# Location Preferences of Women Police Officers in India*

Palak Suri†

March 11, 2025

[Please download the most recent version here]

**Abstract**

The police administration of Madhya Pradesh, India, recently adopted a centralized recruitment system using a deferred acceptance (DA) mechanism to match new recruits with police units. Using administrative data on location matches and the DA mechanism's theoretical properties, I estimate police officers' location preferences. Women value proximity to home more than men, equivalent to 12% vs. 7% of base pay. Both prefer areas with better infrastructure, but women also prefer flexible roles and safer locations. Scarcity of women officers relative to demand under the DA mechanism results in spatial mismatches, leaving the regions that need them most underserved.

*I thank the police administration of Madhya Pradesh for sharing the data used in this paper. I am grateful to Chenyu Yang, Sergio Urzua, JP Chauvin, Jessie Handbury, Brad Humphreys, Adam Nowak, Dan Grossman, Kole Reddig, Qiyao Zhou, Maureen Cropper, Guido Kuersteiner, Sonalde Desai, Suzanne Bellue, Yue Yu, Ian Herzog, Camilo Acosta, Luis Baldomero-Quintana, and Yongjoon Park for their helpful comments and discussions. I am especially thankful to police officers Mr. Anil Kumar, Mr. Sandesh Jain, and Mrs. Anju Gupta for providing important insights about constable recruitment and the duties of women police. I am also thankful to policy specialists and researchers who graciously spared their knowledge and time for this project including Avanti Durani, Pritika Hingorani, Priya Vedavalli, and Neha Sinha.

†palak.suri@mail.wvu.edu; Department of Economics, West Virginia University

# 1 Introduction

Women police and policies that increase the salience of women police are effective at improving reporting and reducing the incidence of crimes against women (Perova and Reynolds (2017), Miller and Segal (2018), Sviatschi and Trako (2024), Amaral et al. (2018), Siwach (2018)). In India, women police officers are a scarce resource but an essential component of the police human capital stock as they perform tasks that cannot be substituted by their men counterparts. The presence of women police in a location is, therefore, an important amenity that can influence workers' geographical sorting behavior (Diamond and Gaubert (2022)) and, consequently, aggregate spatial inequalities (Diamond (2016)). Given that women constitute only 12.3% of the total police force (BPRD (2023)), it is imperative to investigate the determinants of their spatial distribution.

In this paper, I estimate the location preferences of women police officers and trace their implications for the spatial distribution of women police in Madhya Pradesh (M.P.), a state in Central India with a population of over 72 million (Census, 2011).

To create a thick and competitive labor market for police officers, in 2012, M.P. adopted a centralized mechanism to recruit and match newly recruited police constables to police units using a deferred acceptance (DA) mechanism. Candidates are assigned a police unit, based on their stated preferences, home district, merit rank in an entrance exam, affirmative action policies, and police units' requirements. The matching process is outsourced to a small firm to ensure confidentiality and fairness, and preference lists submitted by the candidates are not shared with the police administration. Leveraging administrative data on location assignments, vacancies, and candidate information with the theoretical properties of the DA mechanism, I identify the location preferences of police officers. Specifically, the stability property of the DA mechanism allows me to convert the two-sided matching problem into a revealed preference problem and estimate preferences using discrete choice models.

Preference estimates reveal a strong distaste for distance from home. Practically all positions require candidates to be assigned outside their home districts, and my results indicate that the average woman candidate would be willing to accept a 12% reduction in base pay to be in a district closest to their home. The corresponding number for men is 7%. Women also prefer flexible positions with fixed hours and routine desk duties, while the evidence is weaker for men. All candidates prefer areas with a better formal economy and better infrastructure as measured by health and education facilities and urbanization levels. Women additionally have a strong preference for areas with fewer reports of sexual assaults, but the same is not true for men. They would be willing to accept a 6% reduction

in base pay to go from a district that is at the 95$^{th}$ percentile to one at the 5$^{th}$ percentile.

Due to a scarcity of selected women candidates and the non-uniform spatial distribution of their origin districts, the existing matching mechanism leads to inequitable provisioning of women police services across districts. Addressing this is important from a public finance perspective since large districts with a better formal economy are better able to attract women police than small districts, and therefore, they benefit disproportionately from the thicker market for women police created by centralized recruitment. Women police are instrumental in investigating and deterring crimes against women, and the absence of their service is likely to exacerbate existing spatial inequalities in amenities (Glaeser et al. (2001)). Although not explored in this paper, this can also have political consequences (Singh (2019)).

The benefits from an equitable distribution of women police in this context are difficult to quantify. In the context of Madhya Pradesh, Sukhtankar et al. (2022) using an RCT show that introducing women help desks in police stations lowers barriers to justice for women. Amaral et al. (2018) estimates the impact of special women police stations on crime reporting and finds similar results. I find suggestive evidence that the presence and salience of women police lead to more women applicants in the future, and possibly improve women's labor force participation by improving gender norms. Given India's low and stagnant 27% women's labor force participation rate and the influence of gender norms in determining women's work decisions (Jalota and Ho (2024), Heath et al. (2024)), this is an urgent policy area for the country.

To compute the monetary cost of equitable distribution in this setting, I use estimated preferences to simulate outcomes from a counterfactual mechanism that corrects this inequity by accounting for the ex-post shortfall of recruited women. In the counterfactual, the local demand from each police unit that is input into the matching algorithm is adjusted in proportion to the ex-ante demand such that the ratio of women positions left vacant post-matching is uniform across districts. This adjustment creates a welfare loss for selected women officers– the average affected candidate would need to be paid 15% of the basic pay to compensate for the welfare loss. The total annual cost of this counterfactual mechanism is equivalent to hiring approximately 125 additional women officers, merely 10% of the overall shortfall of women police relative to their demand. Furthermore, this alternative mechanism retains all the desirable properties of a DA mechanism, albeit at the cost of added algorithmic complexity.

This paper adds to the literature examining worker preferences and their implications in specific labor market contexts that facilitate the provisioning of essential services while

facing unique labor supply shortages. To the best of my knowledge, this is the first paper to estimate location preferences of police in India. Ba et al. (2021) is the only other paper in the literature that investigates police officers' preferences for neighborhood characteristics in Chicago. Boyd et al. (2013) investigates teachers' and schools' preferences for each other. Falcettoni (2018) investigates physicians' location preferences and their consequences for physician shortages in rural areas. My results can inform policy on optimal financial and non-financial incentive structures for public personnel (Dal Bó et al. (2013)). This is an active policy area that is still understudied (see Finan et al. (2015) for a review).[1]

Knowledge about workers' preferences can be used to explain geographical sorting and migration (Black et al. (2014), Diamond (2016), Couture and Handbury (2022), Jia et al. (2023)), gender wage gap (Liu and Su (2024), Le Barbanchon et al. (2021)), job mobility trends (Saks and Wozniak (2011)), structural wage differences (Card et al. (2018), Lavetti and Schmutte (2016), Sorkin (2018)), and the variation in the returns to education and skill (Black et al. (2009)). I add to this work by providing women workers' preferences for job and location attributes. The centralized mechanism in my context has a key advantage, allowing the use of a revealed preference framework similar to Agarwal (2015) and Fack et al. (2019) for better identification of preferences than the existing literature. Even though the setting of my paper is of women police in Madhya Pradesh, my results are consistent with worker preferences in other settings, suggesting their generalizability.

