer of managers.

6.3 Estimation using simulated data

I now democtrate the performance of the two-step three-sided estimator described above using simulated data. I assume the number of manager classes $M = 2$, the number of firm classes $K = 2$, and also the number of workers' types $L = 2$. I then simulate data using arbitrary parameter values from Gaussian distribution: $\theta_p, \theta_\pi, \theta_f,$ and $\theta_{f'}$. I also simulate the manager and firm IDs from the discrete classes using random draws from a uniform distribution where the mean is set to an 100 workers per manager and firm. I use this simulated data as the input to my two step, three-sided estimator described in section 6. To compare the means of estimated parameters to original values, I use the Monte Carlo simulation technique. I find that my classification method can recover the true manager and firm classes accurately as shown in appendix A (A.4 and A.5). I also find that the second step of the estimation strategy in section 6.2 produces the productivity distribution and worker proportions are close to "true" parameter values as shown in appendix A (Figure A.6 - Figure A.7).

The asymptotic properties of the estimator are presented using the Monte Carlo simulation approach. In appendix A, I show the distribution of the parameters estimated using randomly drawn data simulated using fixed model parameters. I show the asymptotic normality of the estimator by increasing the sample size, making the number of movers in the data substantially large ($N_m \to \infty$).

Though the mathematical proof of the asymptotic normality of the estimator is not provided formally, the intuition comes directly from past research (Bonhomme and Manresa, 2015; Bonhomme et al., 2019). I satisfy the assumptions used in the Bonhomme and Manresa (2015) to show that asymptotic normality holds for my model. The first assumption in Bonhomme and Manresa (2015) states that mis-classification error in the estimated manager and firm class should approach zero as the sample size grows. This assumption is likely to be satisfied in my model because my estimation methodology for recovering manager and firm classes is similar to a two-sided model where firm classes are recovered. The difference

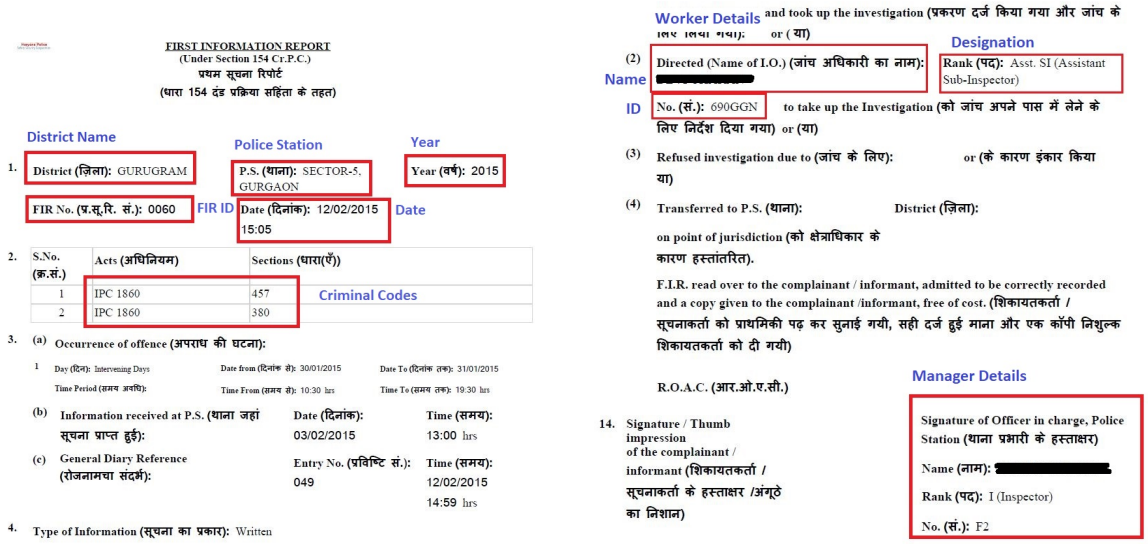
in my setting is primarily nested in nature, because I combine the worker types with firm and manager classes to determine the individual classes. The second assumption states that the properties of estimator in step 2 is like maximum likelihood estimator. This is also true as increasing the degree of freedom by adding the manager classes does not alter the properties of the maximum likelihood estimator (equation 8). Thus using the validity of the two assumptions of Bonhomme and Manresa (2015) and Bonhomme et al. (2019), my three-sided model satisfies the properties that characterize the estimator as asymptotically normal, and the same is shown using Monte Carlo simulation in appendix A (Figure A.8).

7 Data

Crime reports data: I use data on First Information Reports (FIRs) for the state of Haryana in India. Haryana is a state located in the northern part of India (Figure 2), and the Haryana police department has 283 police stations sprawled across 44 thousand square kilometres (Figure 3). I web scraped the individual FIRs for crimes reported between 2015 and 2018. The crime reports have detailed information such as the crime registration date, the administrative district where the crime is registered, the name of the police station, the crime occurrence date, and details of criminal codes applicable as per the Indian law. Each FIR also contains the identity of the Investigation Officer or IO working in the police station, who is responsible for solving the crime. FIR also records the name of the Station Head Officer (SHO), who is the Investigation Officer's (IO's) manager. A sample FIR of the police department is shown in figure 1. The figure highlights the data described above. I web scraped 472,082 of these crime reports for analysis. I then converted the unstructured data in these FIRs to machine-readable data using programmable text extraction techniques.

Productivity measure (time to submit a final report or charge sheet): The Haryana Police Department also publishes the individual case-level final report or charge sheet filing date. I construct the charge sheet data using another web scraping exercise. I

Figure 1: Sample First Information Report (FIR) with highlighted data



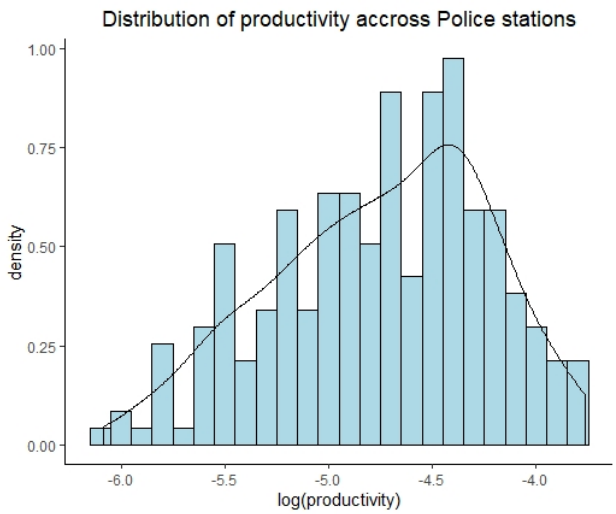
then match all the crime reports using the unique FIR id to their outcome, i.e., the time to submit the final report or charge sheet (Appendix A, Figure A.2) .

