Sources of productivity dispersion in public services: Evidence from Indian police

Amit Chaudhary

University of Warwick

15 January, 2021

Motivation

- Can we increase productivity by **reallocating resources**?
 - Input misallocation Source of productivity gap between firms (Banerjee and Duflo, 2005; Hsieh and Klenow, 2009)
 - Need to know what drives large and persistent productivity gaps (Syverson, 2011)
- Recent paper Unobserved worker and firm heterogeneity shape wage/productivity distribution (Bonhomme et al. 2019)
- This paper
 - Added third source of heterogeneity: Managers (Bloom et. al 2013)
 - Used microdata from the Indian police

Research questions

- What determines productivity dispersion in policing?
 - Contribution of worker, manager and establishment heterogeneity

- Can we increase the total productivity of the police department by reallocating workers?
 - Gains from matching workers with different types of manager and establishment

Contribution

- Novel data: Used half a million webscraped crime reports to create the matched database of employment histories
 - Tracked the workers and managers across police stations
 - Matched the outcome of half-million cases in separate web scraping
 - Used time to submit final report/charge sheet as productivity measure
- New estimator: Extended standard model of how workers and firm contribute to productivity (Abowd et. al, 1999; Bonhomme et. al, 2019)
 - Added managers and the interaction between worker, manager and establishments
 - Taken two-step estimation approach
 - 1. Classification step (k-means clustering)
 - 2. Model estimation (finite mixture model)

Findings

- Worker, firm and manager heterogeneity present and important in determining productivity
- There are substantial complementarities between workers, managers, and establishment/police station
 - Low-type workers are 57% more productive when matched from low-type manager and less productive police station to high type manager and high-class police station
 - High type worker can increase productivity by 86% from better match
- If the current matching level raised using optimal matching rule (positive assortative matching)
 - \bullet Real location of workers could increase the aggregate productivity by 10%

Contribution

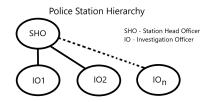
- Estimation using employer-employee matched datasets (Bonhomme et. al 2019; Abowd et. al 1999; Card et. al 2013)
 - Extended the model to three sided heterogeneity: worker, manager and establishment
 - Interaction between worker, manager and establishment
- Bureaucratic efficiency due to mobility (Rasul and Rogger 2014; Khan et al. 2019; Banerjee et al. 2012; Rauch 1995; Rauch and Evans 2000)
 - Reallocation of civil servants has large effects on productivity
- Do civil servants and managers enhance worker productivity (Chetty et al 2014; Lazear et al. 2015; Fenezia 2019)
 - Presence of worker and manager complementarities

Outline

- 1 Introduction
- 2 Background
- 3 Model
 - Identification
 - Estimation
- 4 Data
- 6 Results
- 6 Conclusion

Police structure in India

• Every district¹ is divided into field units called police stations



²A district is an administrative division of an Indian state

Crime reporting and investigation steps

- 1 A crime is reported in the police station
- 2 Details of crimes are documented in a First Information Report (FIR)
- 3 The case gets assigned to an Investigation Officer (IO) by the manager(SHO)
- 4 Investigation starts: collection of evidence, examination of witnesses, searching the property, identification of suspects, and arrests
- 5 Final report/charge sheet is submitted to a magistrate
- 6 Unsolved cases are closed after magistrate's approval and the rest go for trial

Time to submit final report/charge sheet as productivity measure

- Calculated as final report/charge sheet submission date FIR date
- Past research has used it and has advocated its use
 - Probability of case clearance falls with passage of time
 - "Cold case" phenomenon (Regoecz et. al 2008; Addington 2007)
 - Reflects the quality of investigation carried out by police (Iyer et. al 2012; Amaral et. al 2018)
- Time to submit final report/charge sheet has consequences for the case outcome (Law commision of India, 2015)
 - Leads to delay in criminal conviction
 - Acquittals because of delay in investigation

Time to charge sheet as productivity measure

- Delay in charge sheet filing has consequences on criminal justice outcomes
- Most prominent reason for delay in criminal conviction
 - 55% of the pending cases in courts are delayed at investigation stage (charge sheet filing)
 - Law commission of India, 2015 (random sample survey; N=1630 cases)
- Does delay in filing charge sheet adversely affect the prosecution of the case?
 - 100 % of the randomly sampled Judges answered yes (N=50)
 - Report of Bureau of Police Research and Development (BPRD) on increasing acquittals in India, 2013