Finally, this paper adds to the growing empirical mechanism design literature highlighting an important limitation of the candidate-proposer-DA mechanism in the context of many-to-one matching when the proposer side of the market is thin. The scarcity of women police leads to the matching mechanism creating a spatial mismatch between women police and the demand for them due to wages being uniform and fixed. Various matching mechanisms exist around the world across markets including education (Abdulkadiroğlu and Sönmez (2003)), kidney transplants (Roth et al. (2004)), armed forces (Sönmez and Switzer (2013)), civil services (Thakur (2021)), and medical residents (Roth (2003)), among others. Economics research on matching mechanisms has motivated a series of policy changes across these contexts (Niederle et al. (2008), Pathak (2017)). This paper is a first step toward understanding and informing policies that promote women's safety and support their broader participation in the economy.

The paper is organized as follows. Section 2.1 describes the institution and the mechanism. Sections 2.2 discuss the theoretical properties of the mechanism, which are leveraged

[1]The context of centralized police provisioning also draws a parallel with the literature on fiscal federalism (Oates (1999)) as the provisioning of women police is made centrally, whereas benefits of women police are experienced locally.

for identification. Section 2.3 describes the affirmative action rule and its implications for identification. Section 3 describes evidence highlighting the spatial inefficiency. Section 4 describes the theoretical model used for the identification of preferences. Section 5 describes the data and Section 6 the estimation, parametric identification concerns, and the results. Section 7 discusses the implications of the results and an alternative mechanism and Section 8 concludes.

# 2 Context

This section explains the institutional rules of police recruitment and unit assignment, which are leveraged for the identification of individual preferences.

## 2.1 Institution

In 2012, Madhya Pradesh (M.P.) in India started recruiting constables for the state police force through a centrally conducted examination testing candidates' basic general knowledge, algebra, language skills, and physical fitness. Selected candidates are then requested to submit a rank-ordered list of their preferences ranking the 86 police units in the state.[2] Women candidates are ineligible for 23 of the units for armed police and men candidates are ineligible for the only women armed police unit. Selected candidates undergo training at the police academy and are subject to a final medical test before being assigned a police unit for an initial period of 8-10 years.[3]

Assignments are made based on candidates' reported preferences in order of merit, conditional on there being vacancy for categories in which candidates are eligible- as in a serial dictatorship mechanism with merit-based priority. Candidates' categories are determined by whether they benefit from any affirmation action policy. Those that do are eligible for more than one category. Number of vacant posts at each police unit is not known to the candidates. I will argue below that in this context, the mechanism is strategy-proof and the rank-ordered lists submitted by the candidates likely reflect their true preferences.

Broadly, there are three different types of police units– civil police units, which are the most common type and involve field work and interactions with the public regularly while being in a fixed geographical location; armed police or railway police units, where work location and type of work is more uncertain; and communications, cyber security or

[2]There are other posts for which recruitment is done alongside that of general duty police constables like home guard, driver, cook, head constable computer, etc.

[3]Police officers can always submit a formal application requesting a transfer to a different unit and several such requests are made each year.

investigation departments with routine office duties and predictable hours. Postings with routine office duties for women are only available in four districts, of which only the capital of Madhya Pradesh, Bhopal has a sizable number of vacancies. As a rule, police constables are not posted in their home district if they are assigned a duty that requires interaction with the public. Vacancies are recorded with the police departments for each candidate category separately.

The allocation mechanism follows the Gale-Shapley deferred acceptance (DA) algorithm with the candidate as the proposer. Each candidate's priority is determined at each police unit independently based on their merit scores in the entrance examination and the relevant category-specific vacancy at the unit. The category of a selected candidate is independent of the requirements at a specific police unit but is related to the overall type-specific requirements in a given year. Therefore, a category can be oversubscribed at one unit but undersubscribed at another. However, the total number of candidates assigned to a unit never exceeds the total number of vacant positions at that unit.

The matching algorithm works as follows:

*Round 1:* All selected candidates apply to their foremost choice of the police unit. Each unit rejects candidates with the lowest priority as per their merit and reservation category in excess of the unit's capacity and temporarily holds the remaining candidates.

*Round 2:* All students rejected in the previous round apply to the next best police unit on their list of preferences submitted to the computer. Each unit pools all the applicants together and rejects the ones with the lowest priorities in excess of capacity while temporarily holding the rest.

*Round k:* Same as *Round 2*.

The algorithm terminates when no rejections are issued. Candidates are matched to the last police unit holding them.

## 2.2 Properties of the Deferred Acceptance (DA) Mechanism

The DA mechanism is used in many real-world contexts, e.g., school, college, and branch allocation in the military and the medical residency match. The theoretical literature investigating the properties of matching mechanisms has established that the allocation resulting from a candidate-proposing DA mechanism is not necessarily the best outcome for the other side of the market (Roth (1985)). While the DA mechanism is still desirable in the context of schools and colleges where the preferences of institutions are not relevant for policy, since law enforcement is a quasi-public good, it is imperative to understand the

extent to which the final distribution of assignments is aligned with the requirements of the citizens. My analysis tests this.

The following two desirable properties of the DA mechanism form the basis of my analysis in identifying the location preferences of police officers.

First, Roth (1984) and Roth (1985) show that the DA mechanism in a many-to-one matching context produces a stable match. In my context, stability means that after the final match, there is no candidate-police unit pair that would strictly prefer to match with each other over their given assignments. That is, no candidate can be made better off without making other candidates worse off while also simultaneously respecting the priorities of the police units. This property will form the basis of my identification strategy.

Second, truth-telling by candidates under the candidate-proposing DA mechanism is a weakly dominated strategy (Dubins and Freedman (1981) and Roth (1982)). This means that the rank-ordered lists that the candidates submit are a close reflection of their true preferences. Since I only observe the preference ranking of the matched police unit alternative, I use this property in my analysis to refine the set of feasible alternatives identified based on the assumption of stability for a subset of candidates (discussed in Section 4). In similar studies where information on rank-ordered lists is available, this property forms the basis of identification strategy, however, Fack et al. (2019) argues that in most empirical contexts, it is more credible to rely on the stability assumption.