Worker-manager-establishment matched data set and job mobility: My methodology of matched data set construction of the three-sided model (worker-manager-firm) follows the approach extensively used in literature (Abowd et al., 2002; Bonhomme et al., 2020) that uncovers worker and firm fixed effects using matched employee-employer data. The case's name and the unique employee id of both the Investigation officers (IO) and their supervisors, i.e., the Station Head Officers (SHO) are observed in the FIR data. To create the matched data set, I use the anonymised unique ids of the employees of the police department. In a few reports where the employee id is missing or erroneous, I created synthetic employee ids using the officer name. To draw the analogy from the three-sided model described in Section 3 to the crime reports data of the police department, I consider the investigation officer (IO) as the worker, the station head officer (SHO) as the manager, and the police station as the establishment or firm. I infer the job mobility of an investigation officer (IO) when he/she moves to different managers and police stations. Hence, from the data, I can observe the job mobility patterns of workers across managers and police stations.

Figure 2: Map showing India and the state of Haryana is shaded (red)

Figure 3: Location of police stations in Haryana shown as dots (black)

Figure 4: Distribution of log productivity across police stations in Haryana



Note: Log productivity in the x-axis is the negative transformation of log(Time to chargesheet)

Figure 4 shows the distribution of log productivity of employees across police stations in Haryana for the year 2017. There is a large dispersion visible in productivity across police stations. The top decile police station is 2.75 times more productive than the bottom decile police station. Productivity comparison of police is scarce in the past literature. Hence I use the benchmark from firm-level log-productivity distribution from Syverson (2011). It reports the within plant productivity gap of 2.9 in India, comparable to the police productivity gap across police stations.

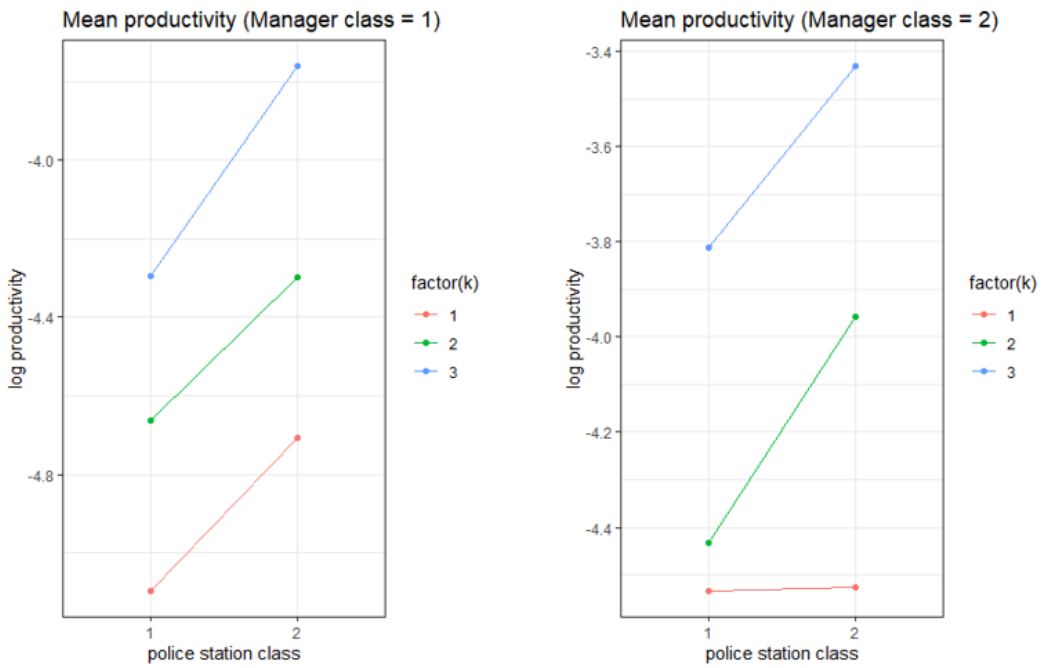
8 Results

I use the worker-manager-police station matched data set of all crime cases registered from 2015 to 2018. I follow the sample selection methodology described in Bonhomme et al. (2019), Friedrich et al. (2019), and Fenizia (2019), which uses weekly wage data in employee-firm matched data sets. Following Friedrich et al. (2019), in my analysis, I consider workers who have worked for at least three months in a police station. I use at least three months for a worker because I want to exclude the temporarily seconded employees. They are

sometimes posted in a police station as a trainee or short-term replacement of a police officer. Occasionally, the employee id column is entered manually, which might cause an error in inferring the job mobility of workers and managers across police stations. In such situations, I use the names of employees to avoid ambiguity in id matching. I track job changes by capturing all state changes of a worker due to his or her matches with managers and police stations. In addition to the model having different distributions of productivity of workers for different periods, the inclusion of such higher frequency job mobility will also be useful in incorporating time fixed effects (Lentz et al., 2020). I use the aggregated case-level outcomes for each employee to derive the employee-level productivity measure. Employees with higher charge sheet time are considered less productive.

I now estimate the model assuming the number of classes of managers as $M = 2$ and classes of police stations as $K = 2$. Two factors guide this assumption of a finite number of classes. The first one relates to past literature, which assumes that a small number of groups can represent the substantial heterogeneity in the classes. For example, Bonhomme et al. (2019) assume ten firm classes in the Swedish data and recently they have also assumed 10 classes in the research on the labour market in the USA and Italy (Bonhomme et al., 2020). The substantial earnings difference across all classes (either manager or firm) is critical for assuming the number of classes. The second criteria relates to the restriction posed by the finite sample available for analysis. The research based on economy-wide employee-employer matched data (Bonhomme et al., 2019) has a large sample size of 0.5 million workers in 42K firms. Therefore, at an average of 140 workers per firm, Bonhomme et al. (2019) can choose to have several firm classes equal to 10 and still have many workers present within each class. The sample size of the police data I use is smaller when compared with economy-wide administrative data used in previous research. The police department sample has 9581 police officers (Investigation officers), 1007 managers (Station head officers) employed within 282 police stations. This amounts to around 9 workers per manager and 30 workers per police station in my data.

Figure 5: Estimates of the static model on police department data of Haryana. Estimates of means of log-productivity, by worker type (IO), manager (SHO), and police station class. I order the manager class (M = 2) and firm class (K=2) (on the x-axis) by mean log-productivity. On the y-axis we report estimates of mean log-productivity for the L = 3 police officer/Investigation Officer types.