Outline

- 1 Introduction
- 2 Background
- 3 Model
 - Identification
 - Estimation
- 4 Data
- 5 Results
- 6 Conclusion

Model

Heterogeneity is three sided

- N Workers (IOs) indexed by i
 - Discreet types: $\alpha_i \in \{1, ..., L\}$
- \bullet H Managers (SHOs) indexed by h
 - h_{it} : worker i at time t employed with manager h
 - Classes: $m_{it} = m(h_{it}) \in \{1, ..., M\}$
- J establishments (police station) indexed by j
 - j_{it} : worker i at time t employed in police station j
 - Classes: $k_{it} = k(j_{it}) \in \{1, ..., K\}$

Model

- Worker draws log productivity Y_{it} from a distribution that is characterised by α_i , m_{it} and k_{it}
- Conditional CDF of log productivity Y_{it} is (Complementarities)

$$Pr[Y_{i1} \le y | m_{i1} = m, k_{i1} = k, \alpha_i = \alpha] = F_{mk\alpha}(y)$$

• Proportion of type- α workers working with manager m and police station k: (Sorting)

$$Pr[\alpha_i = \alpha | m_{i1} = m, k_{i1} = k] = \pi_{mk}(\alpha)$$

• We will also know the transition probabilities

Assumptions

Worker moves: $s_{it} = 1$

$$\underbrace{m, k, \alpha}_{\text{period 1 }(Y_{i1})} \to \underbrace{m', k', \alpha}_{\text{period 2 }(Y_{i2})}$$

Assumption 1 (Mobility)

• Probability of moving **depends on** m, k and α but **not** on Y_{i1}

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

Assumption 2 (Serial independence)

• In T=2 workers draw productivity Y_{i2} from a distribution $F'_{m'k'\alpha}(Y_{i2})$ that **depends on** m', k' and α but **not** on m, k or Y_{i1}

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

Link to other models

- Models where state variables (α, m_t, k_t) are first order Markow
- Models where next period productivity is determined by the current state (Fenezia 2020)
- No human capital accumulation or learning
- Similar to labour market models where wages are the outcomes of match between worker types and firm classes (Card et. al 2013; Lenz, Piyampromdee, and Robin 2020)

Outline

- 1 Introduction
- 2 Background
- 3 Model
 - Identification
 - Estimation
- 4 Data
- 6 Results
- 6 Conclusion

Identification using manager and establishment classes

• Distribution of job movers using assumption 1 & 2

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

- $F_{mk\alpha}(y_1)$: CDF of log-productivity in period 1
- $F'_{m'k'\alpha}(y_2)$: CDF of log-productivity in period 2
- $p_{mm',kk'}(\alpha)$: probability distribution of job movers of α_i types between manager and firm classes

Identification using manager and establishment classes

• Distribution of log-productivity in period 1

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

- $\pi_{mk}(\alpha)$: distribution of α_i worker in manager class m and firm class k
- Parameters to be identified
 - $F_{mk\alpha}$ and $F'_{m'k'\alpha}(y_2) \forall (\alpha, m, k)$
 - $p_{mm',kk'}(\alpha) \ \forall \ \alpha \ \text{and} \ (m,m') \times (k,k') \ \text{pairs}$
 - $\pi_{mk}(\alpha) \, \forall \, (\alpha, m, k)$

Identification: intuition

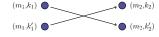
•
$$M = 2$$
 and $K = 2$

Workers move within same firm class

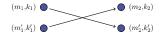
$$(m_1,k_1) \bigoplus (m_2,k_2)$$

$$(m_1',k_1) \bigoplus (m_2',k_2)$$

Moves within manager class



Moves across manager and firm classes



- Assumptions
 - Presence of connecting cycles in graph (implies connectedness)
 - Asymmetry in worker compositions while moving
 - Proof: Worker, manager and firm interaction effects identified

Identification proof

• Proof like below upcoming:

Theorem 1. Let T=2, and let Assumptions 1 and 3 hold. Suppose that firm classes are observed. Then, up to labeling of the types α , $F_{k\alpha}$ and $F_{k'\alpha}^m$ are identified for all (α, k, k') . Moreover, for all pairs (k, k') for which there are job moves from k to k', $p_{kk'}(\alpha)$ is identified for all α , for the same labeling. Lastly, the type proportions $q_k(\alpha)$ in the first period are all identified, again for the same labeling.