## 2.3 Affirmative action rule

All government departments in India follow affirmative action policies laid out in the Constitution and enforced by the Supreme Court of India. There are two types of reservation in this context- horizontal and vertical. There are four vertical categories- Unreserved category (also known as the general category), SC, ST, and OBC categories. Within each vertical category, there are three horizontal categories- open (or gender-neutral), women, and ex-servicemen. Everybody, regardless of their caste membership or sex is eligible to be admitted under the unreserved open category based on merit but only members of SC, ST, and OBC categories can be admitted under their respective vertical categories. Within each vertical category, women of only that category are eligible to be admitted against the women's quota. For instance, an OBC woman can not be assigned a seat meant for an unreserved category woman but she can be assigned an unreserved open category seat. The rules are summarized in Figure 1.

I use these affirmative action rules to determine individual candidates' eligibility for the police units. Sonmez and Yenmez (2019) discusses the possibility of false reporting of

preferences and unstable outcomes resulting from the affirmative action policy. The rest of this section argues why this is not a concern in my setting.

The way in which reservation for women is implemented may give rise to situations where a higher-ranked OBC woman loses her seat to a lower-ranked unreserved category woman which is inconsistent with the original ideology behind reservation and has become a cause for several litigations. This system and the issues are discussed in detail in Sonmez and Yenmez (2019). The authors have termed this situation as justified envy. In the matching literature, the term justified envy is used to describe a situation where an individual prefers a unit and is in turn preferred by the unit over another individual who has been assigned the unit by the system. In this context, the government was mindful of this possibility and maintained that such inefficiencies would be allowed in the system.[4] If units have accepted this affirmative action rule, then it becomes a part of their preference to adhere to it. Hence, given the unconventional preferences ascribed to the units by the affirmative action rule, the question of justified envy does not arise in my context.

However, it is still important to understand the implications of this policy specific to this context. The caste category is revealed by the candidates to the system at the beginning of the recruitment process since there are other benefits associated with reservation in the form of relaxation in physical requirements, exam score cutoff, subsidized application fee, etc. For some women candidates, it may be ex-post optimal to not avail of the benefits of the reservation policy and try to get assigned under the unreserved category instead. This is not a practical possibility because caste category is revealed at a time when the distribution of merit is not known and candidates never find out about unit-specific type vacancies. Since the candidate pool is large and it is not possible to predict ex-ante which individual would be affected by this, it is unlikely that candidates are strategically misreporting their preferences.

In the absence of affirmative action policy requirements, a pareto improvement in outcome is possible in terms of satisfaction of candidate preferences. A simple example adapted from Sonmez and Yenmez (2019) shows this. Suppose there are only two units A and B and an SC category woman with merit rank 17 who prefers Unit A to Unit B. Suppose that Unit A has 10 seats for the open category (no reservation), 3 for women of unreserved category, 4 gender neutral seats for SC, and 2 seats for SC women. Suppose that at Unit A, the cutoff for an open category (no reservation) assignment is 10, the merit rank of the last assigned candidate under the unreserved woman category is 25, and there is 1 seat empty under the unreserved woman category but not under the SC category. Despite the vacant spot at Unit A and the merit rank of the SC candidate being lower than the last unreserved

[4] http://peb.mp.gov.in/rulebooks/RB_2016/PCT_2016_Rule_Book_Revised.pdf

category woman assigned to Unit A, the SC candidate would be assigned Unit B. Assigning her Unit A would be a clear pareto improvement unless the police Unit A has a preference for empty seats over violating the affirmative action policy rule. Since this rule is written in the assignment mechanism rules, the only recourse the candidate has is to challenge the rule in a court of law. Note that I did not even evoke justified envy here as in Sonmez and Yenmez (2019) which would have been relatively difficult for the candidate to detect since the results of the assignment are not revealed publicly.

# 3 Descriptive Evidence

The Police Administration adopted the deferred acceptance (DA) mechanism for centralized assignments likely due to the incentives it provides to the workers. Table 1 shows the percentage of candidates in the four recruitment rounds between 2012-2016 who got their top 5 preferred units. Most candidates submit their preference rankings to the system and those that don't are administratively assigned after all candidates have been placed by the system. In 2016, only 2% of the candidates did not submit their preferences. Due to the scarcity of women candidates, around 90% of women get assigned their top five choices. While there is some variation in this number across the recruitment rounds, it is consistently smaller for men due to a greater competition for positions for which men are eligible. [5]

What determines whether someone will get their topmost-ranked location or a lower-ranked one? 60% of the variation in the order of the assigned unit in the candidates' rank-ordered list across recruitment rounds is explained by gender, caste category, type of posting, distance between residence and work districts, and fixed effects for residence and work districts. An additional 6 percentage points is explained by five quantiles of merit ranking within gender and recruitment round, and fixed effects for the gender and recruitment round-specific match between resident and work districts. This highlights the importance of candidates' residence locations and vacancies in determining the assignment outcome and a potential major weakness of the existing mechanism given the scarcity of selected women candidates.

Theoretically, the DA mechanism only favors the candidates and not the police units (Roth and Sotomayor (1992)). Given the high number of initial applicants and selected candidates that are men, this system does not constrain any police units from filling vacancies for men. However, due to the scarcity of selected women candidates, many positions

[5] Table 1 also highlights the importance of careful identification of the set of feasible options in estimating location preferences.

across police units remain vacant after the assignment procedure. Since the spatial distribution of selected women candidates is not uniform, this can result in a spatial inequity in which police units' demand for women candidates is met. Figures 2 and 3 show the spatial distribution of women applicants and selected women candidates from the 2016 recruitment round. Figure 4 shows the spatial distribution of the ratio of women vacancies that remain unfilled.

Table 2 shows the determinants of the ratio of women vacancies that remain unfilled using OLS.[6] A high proportion of vacant positions is positively associated with a greater number of originally posted vacancies. Civil police unit type, a predicted amenity index, and the number of eligible applicants within 100 km of the district are all negatively correlated with the vacancy ratio. There is also suggestive evidence of a positive association between assaults per 1000 women, a proxy for crimes against women and the vacancy ratio.

These results highlight an important spatial mismatch created by this police assignment mechanism. Police units located in districts with better formal economy and infrastructure are more likely to attract scarce women candidates. Since the presence of women police helps deter crimes against women (Miller and Segal (2018) and Sviatschi and Trako (2024)), this is likely to perpetuate the existing inequality among districts unless the number of women applicants rises. There is also suggestive evidence that the presence of women constables and head constables is associated with a greater number of future women applicants. Therefore, it becomes important to understand women police officers' location preferences and their spatial equity implications. The estimated preferences can also be used to understand the consequences of alternative mechanisms.