Note: Log productivity in the x-axis is the negative transformation of log(Time to chargesheet)

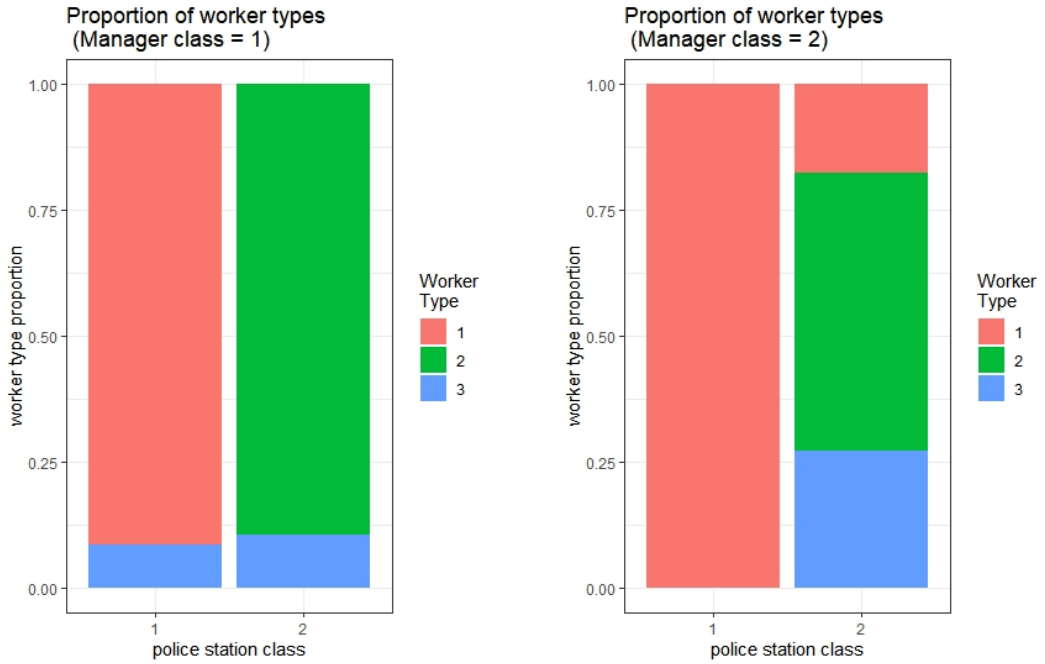
I estimate the manager and firm classes by weighted k-means clustering described in Section 6.1. I estimate the model parameters using step 2, shown in Section 6.2. My estimates are based on the number of manager classes M=2 and police station classes K=2. I use the Gaussian finite mixture in equation 1 assuming the number of worker types L = 3. I estimate the productivity distribution and proportion of job movers across managers and firms using equation 8, and then I estimate the worker proportions using equation 9. As described in Section 6.2, I estimate the finite mixture model using the ECM (Expectation Conditional Maximisation) algorithm. A well-known problem associated with the ECM algorithm is that it can converge at local maxima and consequently fail to reach global maxima (Wu, 1983). To alleviate this concern, I estimate the parameters using multiple starting points using a grid-based parameter search methodology (Biernacki et al., 2003). I then choose the result

(global maxima) that has the maximum likelihood among the converged solutions.

The results of mean worker productivity are presented in Figure 5. The estimation results show the mean log-productivity of workers when matched with low type managers (m=1) and high type managers (m=2) separately. Within the manager classes, police station classes are shown ordered by productivity. So police station class (k=1) shown on the x-axis has lower productivity than the class k=2. In both panels in the Figure 5, I show the mean log productivity of the 3 types of workers in each type of police station class. The difference in log productivity (Figure 5) among different worker types across the manager and police station classes shows the worker-manager-police station heterogeneity. The estimates indicate the complementarities between worker, manager, and police station types, as the mean productivity of the same worker type plotted across police station types is not parallel. There is growth in the productivity of high-type workers when matched with the high-type police station and manager. For example, suppose I match the high type worker ($\alpha = 3$) to a highly productive manager and police station. In that case, there is a 40% benefit in doing so when compared with matching the lower type worker ($\alpha = 1$) to the highly productive manager and police station. Thus, the match-related complementarities are large in magnitude, suggesting that workers can gain immensely by matching with the right type of manager and police station.

I also present the estimates of $\pi_{\alpha mk}$ or the proportions of workers in the manager and police-station classes. Figure 6 represents the worker proportions. I show worker proportions within low productivity managers (m = 1) in the left figure, whereas in the right figure, they are for high productive managers (m=2). I observe that, within the less productive manager class (m=1) and least productive police station (k=1), most of the workers are of the lowest type $\alpha = 1$. These figures also then, show the sorting pattern that exist in the police force. The highest type worker proportion monotonically increases with the higher productive classes, i.e., the positive sorting of workers to manager-firm classes. The proportion of high type workers is 10% with the lowest manager type and firm class, whereas 70% of the workers

Figure 6: Estimates of the proportions of worker types. Worker proportion in manager class m =1 (left) and m=2 (right) across police station classes



have the highest manager and firm classes. This positive sorting of workers resembles what has been found in the wage heterogeneity literature. The variation in log earnings is due to sorting patterns, i.e., high-productivity firms employ high type workers disproportionately. My model of three-sided heterogeneity reveals a large difference in worker productivity due to the strong presence of complementarities that are not observed in the wage dispersion literature.

8.1 Variance-Covariance decomposition of productivity

In this subsection, I propose a variance decomposition of the productivity. I extend the methodology of Abowd et al. (1999), Card et al. (2013), and Fenizia (2019) to my model where the heterogeneity comes from three sides. In my model, the variation in productivity is explained by worker quality α_i, manager m_i, and firm heterogeneity k_i. I follow the Bonhomme et al. (2019) methodology, in which I perform the three-sided decomposition of productivity by linearly projecting the log productivity on the worker, manager, and police

station classes indicators, without interaction. The variance-covariance decomposition of the linear model is given below.

$$Var(Y_{it}) = Var(\alpha_i) + Var(m_{it}) + Var(k_{it}) + Var(\epsilon_{it}) + 2Cov(\alpha_i, m_{it}) + 2Cov(\alpha_i, k_{it}) + 2Cov(m_{it}, k_{it}) \quad (10)$$

where α_i is employee type, m_{it} is manager class and k_{it} is police station class. Equation 10 decomposes the variance of log productivity into the variances of worker type effect α , manager class m, police station k, the combination of the covariances, and residual variation. The results of the variance-covariance decomposition from equation 10 are shown in Table 1. The worker productivity component explains 57% of the total variation. The worker share is high and comparable to recent estimates of worker effects in wage dispersion (Bonhomme et al., 2019; Lentz et al., 2020; Bagger et al., 2013; Abowd et al., 1999; Card et al., 2013). Managers explain 6.2% of the productivity described in the three-sided model, and the effects of the police station are similar (6.4%). Thus the effects of the manager on productivity are comparable to firm-specific effects on productivity. In my estimation, the manager (Station Head Officer) effect is in line with the literature discussing the role of management on firm productivity (Fenizia, 2019; Lazear et al., 2015; Bloom and Van Reenen, 2007; Bloom et al., 2013).