Identification of manager and firm classes

• Distribution of log-productivity in manager h and firm j does not depend beyond it's manager class m and firm class k

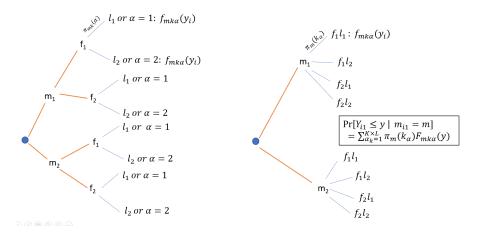
$$Pr[Y_{i1} \le y_1 | h_{i1} = h, j_{i1} = j] = \sum_{\alpha=1}^{L} \pi_{mk}(\alpha) F_{mk\alpha}(y_1)$$

• Combining the firm and worker into $K \times L$ classes: Nested BLM

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

Identification of manager and firm classes:Intuition

• Example: (No. of) manager classes (M) = 2, firm classes (K) = 2 and worker types (L) = 2



Outline

- 1 Introduction
- 2 Background
- 3 Model
 - Identification
 - Estimation
- 4 Data
- 6 Results
- 6 Conclusion

Step 1: estimating manager and firm classes

- Manager classes estimated by combining the firm and worker to $K \times L$ classes
- \bullet Managers are clustered in M groups
- Manager classes are identified (BLM 2019) and recovered using k-means clustering below

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

- Minimize the least square error of the within-cluster (weighted k-means)
 - \hat{F}_h is the empirical CDF of log-productivity of manager h
 - $H_{m(h)}$ are CDF's of the manager classes

Step 1: estimating firm classes

- Similar approach:
- Firm classes estimated by combining the manager and worker to $M \times L$ classes
- Threat to identification when two manager or firm classes have same productivity distribution

Step 2: estimating model parameters

- In step 1, I estimated the manager and firm class membership : \widehat{m}_{it} and \widehat{k}_{it}
- Log likelihood of job movers (L1)

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

- $f_{mk\alpha}(y_1; \theta_f)$ and $f'_{m'k'\alpha}(y_2; \theta_{f'})$: first and second period productivity
- $p_{mm',kk'}(\alpha;\theta_p)$: job movers probability
- Parameter vectors : $\theta_p, \theta_f, \theta_{f'}$

Step 2: estimation method

- I estimate $\widehat{\theta}_p, \widehat{\theta}_f, \widehat{\theta}_{f'}$ by maximising **L1**
- I use EM algorithm for mixture model estimation (Dempster et al., 1977)
- Used Conditional maximisation algorithm for fast convergence (Meng and Rubin, 1993)
 - Log-normal specification: Assumed mean and sd (θ_f) and then find optimal values of p using classical EM
 - Conditional on last parameter values (p), find the optimal values of θ_f in a sequential manner
 - Last step is unconstrained optimisation
- Estimated parameters with multiple initial values and then chosen the best model to reach global maxima. why?
 - Because EM algorithm converges to local maxima

Step 2: estimating model parameters

• Log likelihood of workers in period 1 (L2)

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

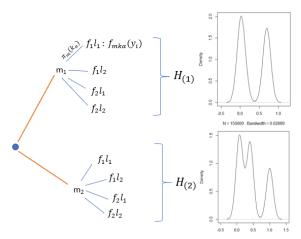
- I estimate $\widehat{\theta}_{\pi}$ for all worker types in each manager class m and firm class k by maximising **L2**
- Remember earlier we used a computationally intensive modified EM algorithm
- Here I use linear programming as $f_{mk\alpha}$ is known

Experiments with simulated data

- Assumed the number of manager classes (M=2) and number of firm classes (K=2) and workers types (L=2)
- Simulated the data using arbitrary parameter values : $\theta_p, \theta_\pi, \theta_f, \text{ and } \theta_{f'}$
- Also simulated the manager and firm id's from the discrete classes
- Used the simulated data as input to my **two step three-sided** estimator
- Monte Carlo simulation technique with means of estimated parameters reported in the following slides