# 4 Model and Identification

Suppose there are $N$ candidates $i = \{1, 2, ..., N\}$ with types $t \in \{$UR Female, UR Open, OBC Female, OBC Open, SC Female, SC Open, ST Female, ST Open$\}$ defined in Figure 1, and $J$ police units $j = \{1, 2, ..., J\}$. Suppose that each candidate has a utility function $u_{ij}$.

Based on relative values of $u_{ij}$ across $j$ for an individual, he submits a rank-ordered list to the system, denoted by $r_i$. Each police unit $j$ has type-specific vacancies, denoted by $v_{jt}$ in accordance with the various affirmative action policies. The mechanism assigns every candidate a priority rank $p_{ijt}$ based on his merit rank and whether he is eligible for assignment

[6] 2016 is the only recruitment round for which initial vacancy listings were posted for each gender and caste category. Vacancy information was not available for seven units including the police training academy, Crime Investigation Department, Telecommunications, Special Branch, and a women-only special armed forces unit, all of which are non-civil-police postings. Assignments to these units are observed, making it difficult to assess the determinants of post-assignment vacant positions.

under any affirmative action policy. Recall that the mechanism is a candidate-proposing deferred acceptance algorithm where candidates propose to their most preferred units and the units accept proposals in order of priority ranks of each proposer conditionally. In the second step, rejected candidates propose again and units select candidates based on their priority rank from the initial pool of conditionally accepted proposals as well as the new pool. The process continues until all candidates are matched to a unit.

The mechanism produces a match $\mu(u_{ij}, p_{ijt}, v_{jt}) = j$. In equilibrium, cut-offs for priority rank are determined for each police unit. Priority rank is simply the vector of merit rank for each candidate category. The cut-off defines the merit rank threshold below which applicants were rejected by the unit. Denote these cut-offs by $C(\mu)$. This is a random vector determined by utility shocks $\epsilon$.

The stability of the DA mechanism implies that the candidates are assigned their most preferred available alternative as per the submitted rank-ordered lists. This outcome is the same as that arising from a discrete choice model where each candidate would choose the best possible option from his personalized set of feasible alternatives. Recall that an alternative is feasible if the candidate is eligible for assignment to a particular unit based on the cutoff score of the unit and his priority rank. The set of feasible alternatives can be written as $F(p_{ijt}, C(\mu))$, or $F_i$ for short. It is generally not possible to know the preference relation between the chosen alternative and an alternative for which the candidate was not eligible when relying only on observed matches (Agarwal and Somaini (2019), Fack et al. (2019)). However, in cases where the candidate's assigned unit was first or second in his rank-ordered list, the set of feasible alternatives can be expanded to include all possible units.

Even though in theory the cut-off in any given year will depend on the pool of candidates for that year, the large size of the market and stability imply that cut-offs and the individual-specific feasible set can be treated as exogenous (Fack et al. (2019)). It is possible to account for equilibrium effects, and doing so would lead to efficiency gains but I ignore that due to the complication introduced by the affirmative action rule in determining the cutoffs and due to incomplete information on category-specific vacancy information, which is only available for the 2016 round.[7]

**Utility Function:** A police candidate $i$ selects a police unit $j$ located in district $d$ out of the set of feasible alternatives based on their preferences for different amenities available in each location including distance from hometown, and economic and physical infrastruc-

[7] In the context of school choice in Paris, Fack et al. (2019) notes that the efficiency gains were negligible from accounting for the endogenous determination of priority ranks.

ture of each district. I assume a linear additive random utility function $u_{ijd}$.[8]

$$u_{ijd} = \beta_i * X_{id} + \zeta * K_{jd} + \gamma * Z_d + \nu_j + \epsilon_{ijd} \tag{1}$$

$X_{id}$ represents the distance between the headquarters of police unit $j$ district and the centroid of candidate $i$'s home district. $K_{jd}$ represents the two types of police units based on job flexibility. $Z_d$ represents district amenities reflective of the formal economy, urbanization, health and education infrastructure, crimes against women, and a measure of wages adjusted for housing rent allowance. $\nu_j$ represents mean preferences for each police unit. Since only a limited number of districts have more than one police unit, I estimate mean preferences for districts instead of units. In some specifications, I estimate mean preferences for police zones instead of districts to allow for a greater variation in amenities across locations.[9] $\epsilon_{ijd}$ is an i.i.d. random utility shock assumed to follow a Type-I extreme value distribution. $\beta_i$ denotes individual specific taste for features in distance. $\zeta$ represents average preference for a civil unit posting. $\gamma$ represents average preferences for district amenities that are common across candidates. In the main specification, I model $\beta_i$ as a function of observable individual traits, denoted by $E_i$ including age, education, marital status, and rural place of origin.

$$\beta_i = \bar{\beta} + E_i * \beta \tag{2}$$

The probability that an individual $i$ is matched with unit $j$ is given by

$$Pr\big(j = \mu(u_{ijd}, p_{ijt}, v_{jt})\big) = argmax_{j \in F_i} u_{ijd} | X_{id}, K_{jd}, Z_d, E_i, F_i; \beta_i, \zeta, \gamma, \nu_j \tag{3}$$

**Identification assumptions:** Assuming a correct utility function specification, identification of preference parameters relies on the following two assumptions.

1. Given candidate-level observables, preferences and priority indices are independent and no individual can manipulate the cut-offs that determine priority indices. That is, $p \perp \epsilon | X, Z, K, E$.

2. Each candidate's preferences are independent of their own set of feasible alternatives, conditional on observables. That is, $F(p, C(\mu)) \perp \epsilon | X, Z, K, E$.

Since the total number of candidates and vacancies is high and it is not possible for candidates to predict unit-specific vacancies or influence the cut-offs, both of these assumptions are easily satisfied. Each candidate is a very small part of this market. Additionally, a

[8] Since I am interested in modeling location preferences, $u_{ij}$ is henceforth denoted as $u_{ijd}$.
[9] Empirical concerns are discussed in Section 6.

candidate's eligibility is determined by the preferences submitted by candidates with a higher priority order. By assumption, candidates' preferences are independent of each other and therefore, assumption 2 holds.

The above arguments allow me to transform the matching problem into a discrete choice problem. Similar to Fack et al. (2019), as long as the individual-specific choice sets are exogenous (Assumption 2 above), this transformation is valid. Identification concerns related to the parametric identification of preferences are discussed in Section 6.

Units differ in their requirements for candidates of different categories and since these are not known in detail to the candidates at the time they submit their rank-ordered list to the system, the set of feasible alternatives can be considered as exogenous even though it is determined in equilibrium. Outside the system, candidates may have additional sources of information but that is an exception rather than the norm and hence, I do not consider that possibility here.