Table 1 also shows the share of productivity variation explained by the covariance between workers, managers, and police stations. The covariances explain 30% of productivity variation. The correlation between the manager and police station is 72%, which shows a high degree of sorting between the managers and police stations. The degree of correlation between workers and managers is 27%, similar to the correlation between workers and police stations. This moderate correlation between worker-manager and worker-police station types shows the presence of positive sorting. Similar positive sorting of workers is reported in the

Table 1: Variance decomposition exercise

	Variance share
Var(Worker)	57.3
Var(Manager)	6.2
Var(Police station)	6.4
2Cov(Worker, Manager)	10.3
2Cov(Worker, Police station)	10.6
2Cov(Manager, Police station)	9.1
Corr(Worker, Manager)	27.4
Corr(Worker, Police station)	27.9
Corr(Manager, Police station)	72.1
R squared	31.9

Notes: Linear regression $Y_{it} = \alpha_i + m_{it} + k_{it} + \epsilon_{it}$ on the estimated values of model

literature on wage determination, where high-wage workers tend to sort to firms that offer high wages (Bagger et al., 2013; Card et al., 2013; Abowd et al., 1999; Lentz et al., 2020). The presence of moderate sorting of workers will be important in the counterfactual exercise shown in the next step. In the next section, I increase the degree of positive assortative match of workers with managers and police stations.

9 Reallocating workers: Counterfactual simulations

I have estimated the underlying parameters of the productivity distribution in the previous section. The estimation of equations 8 and 9 gives estimates of the underlying structural parameters of the model described in Section 3. In this section, I execute a counterfactual exercise to show that the changes in the matching pattern of workers with managers and firms can increase aggregate productivity. I change the matching rule by reallocating workers to different managers and police stations while keeping the number of workers and their quality fixed. While performing the counterfactual exercise, I rely on the complementarity of production shown in the estimates of mean log productivity in Figure 5. The constraint on the number of workers of specific types within manager and police station classes is taken from the estimates shown in Figure 6. I assume that the log productivity distribution

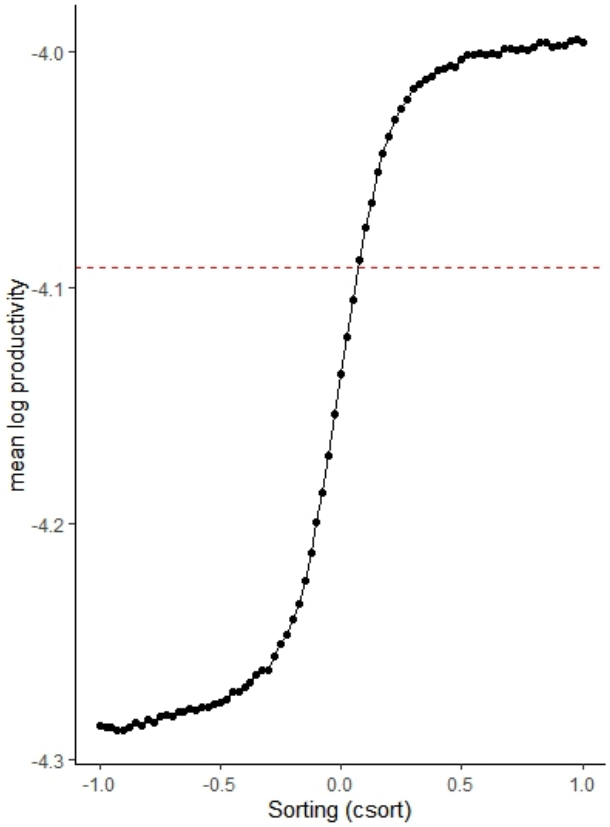
remains identical when workers are matched with different classes of managers and police stations/firms.

There are certain matching rules defined in the literature that provide optimal aggregate productivity when the production function has complementarities between input factors. Becker (1973), and Eeckhout and Kircher (2011) both states that positive assortative matching is optimal when the production function or the match surplus exhibits supermodularity (a strong form of positive complementarity). In the Indian police, productivity gains from matching the high-quality worker with a highly productive manager are higher (87%) when compared with matching a low-quality worker with a highly productive manager (57%). So the optimal matching rule will be to pair high-type workers with high-type managers and low-type workers with the less-productive manager (Topkis, 2011; Eeckhout and Kircher, 2011). On the contrary, if the complementarities are negative or the submodularity exists in the production function, then the optimal matching rule should be negative assortative matching.

I follow the following algorithm to vary the degree of assortative matching. I compute the counterfactual worker proportion in each manager and firm type for two matching rules, namely positive and negative assortative matching. In Figure A.3, worker proportions are calculated using the pure positive assortative matching rule. I compute the counterfactual proportions in Figure A.3 by rank ordering the worker, manager, and police station by their types/classes and then allocating the workers to manager and police station classes as per their rank. Similarly, counterfactual allocation of worker types $\pi_{mk}^{nam}(\alpha)$ for negative assortative matching is calculated. I then simulate the intermediate sorting patterns by randomly allocating some workers within the manager and police station classes. To get a sequence of multiple sorting patterns, I increase these randomly chosen worker proportions iteratively (by 0.5% of the total worker population). The degree of positive assortative is matching increasing in these sequences.

I use the simulated sequence of sorting patterns $\pi_{mk}^{cf}(\alpha)$ described above to generate the

Figure 7: Haryana police department's aggregate productivity (Y-axis) calculated using multiple degrees (X-axis) of workers matching with managers and police stations.



Note: csort in x-axis is the simulated sorting pattern. Corner values represent the positive (+1) and negative (-1) assortative matching rule

counterfactual productivity by using the parameter estimates from Section 6. Figure 7 shows the counterfactual simulation results. The x-axis shows the simulated sorting pattern, which I generated using the algorithm described above. In the x-axis, corner values represent the positive (+1) and negative (-1) assortative matching rule. I estimate the benefit of reallocating police officers by using the productivity distribution of all the simulated match rules (x-axis) and find the optimal sorting pattern of workers that maximises total productivity in the police department.

The results in Figure 7 show that the police department's aggregate productivity increases monotonically as the degree of workers matching with managers and police stations changes from negative assortative to positive assortative. The supermodular nature of the production

function in the three-sided case shows pure positive assortative matching as the optimal solution. Table 2 shows the estimates of change in productivity between the original and counterfactual states. Results show that the police department can increase productivity by 9.2% by reallocating the workers using a positive assortative match rule, i.e. matching high-quality workers with highly productive managers and police stations. I also compare the counterfactual distribution of productivity with the current productivity distribution in the police department. The optimal match leads to higher benefits in the top 90% percentile of the productivity distribution. Table 2 shows that the 90% percentile receives a 30% improvement in productivity. This is because the current allocation of high-quality workers in the police department is sub-optimal in terms of matching. High gains can be achieved by leveraging the strong complementarities between workers, managers, and police stations.

Table 2: Estimates of productivity at optimal matching rule

Reallocation exercise ($\times 100$)			
Mean	Median	10% quantile	90% quantile
Positive Assortative matching			
9.2	6.7	-3.9	30.7

Differences in the means, quantiles of log productivity between two samples: counterfactual sample where workers are reallocated optimally, and the original sample

10 Conclusion

In this paper, I decouple the effects of workers, managers, and firms on productivity and show that heterogeneity matters. The worker, manager, and firm types heterogeneity are essential determinants of productivity. The empirical analysis uses data from an Indian police department. It shows that a manager's contribution to productivity is significant, and a central planner can increase productivity by leveraging the complementarities between worker and manager-firm matching. Counterfactual simulation shows that if high-type workers are matched to high-type managers and highly productive police stations, then aggregate produc-

tivity of the police department can be increased by ten percent. This suggests misallocation of resources within the police department.