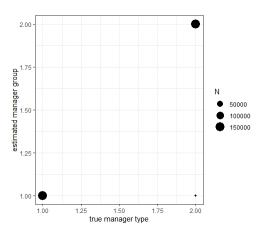
simulated data: recovering manager classes

- Manager classes estimated by combining firm and worker classes together. $(K \times L \text{ or } 2 \times 2)$
- k-means clustering aim to group the manager id's into 2 classes



simulated data: recovering the manager classes

• Low misclassification rate (< 1%)



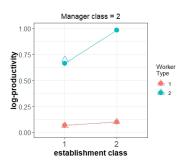
• Estimated firms classes (link)

Simulated data: estimating model parameters

• Estimated of θ_f : means only for $f_{mk\alpha}(y_1; \theta_f)$

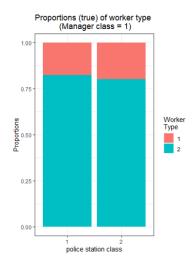
Figure: Estimates of Step 2: Model parameters - Bold(circles) lines are true parameter values and dotted (triangles) are estimated values

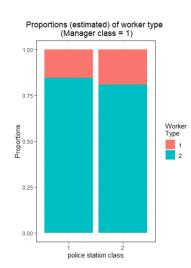




Simulated data: estimating model parameters

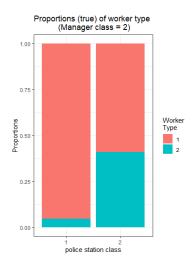
• recovering $\pi_{mk}(\alpha)$: worker proportions

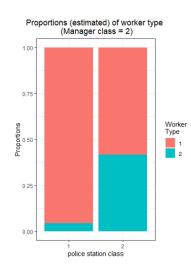




Simulated data: estimating model parameters

• recovering $\pi_{mk}(\alpha)$: worker proportions

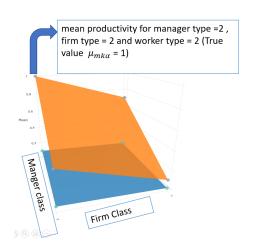


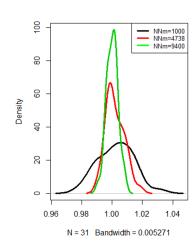


Asymptotic properties

- Asymptotic normality of the estimator
- When sample size tends to infinity, under following properties shown in BLM 2019
 - 1) Misclassification error in estimated manager and firm classes approaches zero
 - 2) Estimation in Step 2 behaves like that of Maximum liklihood estimator
- Asymptotic properties shown using Monte Carlo simulation: computational approach
- Standard errors using the parametric bootstrap method

Asymptotic properties: Monte Carlo simulations





Recap: Estimation

Step 1

- Recover manager and firm classes
 - Recover manager classes: combine firms and workers then run BLM step 1 (clustering)
 - Recover firm classes: combine managers and workers then run BLM step 1 (clustering)



Step 2

- Estimate productivity distributions
 - Three-sided estimator with productivity distribution $f_{mk\alpha}$ (T=1) and $f_{mk\alpha}^{m}$ (T=2)
 - Mixture model where workers move across manager and firm classes
- Estimate worker proportions : $\pi_{mk}(\alpha)$

Outline

- 1 Introduction
- 2 Background
- 3 Model
 - Identification
 - Estimation
- 4 Data
- 5 Results
- 6 Conclusion

Data: Crime reports

• I use all the web scraped crime reports for the state of Haryana between $2015-2018^2$ (N=472,082)



¹Ongoing collaboration with Bhatia, Haseeb and Joshi 2019

Crime outcomes

- In another web scraping, I match all reports to their outcomes
- Time taken to submit final report/charge sheet as productivity measure : Charge sheet date FIR date

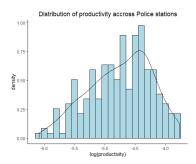


Data: Police productivity in Haryana

0

Location of police stations





Data: Inferring job mobility in Haryana police

- Created Worker-Manager-Establishment matched dataset
- Data mappings:
 - Investigative officer (IO) is worker (count: 9581)
 - Station Head officer (SHO) is manager (count: 1007)
 - Police station is establishment (count: 282)