# 5 Data

To estimate candidates' location preferences, I use administrative data from four recruitment rounds, one each in 2012 and 2016, and two in 2013. The data has information on individual characteristics, the reservation category for which candidates are eligible, the category under which they are assigned (it is possible for a reserved category candidate to obtain assignment under open category based on merit alone), the police unit assigned to each candidate, the rank of this unit in the list of preferences submitted by the candidate, the total number of seats to be filled for each police unit and the disaggregated vacancy by reservation rules for each unit. Note that I do not observe the entire rank-ordered list submitted by each candidate.

**Individual Characteristics:** I observe individual characteristics such as home district, sex, and age for each recruitment round. For the 2016 round, I additionally observe education level, whether candidates' residence is classified as rural or urban, marital status, and the number of children, if any. The 2016 round also has the most number of women candidates, therefore, only 2016 data is used for understanding heterogeneity in preferences by individual characteristics.[10] The analysis is restricted to in-state candidates who submitted their preference ranking to the system.[11]

A description of women and men candidates in my estimation samples is in Table 3.

[10] 98% of the candidates submitted their preferences.

[11] Residence of out-of-state candidates is not known and they are dropped from the sample. An insignificant proportion of women candidates in each recruitment round are from out-of-state.

Candidates are between 18 and 33 years old with the median age being 22. 86-93% women are in civil police units, while 54% of men are in civil police units. Most of the remaining men are in special armed forces, which is a more rigorous posting than the civil units while most of the remaining women are in more flexible positions. In the 2016 round, 30% have Above HS education, 70-76% are from rural areas and 3-6% are married. These candidates were assigned to up to 85 police units.[12]

**Amenities:** Proximity to home has been found to be an important indicator of work location, especially for women (Roback (1982), Boyd et al. (2013), Black et al. (2014), Le Barbanchon et al. (2021), Liu and Su (2024), Barwick et al. (2021)). Therefore, I compute geodetic distances between the centroid of the candidates' home district and the headquarters of the district assigned to them as a measure of proximity to the workplace. Figure 3 shows the location of selected candidates by gender. The average woman is between 124-163 km from her hometown, while the average man is 205 km from his hometown (Table 3). I also obtain information on various infrastructure amenities, economic activity, and crime at the district level from different sources.

Literacy rate, sex ratio, population density, proportion of households with access to water within the household premises, and the proportion of households with TV are from the Census of India, 2011. The numbers of formal establishments and men and women workers working at such establishments are from the Sixth Economic Census 2013-14. Information on the number of schools affiliated with the Central Board of Secondary Education (CBSE) is from a portal maintained by the CBSE. Information on the number of government hospitals and the number of institutional deliveries as a proportion of total deliveries from the District Level Household and Facility Survey 2012-13. Finally, official crime reports are from the National Crime Records Bureau (NCRB) and the state's crime bureau.

Basic pay in this context is fixed across locations by Central Government Regulations. However, there is a cost of living adjustment in the form of house rent allowance determined by urban population levels. I use urban population at the district level from the 2011 Census to construct a variable measure of wages allowing for the house rent allowance adjustment. There is limited variation in this variable, which leads to limited statistical power in some models, but it is sufficient to monetize preferences for most amenities of interest.

[12]Data on amenities is not available for one of the districts as that district was formed post the 2011 Census when such data was compiled. Therefore, the candidates who hail from that district or who matched with the police unit in that district are not in my estimation samples.

# 6 Estimation and Results

The objective is to investigate the determinants of women police officers' location choices and measure the heterogeneity in preferences. I estimate preferences for two broad categories of amenities: those that vary by individuals and police unit locations and those that only vary across police unit locations but are common across individuals. I first estimate the conditional logit model in equation 1 using maximum likelihood estimation for candidates in the 2016 recruitment round separately by gender. Results are in Columns 1-3 of Table 4 for women candidates and Columns 4-6 for men. Data from prior recruitment rounds is later used to assess out-of-sample performance of these preference estimates.[13]

**Preferences for Proximity to Hometown:** Results indicate that all candidates have a distaste for distance from hometown and women have a stronger distaste for distance than men.[14] A concern with the model in Columns 1 and 4 of Table 4 is the presence of unobserved amenities that are correlated with distance. To allay concerns about endogeneity, I add police-zone intercepts in Columns 2 and 5.[15] Since zone intercepts may not completely account for unobserved location-specific preferences, I also estimate specifications with district-level intercepts in Columns 3 and 6. This is similar to the first step of the two-step approach of Berry et al. (1995) and Bayer et al. (2007), with the second-step in Table 4. Women's valuation of distance is Rs. 44/km, while men's valuation is 50-60% that amount. The average woman candidate would be willing to forego 12% of their basic pay to be in the district closest to their home, and the corresponding number for men is 7%. Results are not included in this draft, but similar valuation results using data from previous recruitment rounds.

The strong distaste for distance from one's hometown reflects the importance of proximity to social networks. Most of the candidate population is between 18 and 25 years of age, with little work or education experience, and most of their social circle is likely in their hometown.[16] Typically, this is their first job after graduating from school or college. About 75% are from rural areas. In this setting, being close to one's family becomes extremely important as a social safety net for both, the families and the candidates. In a society where women's mobility has traditionally been low, women officers or their families may be even less confident than men officers about an assignment in a district that is far from home.

[13] Information on candidate characteristics that can be used to account for preference heterogeneity and construct accurate feasible sets is limited in other recruitment rounds.

[14] The comparison between men and women candidates is based on a combined model. Results are not shown here for brevity.

[15] The 50 districts in the estimation sample comprise 11 police zones.

[16] In the current sample, 90% of the candidates are below 26, with the median age being 22 and 94% of them are unmarried.

Additionally, since government jobs are highly sought after in this setting and increase one's marriage prospects immensely, being close to home makes the process of 'arranged marriage' easy for the officers' families which is a common social norm in India.

There is evidence of heterogeneity in women's preferences for proximity to their hometown, which is robust across Models 1-3 in Table 4. Women with above-high-school education have a smaller distaste for distance. This could be due to targeting locations not just based on flexibility but the potential for promotion. Married women have a stronger distaste for distance likely reflecting familial constraints. Women candidates from rural areas also have a stronger distaste for distance. This could be due to limited information about the various police units or reflect the additional mobility constraints for rural residents. The results are somewhat different for men, likely due to differential gender norms and selection into the occupation for men and women.

**Preferences for Flexibility:** Civil police units are the most common types of police units and each district has one civil police unit. 6 out of 51 districts have other police units as well including special branches, which offer more flexible hours. Special branch positions are particularly sought after by those planning to prepare for other police or government exams. Officers who work in special branches can also hail from the same district, however, officers working in civil police units can't work in their home district.