Identifying the roles of workers, managers, firms, and their interaction relies on the job moves observed in my micro-level data. The estimation methodology adopted in the paper has the advantage of beingmore robust than linear fixed effects models, since it does not have a functional assumption like additive and linear assumptions of the AKM's fixed-effects model. The model does not restrict the match-related complementarity arising from workers placed with different types of managers and police stations. This methodology helps me recover the structural estimates of the parameters that define heterogeneity in the production function. Moreover, this methodology adopts an approach that classifies managers and firms into discrete types, thereby circumventing the limited mobility bias issues debated in the literature of two-sided heterogeneity.

This paper brings new insights into the productivity of public institutions like police departments, which is difficult to measure, especially in developing countries like India. The enormous productivity gap across the police stations arises from the underlying heterogeneity of workers, managers, and firms. This paper shows that similar to the private sector, managers are relevant in police departments too. The significant manager effect helps to understand the functioning of public institutions from the perspective of managerial talent and leadership. The results also reconcile the theoretical framework where complementarities in production function can lead to different optimal matching rules. This study shows that the optimal matching rule in the Indian police department is positive assortative matching. The positive assortative matching rule is due to positive complementarity between worker and the manager-firm match types.

The methodology adopted in this paper is general for scenarios when the outcome is generated from a process where heterogeneity is three-sided. However, I use police department productivity to show the presence of heterogeneity. Therefore, the empirical estimates of the magnitude of misallocation can be extrapolated to police departments only. Future

research can adapt this methodology to specific sectors of the economy and the public sector departments.

As previously discussed, this paper uses job mobility to identify misallocation of resources in the public sector in India and also persistent productivity gap across locations. How this persistent gap remains incentive-compatible in the public sector, remains an open question.

References

Abowd, J. M., R. H. Creecy, and F. Kramarz (2002, March). Computing Person and Firm Effects Using Linked Longitudinal Employer-Employee Data. Longitudinal Employer-Household Dynamics Technical Papers 2002-06.

Abowd, J. M., F. Kramarz, and D. N. Margolis (1999). High wage workers and high wage firms. *Econometrica 67*(2), 251–333.

Addington, L. A. (2008). Assessing the extent of nonresponse bias on nibrs estimates of violent crime. *Journal of Contemporary Criminal Justice 24*(1), 32–49.

Adhvaryu, A., V. Bassi, A. Nyshadham, and J. A. Tamayo (2020, April). No Line Left Behind: Assortative Matching Inside the Firm. NBER Working Papers 27006.

Alvarez, J., F. Benguria, N. Engbom, and C. Moser (2018). Firms and the decline in earnings inequality in brazil. *American Economic Journal: Macroeconomics 10*(1), 149–89.

Amaral, S., S. Bhalotra, and N. Prakash (2019, April). Gender, Crime and Punishment: Evidence from Women Police Stations in India. Boston University - Department of Economics - The Institute for Economic Development Working Papers Series dp-309, Boston University - Department of Economics.

Andrews, M. J., L. Gill, T. Schank, and R. Upward (2008). High wage workers and low wage firms: negative assortative matching or limited mobility bias? *Journal of the Royal Statistical Society: Series A (Statistics in Society) 171*(3), 673–697.

Bagger, J., K. L. Sørensen, and R. Vejlin (2013). Wage sorting trends. *Economics Letters 118*(1), 63–67.

Banerjee, A. and E. Duflo (2005). Growth theory through the lens of development economics. In P. Aghion and S. Durlauf (Eds.), *Handbook of Economic Growth* (1 ed.), Volume 1, Part A, Chapter 07, pp. 473–552. Elsevier.

Bayley, D. H. (2015). *Police and political development in India*. Princeton University Press.

Becker, G. S. (1973). A theory of marriage: Part i. *Journal of Political Economy 81*(4), 813–846.

Bertrand, M. and A. Schoar (2003, 11). Managing with Style: The Effect of Managers on Firm Policies*. *The Quarterly Journal of Economics 118*(4), 1169–1208.

Best, M. C., J. Hjort, and D. Szakonyi (2017, April). Individuals and Organizations as Sources of State Effectiveness. NBER Working Papers 23350, National Bureau of Economic Research, Inc.

Biernacki, C., G. Celeux, and G. Govaert (2003). Choosing starting values for the em algorithm for getting the highest likelihood in multivariate gaussian mixture models. *Computational Statistics & Data Analysis 41*(3-4), 561–575.

Blanes Vidal, J. and T. Kirchmaier (2018). The effect of police response time on crime clearance rates. *The Review of Economic Studies 85*(2), 855–891.

Bloom, N., B. Eifert, A. Mahajan, D. McKenzie, and J. Roberts (2013). Does management matter? evidence from india. *The Quarterly Journal of Economics 128*(1), 1–51.

Bloom, N. and J. Van Reenen (2007, 11). Measuring and Explaining Management Practices Across Firms and Countries*. *The Quarterly Journal of Economics 122*(4), 1351–1408.

Bonhomme, S., K. Holzheu, T. Lamadon, E. Manresa, M. Mogstad, and B. Setzler (2020, June). How Much Should we Trust Estimates of Firm Effects and Worker Sorting? NBER Working Papers 27368, National Bureau of Economic Research, Inc.

Bonhomme, S., T. Lamadon, and E. Manresa (2019). A distributional framework for matched employer employee data. *Econometrica 87*(3), 699–739.

Bonhomme, S. and E. Manresa (2015). Grouped patterns of heterogeneity in panel data. *Econometrica 83*(3), 1147–1184.

Card, D., J. Heining, and P. Kline (2013, 05). Workplace Heterogeneity and the Rise of West German Wage Inequality. *The Quarterly Journal of Economics 128*(3), 967–1015.

Cook, P. J. (1979). The clearance rate as a measure of criminal justice system effectiveness. *Journal of Public Economics 11*(1), 135–142.

Council, N. R. (2004). *Fairness and Effectiveness in Policing: The Evidence*. Washington, DC: The National Academies Press.

Das, D. K. and A. Verma (1998). The armed police in the british colonial tradition: The indian perspective. *Policing: An International Journal of Police Strategies & Management*.

Delacroix, A. and S. Shi (2006). Directed search on the job and the wage ladder. *International Economic Review 47*(2), 651–699.

Dempster, A. P., N. M. Laird, and D. B. Rubin (1977). Maximum likelihood from incomplete data via the em algorithm. *Journal of the Royal Statistical Society: Series B (Methodological) 39*(1), 1–22.

Eeckhout, J. and P. Kircher (2011). Identifying sorting—in theory. *The Review of Economic Studies 78*(3), 872–906.

Eeckhout, J., N. Persico, and P. E. Todd (2010). A theory of optimal random crackdowns. *American Economic Review 100*(3), 1104–35.