Worker's (id=303) job mobility across police stations



Outline

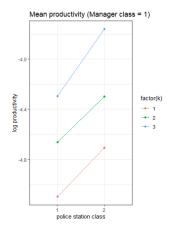
- 1 Introduction
- 2 Background
- 3 Model
 - Identification
 - Estimation
- 4 Data
- 6 Results
- 6 Conclusion

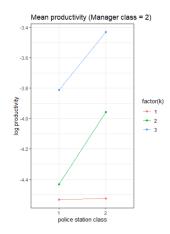
Results

- Estimated the model assuming
 - No. of worker types: L=3
 - No. of manager classes: M=2
 - No. of police station classes: K=2
 - Gaussian finite mixture model
- I estimated the productivity distribution and proportion of job movers of different types
- Finally, I estimated the worker proportions
- Estimation involves multiple starting points (typically 100) done using parallel computation because Maximum likelihood estimation of finite mixture models is often subject to local maxima
- The results are subjected to change as done using 2 starting point on local machine! (Orac was on maintenance)

Results

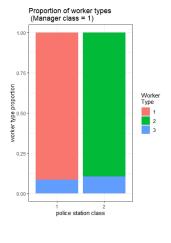
- Evidence of worker-manager-firm heterogeneity
- Complementarity

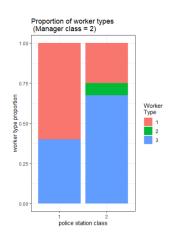




Results

• Sorting patterns shows scope of benefit from re-allocations





Variance-Covariance decomposition

- Wage literature: Variance decomposition (Abowd et al. 1990, Card et al. 2013)
- Classical way to relate variance of log productivity and heterogeneity (Fenezia 2019)
- Linear projection of non linear model: BLM 2019

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

0 -

Variance-Covariance decomposition

0

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

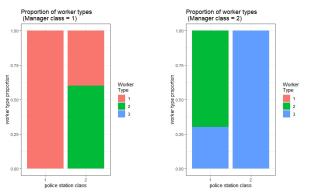
Notes: Linear regression $Y_{it} = \alpha_i + m_{it} + k_{it} + \epsilon_{it}$ on the estimated values of model

Revisiting research question 2

- What determines productivity dispersion in policing?
 - Contribution of worker, manager and establishment heterogeneity
- Can we increase the total productivity of police department by reallocating workers?

Counterfactual reallocations

• Allocating worker using pure assignment (Eg. PAM below)

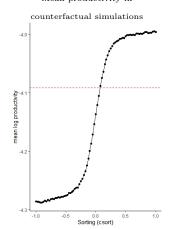


- I generate intermediate sorting patterns using an algorithm
 - Choose the proportion of workers in each type
 - Randomly allocate each worker type to manager-establishment class

Optimization results

• 9.2% increase in productivity of police department by reallocating police officers

Mean productivity in



Estimates of productivity at optimal matching rule

Reallocation exercise (×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

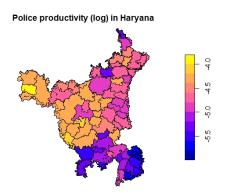
Conclusion

- Uncovered the heterogeneous effects of workers, managers and police stations on productivity using:
 - Novel microdata from the Indian police
 - New estimator (two-step approach)
- Findings
 - There are substantial complementarities between workers, managers and establishment
 - Gains from **reallocation** are possible by matching **high type** worker with **high type** of manager and police station

Appendix

Data: Police productivity in Haryana

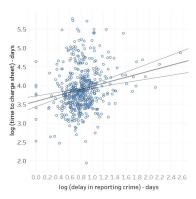
• Variation in productivity across sub-districts



Managers/SHOs matter

- SHOs have authority to decide when to record crime as FIR
 - Local manager (SHOs) cooperation is necessary for the implementation of police reforms (Banerjee et. al 2019)

Plot showing the manager (SHO) wise relationship between delay in reporting crime and time to charge sheet (2017) t-value =4.75



Mean transfer rates

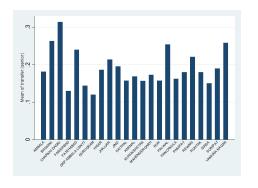


Figure: Mean transfer rate of police officers (SHO) district wise

Internal transfer orders (sample)



Figure: Internal transfer order of police officers



Map of Police stations



Figure: Location of police stations (Circle)

Data: Crime rate

• Haryana has high crime rate - 802 per 100,000 population when compared with India's average of 302