The preference for flexibility is measured by the coefficient on the interaction between flexible posting type and a main city indicator. This is because only the four major cities– Bhopal, Gwalior, Indore, and Jabalpur– employ a considerable number of women officers in both flexible and inflexible types of police units. Results indicate that conditional on being in a major city, women have a strong preference for flexible positions, valued at 11% of the base pay. The same is not true for men. A distaste for extreme inconvenience including danger to life, the uncertain nature of duties, and hours is indicated in the strong negative coefficients on indicators for armed forces and railway posts.

**Preferences for Economic Infrastructure and Other Amenities:** Levels of economic and public infrastructure amenities are often highly positively correlated, making it difficult to credibly identify independent variation to look at substitution patterns between different amenities. For example, districts with greater economic progress and urbanization are also more likely to have a greater proportion of households owning a television or with access to tap water within the household premises, and likely better schools and health facilities as well. Identifying the importance of an individual amenity is also complicated by the possibility of differential levels of information about amenities. Amenity levels observed in data can be treated as proxies for real amenity levels, at best.

To account for preferences for district-level amenities, I construct a district-level amenity index which is the first principal component of the district-level amenities for which data is available. These include– the number of workers employed in the formal sector, an indicator for whether a district has a big hospital, the proportion of births delivered in a health institution, the proportion of households with access to tap water within the premises, the proportion of households with a television, literacy ratio, and the proportion of population classified as urban. Factor loadings indicating the importance of each of these variables in the index are in Appendix Table 7. Both women and men have a strong and robust preference for district amenities, however, women's preference is 64% stronger than men. The results are weaker with zone and district, likely due to limited statistical power.

**Preferences for Crime:** Another district-level amenity that police candidates may care about is crime. Ba et al. (2021) estimates Chicago police officers' preferences for community-level crime and shows that senior officers have a strong distaste for communities with high levels of violent crime, but not so for property crime. In my context, it is difficult to estimate police officers' preferences for crimes since I only have official data on district-level crime reports and not incidences. Certain crime categories are likely to have less underreporting than others. For example, the level of underreporting in property crimes, murders, and riots is much lower than for rapes or sexual assaults. This is consistent with the evidence found by a victimization survey conducted by an international NGO in a different part of the country (CHRI (2015)). Underreporting of crimes is also likely correlated with district infrastructure and resources.

Additionally, candidates may also be differentially aware of certain crimes from local sources of information or from news reports. For example, candidates may have better information about organized gangs, the general prevalence of crime and gang rapes than about reported dowry deaths, other types of rapes, and specific types of thefts.

Therefore, I only focus on one major category of crimes against women: assaults with the intent to outrage the modesty of women.[17] Assaults are likely easier to report and prosecute, and therefore, may be closer to the true situation than other recorded categories. It is unclear how people form perceptions about this measure but it is likely to be correlated to the reported levels of crimes specific to women. I use the number of assault per 1000

[17]Other important categories are rape and gang rapes. The reported incidence of gang rapes is quite low in the data and the rape reporting is very highly correlated with district-level amenities, suggesting a pattern of systematic underreporting in resource-constrained locations. This is reasonable and important since rape prosecutions require numerous steps to be followed by police officials at the initial stages and especially require women police officers. Lastly, crimes related to domestic violence may be irrelevant for women police officers concerned with their own safety in the workplace.

women in a district as a proxy for crimes that women may be especially concerned with.

Women have a preference for districts with lower reports of assaults against women, unlike men (Columns 1 and 2, Table 4). Results are weaker under the two-step BLP-approach due to limited statistical power. Women would be willing to accept a 6% reduction in base pay to go from a district that is at the 95$^{th}$ percentile to one at the 5$^{th}$ percentile.

**Robustness of Estimated Preferences:** To test the robustness of preference estimates obtained from the 2016 recruitment round, I assess the out-of-sample fit of a model similar to Column 3, Table 4.[18] I use parameter estimates implied by the 2016 recruitment round data along with individual data from 2012-13 recruitment rounds to obtain predicted shares and compare them with the true shares. Most deviations of predicted shares from the true share are economically small (Figure 5). 90% of the variation in true shares is explained by the out-of-sample predicted shares.[19] Similar to Pathak and Shi (2021), I conclude that parameter estimates from the discrete choice model perform well, suggesting their suitability for policy simulations in this and other contexts.

**Caveats:** The results on preferences for proximity to hometown are internally robust within this context. Even when the feasible alternatives are misclassified to be the same for all candidates, the coefficients on distance change only slightly. Some of the other preference parameters are affected slightly more than distance, but the overall conclusions about preferences remain unchanged. Due to this observation, I ignore two sources of information that would help identify feasible sets more accurately.

One, the current set of results does not account for vacancies while specifying feasible alternatives and hence, suffers from a slight loss of efficiency. Vacancy information by caste and gender is available only for 2016 and is missing for some of the police units. Therefore, accounting for it limits the comparison across rounds. In several cases, there are slots vacant for different categories. This means that the merit rank of the last assigned candidate does not give the correct cutoff. Ignoring this means that I have smaller feasible sets than the actual ones. This leads to a loss in efficiency and conservative estimates. The preference estimates are still consistent due to the stability property of DA.[20]

Two, I have information about the order of the assigned unit in candidates' ranking of preferences submitted to the system. For units for which someone is ineligible on the basis

[18] There are two differences between this specification and the one in Column 3, Table 4: first, age is the only taste-shifter for distance used in this estimation since 2012-13 recruitment rounds do not have information on other individual characteristics and second, indicator variable for special armed forces is dropped since this position was not available in the 2012-13 rounds and so, the variable cannot be defined.

[19] A limitation of this comparison is that the demand side is unknown for the 2012-2013 recruitment rounds, so it is possible for the predicted share to be positive even though the actual share is zero.

[20] Proof is a simple adaptation of that in Fack et al. (2019) and will be included in the next draft.

of their priority rank, I know the number of alternatives that would lie in the space where comparison with the chosen alternative is not possible. For example, consider a candidate who got assigned her second-most favorite unit. On the basis of her merit rank, she only qualifies for 10 units. So, in my analysis, her feasible set contains only 10 units even though I know that out of the remaining 40 units, there is only 1 unit that she prefers more to the one that she got assigned. So, in principle, there should be 39 additional units in her feasible set.