Fenizia, A. (2019). Managers and productivity in the public sector. Working paper.

Finan, F., B. A. Olken, and R. Pande (2017). The personnel economics of the developing state. *Handbook of economic field experiments 2*, 467–514.

Finkelstein, A., M. Gentzkow, and H. Williams (2016, 07). Sources of Geographic Variation in Health Care: Evidence From Patient Migration. *The Quarterly Journal of Economics 131*(4), 1681–1726.

Friedrich, B., L. Laun, C. Meghir, and L. Pistaferri (2019). Earnings dynamics and firm-level shocks. NBER Working Papers 25786, National Bureau of Economic Research.

Goldschmidt, D. and J. F. Schmieder (2017). The rise of domestic outsourcing and the evolution of the german wage structure. *The Quarterly Journal of Economics 132*(3), 1165–1217.

Hsieh, C.-T. and P. J. Klenow (2009, 11). Misallocation and Manufacturing TFP in China and India*. *The Quarterly Journal of Economics 124*(4), 1403–1448.

Iyer, L., A. Mani, P. Mishra, and P. Topalova (2012). The power of political voice: women's political representation and crime in india. *American Economic Journal: Applied Economics 4*(4), 165–93.

Jackson, C. K. (2013). Match quality, worker productivity, and worker mobility: Direct evidence from teachers. *Review of Economics and Statistics 95*(4), 1096–1116.

Janke, K., C. Propper, and R. Sadun (2019, May). The Impact of CEOs in the Public Sector: Evidence from the English NHS. Technical Report 13726, C.E.P.R. Discussion Papers.

Krishnan, J. K. and C. R. Kumar (2010). Delay in process, denial of justice: the jurisprudence and empirics of speedy trials in comparative perspective. *Georgetown Journal of International Law 42*, 747.

Kumar, S. and S. Kumar (2015). Does modernization improve performance: evidence from indian police. *European journal of law and economics 39*(1), 57–77.

Lambert, E. G., H. Qureshi, N. L. Hogan, C. Klahm, B. Smith, and J. Frank (2015). The association of job variables with job involvement, job satisfaction, and organizational commitment among indian police officers. *International Criminal Justice Review 25*(2), 194–213.

Law, J. (2015). Law commission of india : Arrears and backlog - creating additional judicial manpower. Report no. 245, Government of India.

Lazear, E. P., K. L. Shaw, and C. T. Stanton (2015). The value of bosses. *Journal of Labor Economics 33*(4), 823–861.

Lentz, R., J.-M. Robin, and S. Piyapromdee (2020). On worker and firm heterogeneity in wages and employment mobility: Evidence from danish register data. Working paper.

Mastrobuoni, G. (2020). Crime is terribly revealing: Information technology and police productivity. *The Review of Economic Studies 87*(6), 2727–2753.

Meng, X.-L. and D. B. Rubin (1993). Maximum likelihood estimation via the ecm algorithm: A general framework. *Biometrika 80*(2), 267–278.

Mitra, R. and M. Gupta (2008). A contextual perspective of performance assessment in egovernment: A study of indian police administration. *Government Information Quarterly 25*(2), 278–302.

Ostrom, E. (1973). On the meaning and measurement of output and efficiency in the provision of urban police services. *Journal of Criminal Justice 1*(2), 93–111.

Raghavan, R. K. (2003). The indian police: Problems and prospects. *Publius: the journal of federalism 33*(4), 119–134.

Rasul, I. and D. Rogger (2017, 04). Management of Bureaucrats and Public Service Delivery: Evidence from the Nigerian Civil Service. *The Economic Journal 128*(608), 413–446.

Read, J. D. and D. A. Connolly (2017). The effects of delay on long-term memory for witnessed events. *The Handbook of Eyewitness Psychology: Volume I: Memory for Events*, 4.

Regoeczi, W. C. and D. J. Hubbard (2018, 09). The impact of specialized domestic violence units on case processing. *American Journal of Criminal Justice : AJCJ 43*(3), 570–590.

Restuccia, D. and R. Rogerson (2017). The causes and costs of misallocation. *Journal of Economic Perspectives 31*(3), 151–74.

Shimer, R. (2005). The assignment of workers to jobs in an economy with coordination frictions. *Journal of Political Economy 113*(5), 996–1025.

Shimer, R. and L. Smith (2000). Assortative matching and search. *Econometrica 68*(2), 343–369.

Steinley, D. (2006). K-means clustering: a half-century synthesis. *British Journal of Mathematical and Statistical Psychology 59*(1), 1–34.

Syverson, C. (2011, June). What determines productivity? *Journal of Economic Literature 49*(2), 326–65.

Topkis, D. M. (2011). *Supermodularity and complementarity.* Princeton university press.

Verma, A. (2010). *The new khaki: The evolving nature of policing in India.* Routledge.

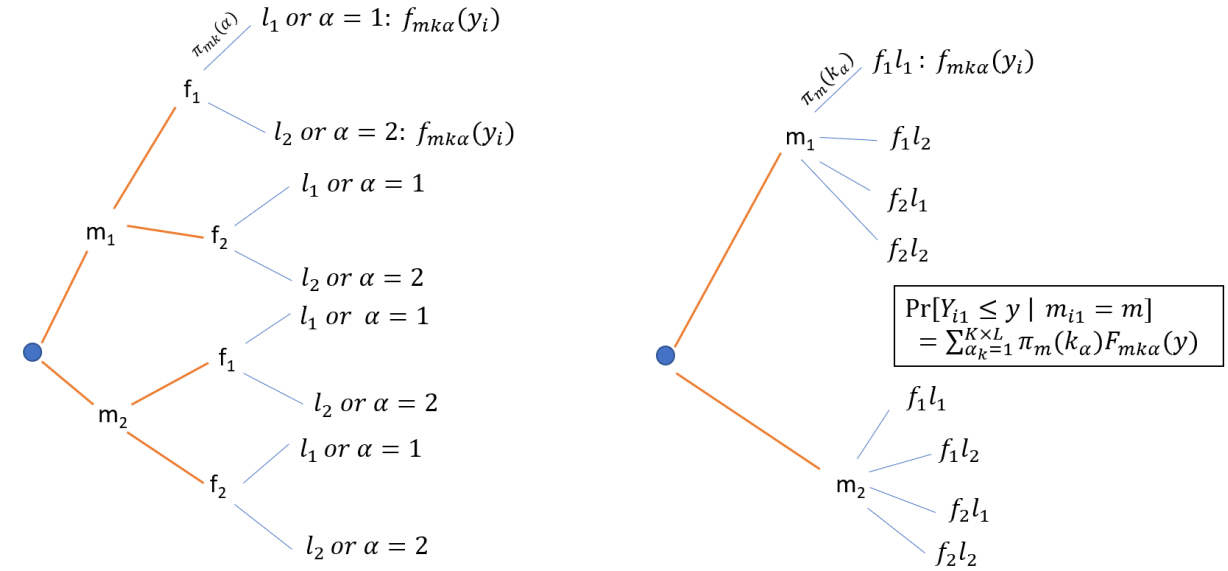
Verma, A. and S. Gavirneni (2006). Measuring police efficiency in india: an application of data envelopment analysis. *Policing: An International Journal 25*(1), 125–145.