# 7 Discussion

## 7.1 The Appeal of the Deferred Acceptance Mechanism

Before the centralized police recruitment system in Madhya Pradesh, zonal authorities managed the recruitment and assignments in a decentralized manner for small groups of districts. The decentralized system limited the scope of the market and allowed for corruption at the hands of recruiters. The centralized recruitment and assignment mechanism helped resolve these issues and led to an increase in the number of applications and recruits. The choice of deferred acceptance mechanism (DA) was likely motivated by its use in the most prominent education institutions and government jobs in India: the assignment of India's civil servants recruited through the elite Union Public Service Commission exams, the assignment of engineering aspirants to majors, and numerous other college assignments. The DA mechanism rules promote competition by incentivizing aspirants' performance, which probably attracts quality candidates.

While the DA mechanism seems intuitive and fair, it is not similarly suitable for every context. First, in the context of police constable candidates, there isn't widespread agreement that constable quality as measured by their exam scores is relevant, especially since this is the most junior level of civil police. Uttar Pradesh Police also switched to a centralized assignment mechanism in 2011 to improve their candidate pool, but do not consider relative exam scores or candidate preferences in the unit assignment process. They make assignments based on distance from candidates' hometowns and the demand across different units. As long as the number of candidates is sufficiently large and police units do not have a preference for candidate's ranking in exams, the DA mechanism offers an improvement by improving candidate welfare at no extra cost to the police units. However, when the number of candidates is small relative to the demand for them, the DA mechanism can prove costly for locations that are less attractive as workplaces. In my context, there are numerous women applicants, but few selected candidates relative to the demand for them.

## 7.2 Alternative Mechanism: Vacancy Adjustment for Equity

Since women candidates have a strong preference for locations close to their home district and with better infrastructure, a DA mechanism would lead to a more equitable distribution only if all districts offered similar levels of amenities and if the number of selected candidates is large and uniformly distributed across the state.

An alternative mechanism that adjusts the number of unit-specific vacancies based on the number of qualified candidates can help address inequity across police units. Based on the number of qualified applicants, the initial reported vacancies input into the DA algorithm can be adjusted to ensure that the proportion of vacant positions is equal across police units after the assignment. This adjustment would ensure that the proportion of seats left vacant is uniform across all units and not biased against districts based on their initial amenity levels (as in Table 2). Since the rules of the mechanism remain unchanged, the desirable properties of the mechanism will be retained.

I assume that the extensive margin of women's employment decision will not be affected by this adjustment to the algorithm. Theoretically, this would depend on the outside option of women candidates and selection into the police force at the constable level suggests less attractive economic opportunities available in the local labor market. Additionally, vacancies are never advertised to the candidates so the decision to select into the police force cannot be a function of it. The vacancy adjusted mechanism might lead to more women candidates reneging. However, there is anecdotal evidence that in the current system early quits are being countered by withholding a part of the salary during the probationary period. Since women select into the system with this knowledge, an adjustment on the extensive margin is not likely. Additionally, newer recruitment rounds have adopted a more sophisticated preference elicitation interface allowing candidates to choose districts that are not acceptable to them.

I compute the monetary cost of welfare loss to women candidates under this vacancy-adjustment algorithm using preference estimates from Column 3 Table 4. I simulate outcomes under the counterfactual mechanism and compute the expected compensating variation that would make women candidates as well off as under the current mechanism.[21] For a conditional logit model with wage entering the utility function linearly, the exact formula for expected compensating variation is given in Small and Rosen (1981). Conditional on being assigned a different police unit, the average candidate experiences a loss valued at 15% of the basic pay of Rs. 26,000 based on the Sixth Pay Commission.

[21] In practice, the vacancy adjustment is not straightforward due to affirmative action rules described in Section 2.3, and the current set of results do not account for them. Simulation of outcomes accounting for this will be included in the next draft.

To examine who gets adversely affected in this counterfactual allocation, Table 6 compares the preference rank of the assigned unit in candidates' true and simulated rank-ordered lists. Under both cases, most individuals get one of their top five preferences. The distribution of the assigned preference order is less positively skewed under the adjusted rule, highlighting the welfare loss some candidates face. The total annual cost of compensating all affected candidates is equivalent to hiring 125 additional women officers. The total number of positions left vacant are about ten times this number, suggesting the relative importance of improving allocation vs increasing the number of selected candidates.

The vacancy-adjusted counterfactual mechanism is likely the least cost way of improving outcomes for the police units and improves allocation equity without compromising on stability. According to the rural hospital theorem, in the presence of strict preferences, there is no other stable mechanism that is welfare-improving for both candidates and the citizens served by the police units (Roth (1986)). Therefore, any alternative mechanism that improves citizen welfare on account of more police resources will also lower the welfare of police candidates in the absence of a variable wage schedule.

Quantifying the benefits of equitable distribution is beyond the scope of this paper, but an equitable distribution is likely to improve the future amenity values of locations through improvements in crime (Sukhtankar et al. (2022), Amaral et al. (2018), Miller and Segal (2018), and Sviatschi and Trako (2024)) and gender norms, and can lead to a more uniform spatial distribution of future applicants. There is suggestive evidence that the presence of women officers encourages more women applicants from the area in the future. The presence of women police constables and head constables in 2014 and the number of women employed in the formal sector explain about 50% of the variation in the location of 2016 women applicants. The presence of 10 women constables or head constables is crudely associated with an increase of 5-10 women applicants in the future from that district. This is likely to have positive spillovers on women's employment in other areas as well by improving location-specific gender norms, which have been shown to be critical for women's labor force participation (Jalota and Ho (2024), Heath et al. (2024)).

# 8 Conclusion

This paper exploits a novel policy setting to study work location preferences of women. I collect and use administrative data on police candidates in Madhya Pradesh, who are assigned to police units according to the Deferred Acceptance (DA) mechanism based on their rank in an entrance exam and preference lists submitted to the system. The theoretical properties of the mechanism allow me to convert the matching problem into a discrete

choice problem, in which candidates select the best alternative from their personalized set of feasible alternatives. This allows me to identify worker preferences for various job and location attributes.

All workers have a strong preference for proximity to home, however, women more so than men. Women's valuation of distance is Rs. 44/km relative to Rs. 27/km for men. All workers prefer areas with a better formal economy and better infrastructure as measured by health and education facilities and urbanization levels. Additionally, women also prefer more flexible job roles and areas with fewer reports of sexual assaults. These preferences are robust across time and are consistent with the larger literature on compensating wage differentials.

Due to the scarcity of women officers, these preferences create a spatial mismatch under the DA mechanism, leading to their disproportionate concentration in districts with better amenities while leaving other areas underserved. This imbalance may exacerbate existing inequalities in law enforcement resources and, consequently, gender disparities in workforce participation. To address this, I propose an alternative assignment mechanism that adjusts vacancy inputs into the algorithm to promote a more equitable distribution of women officers while preserving the desirable properties of the DA mechanism. The estimated welfare loss under this adjustment is relatively small—equivalent to filling just 10% of currently vacant positions—compared to the broader benefits of improved spatial equity. Prior research suggests that a greater presence of women officers not only facilitates crime reporting and reduces the incidence of crimes against women but also enhances women's workforce participation and attracts household in-migration, making this issue particularly relevant within India's social and economic landscape.