Wu, C. F. J. (1983). On the convergence properties of the em algorithm. *The Annals of Statistics 11*(1), 95–103.

Zhang, Y., H. J. Wang, and Z. Zhu (2019). Quantile-regression-based clustering for panel data. *Journal of Econometrics 213*(1), 54–67.

A Figures

Figure A.1: Tree diagram of the distribution represented when together (No. of) manager classes (M) = 2, firm classes (K) = 2 and worker types (L) = 2 *(right)*. Tree diagram after combining the firm classes and worker type *(left)*



Note: I combine firm and worker class from figure in right to figure in left. For example, within manager class 1 (m_1), now there are four types of worker-firm classes namely $k_1 l_1$, $k_1 l_2$, $k_2 l_1$ and $k_2 l_2$.

Figure A.2: Charge sheet date sample from Haryana Police department webpage

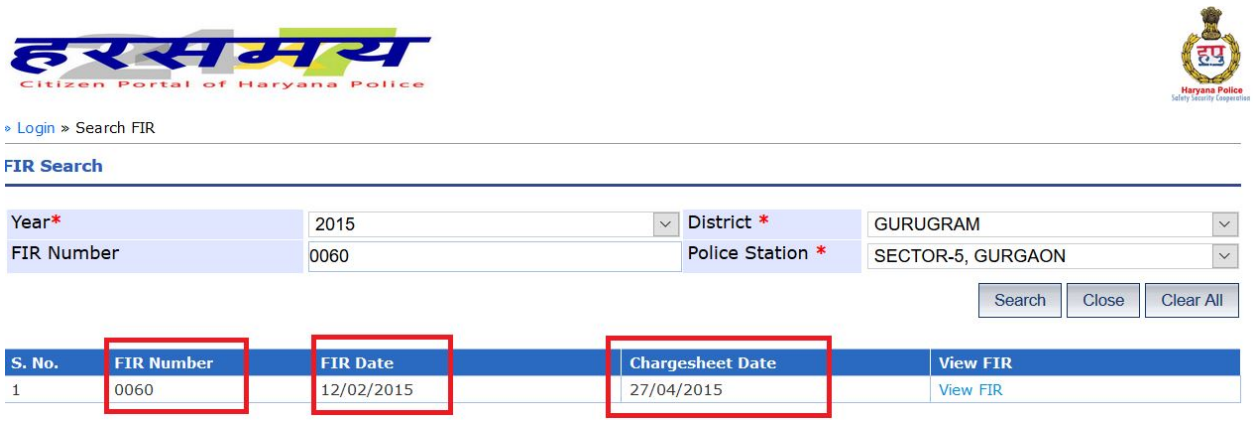


Figure A.3: Counterfactual allocation of workers using positive assortative matching

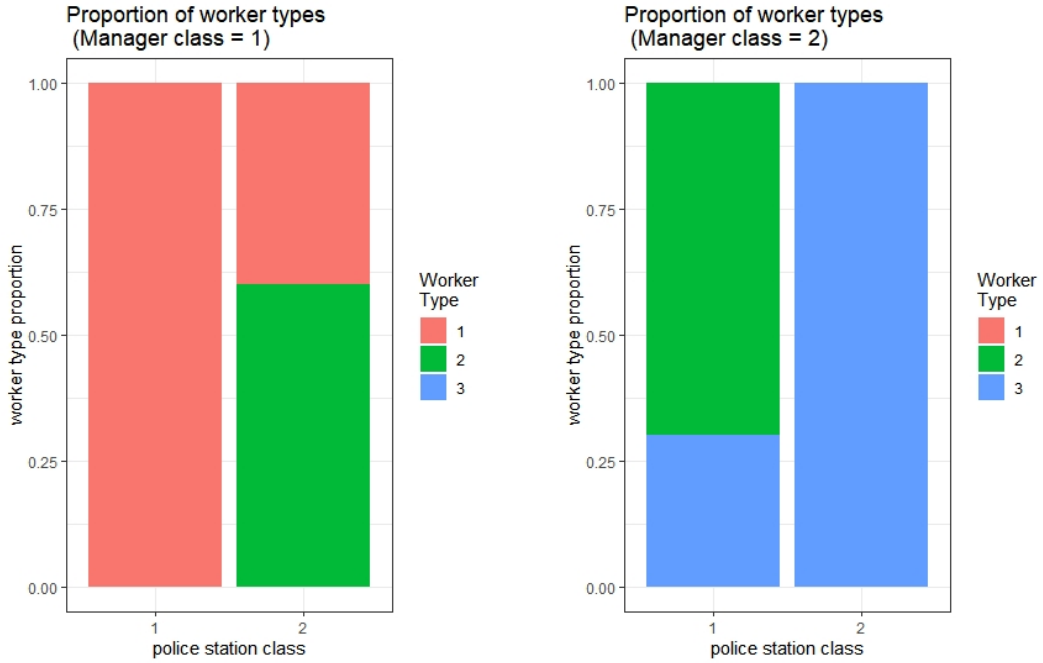


Figure A.4: Simulated data: Manager classes estimated by combining firm and worker classes together. ($K \times L$ or 2×2)

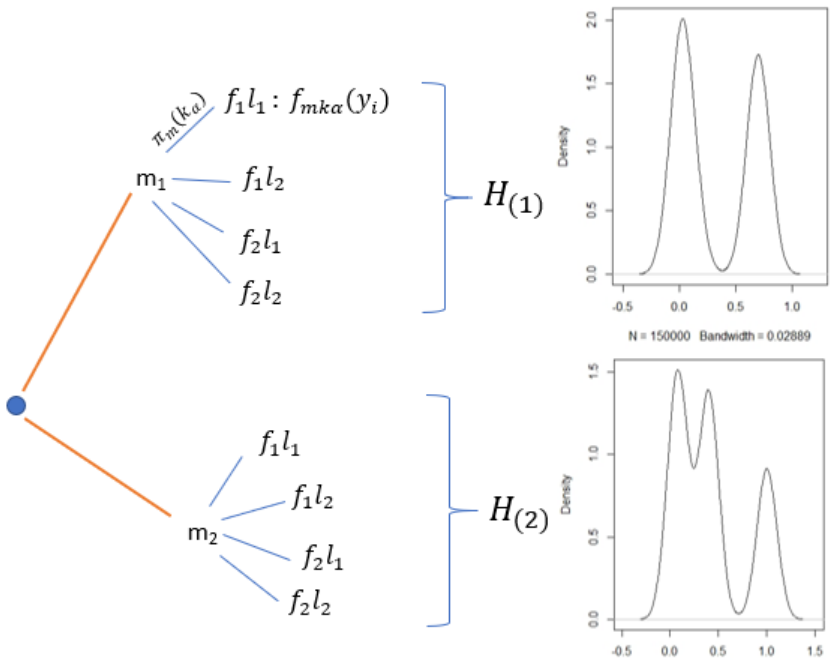


Figure A.5: simulated data: recovering the manager classes (Low misclassification rate (less than 1%))

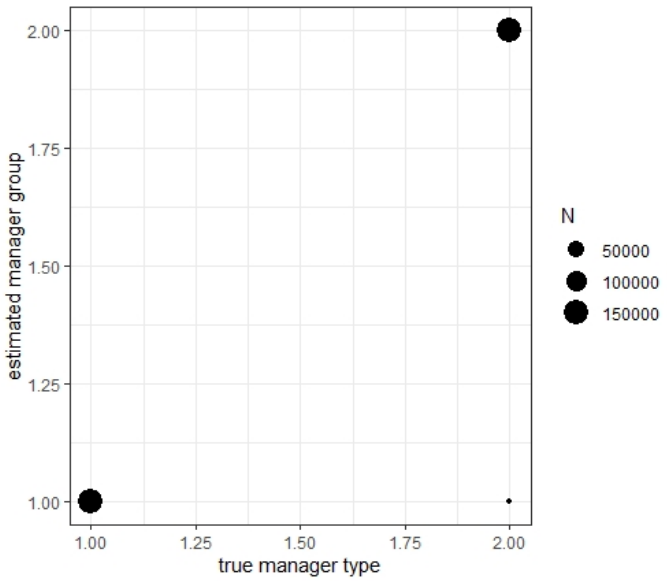


Figure A.6: Estimates of Step 2 on simulated data : Model parameters - Bold(circles) lines are true parameter values and dotted (triangles) are estimated values

Figure A.7: Simulated data: estimating model parameters: recovering $\pi_{mk}(\alpha)$: worker proportions for manager class = 2

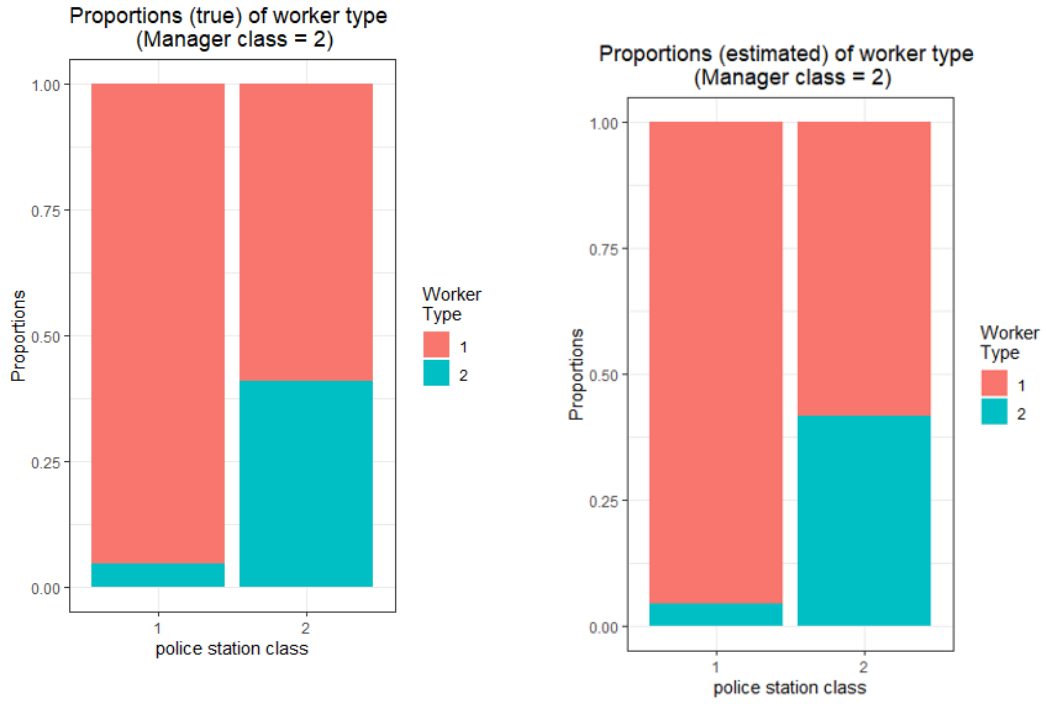


Figure A.8: Asymptotic properties: Monte Carlo simulations

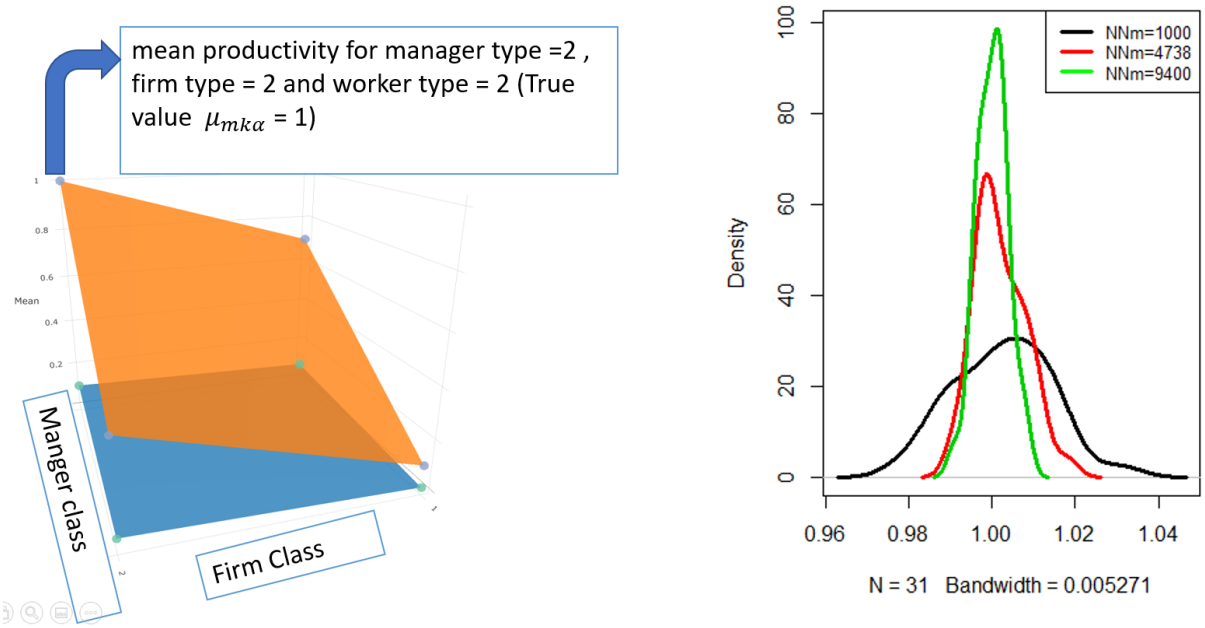


Figure A.9: Simulated data: estimating model parameters: recovering $\pi_{mk}(\alpha)$: worker proportions for manager class = 1