# References


Abdulkadiroğlu, Atila and Tayfun Sönmez (2003), "School choice: A mechanism design approach." *American economic review*, 93, 729–747.

Agarwal, Nikhil (2015), "An empirical model of the medical match." *American Economic Review*, 105, 1939–78.

Agarwal, Nikhil and Paulo J Somaini (2019), "Revealed preference analysis of school choice models." Technical report, National Bureau of Economic Research.

Amaral, Sofia, Prakash Nishith, and Sonia Bhalotra (2018), "Gender, crime and punishment: Evidence from women police stations in india." *Cornell University: NEUDC Retrieved from http://barrett. dyson. cornell. edu/NEUDC/paper_32. pdf*.

Ba, Bocar, Patrick Bayer, Nayoung Rim, Roman Rivera, and Modibo Sidibé (2021), "Police officer assignment and neighborhood crime." Technical report, National Bureau of Economic Research.

Barwick, Panle Jia, Shanjun Li, Andrew R Waxman, Jing Wu, and Tianli Xia (2021), "Efficiency and equity impacts of urban transportation policies with equilibrium sorting." Technical report, National Bureau of Economic Research.

Bayer, Patrick, Fernando Ferreira, and Robert McMillan (2007), "A unified framework for measuring preferences for schools and neighborhoods." *Journal of political economy*, 115, 588–638.

Berry, Steven, James Levinsohn, and Ariel Pakes (1995), "Automobile Prices in Market Equilibrium." *Econometrica: Journal of the Econometric Society*, 841–890. JSTOR.

Black, Dan, Natalia Kolesnikova, and Lowell Taylor (2009), "Earnings functions when wages and prices vary by location." *Journal of Labor Economics*, 27, 21–47.

Black, Dan A, Natalia Kolesnikova, and Lowell J Taylor (2014), "Why do so few women work in new york (and so many in minneapolis)? labor supply of married women across us cities." *Journal of Urban Economics*, 79, 59–71.

Boyd, Donald, Hamilton Lankford, Susanna Loeb, and James Wyckoff (2013), "Analyzing the determinants of the matching of public school teachers to jobs: Disentangling the preferences of teachers and employers." *Journal of Labor Economics*, 31, 83–117.

BPRD (2023), "Data on police organizations." Technical report, Bureau of Police Research and Development.

Card, David, Ana Rute Cardoso, Joerg Heining, and Patrick Kline (2018), "Firms and labor market inequality: Evidence and some theory." *Journal of Labor Economics*, 36, S13–S70.

CHRI (2015), "Crime victimisation and safety perception: A public survey of delhi and mumbai." Technical report, Commonwealth Human Rights Initiative.

Couture, Victor and Jessie Handbury (2022), "Spatial sorting within cities."

Dal Bó, Ernesto, Frederico Finan, and Martín A Rossi (2013), "Strengthening state capabilities: The role of financial incentives in the call to public service." *The Quarterly Journal of Economics*, 128, 1169–1218.

Diamond, Rebecca (2016), "The determinants and welfare implications of us workers' diverging location choices by skill: 1980-2000." *American Economic Review*, 106, 479–524.

Diamond, Rebecca and Cecile Gaubert (2022), "Spatial sorting and inequality." *Annual Review of Economics*, 14, 795–819.

Dubins, Lester E and David A Freedman (1981), "Machiavelli and the gale-shapley algorithm." *The American Mathematical Monthly*, 88, 485–494.

Fack, Gabrielle, Julien Grenet, and Yinghua He (2019), "Beyond truth-telling: Preference estimation with centralized school choice and college admissions." *American Economic Review*, 109, 1486–1529.

Falcettoni, Elena (2018), "The determinants of physicians' location choice: Understanding the rural shortage." *Available at SSRN 3493178*.

Finan, Frederico, Benjamin A Olken, and Rohini Pande (2015), "The personnel economics of the state."

Glaeser, Edward L, Jed Kolko, and Albert Saiz (2001), "Consumer city." *Journal of economic geography*, 1, 27–50.

Heath, Rachel, Arielle Bernhardt, Girija Borker, Anne Fitzpatrick, Anthony Keats, Madeline McKelway, Andreas Menzel, Teresa Molina, and Garima Sharma (2024), "Female labour force participation." *VoxDevLit*, 11, 1–43.

Jalota, Suhani and Lisa Ho (2024), "What works for her? how work-from-home jobs affect female labor force participation in urban india." *How Work-from-Home Jobs Affect Female Labor Force Participation in Urban India (January 4, 2024)*.

Jia, Ning, Raven Molloy, Christopher Smith, and Abigail Wozniak (2023), "The economics of internal migration: Advances and policy questions." *Journal of economic literature*, 61, 144–180.

Lavetti, Kurt and Ian M Schmutte (2016), "Estimating compensating wage differentials with endogenous job mobility."

Le Barbanchon, Thomas, Roland Rathelot, and Alexandra Roulet (2021), "Gender differences in job search: Trading off commute against wage." *The Quarterly Journal of Economics*, 136, 381–426.

Liu, Sitian and Yichen Su (2024), "The geography of jobs and the gender wage gap." *Review of Economics and Statistics*, 106, 872–881.

Miller, Amalia R and Carmit Segal (2018), "Do Female Officers Improve Law Enforcement Quality? Effects on Crime Reporting and Domestic Violence." URL `https://dx.doi.org/10.1093/restud/rdy051`.

Niederle, Muriel, Alvin E Roth, and Tayfun Sonmez (2008), "Matching and market design." *The New Palgrave Dictionary of Economics*, 5, 436–445.

Oates, Wallace E. (1999), "An essay on fiscal federalism." *Journal of Economic Literature*, 37, 1120–1149.

Pathak, Parag A (2017), "What really matters in designing school choice mechanisms." *Advances in Economics and Econometrics*, 1, 176–214.

Pathak, Parag A and Peng Shi (2021), "How well do structural demand models work? counterfactual predictions in school choice." *Journal of Econometrics*, 222, 161–195.

Perova, Elizaveta and Sarah Anne Reynolds (2017), "Women's police stations and intimate partner violence: Evidence from brazil." *Social Science & Medicine*, 174, 188–196.

Roback, Jennifer (1982), "Wages, rents, and the quality of life." *Journal of political Economy*, 90, 1257–1278.

Roth, Alvin E (1982), "The economics of matching: Stability and incentives." *Mathematics of operations research*, 7, 617–628.

Roth, Alvin E (1984), "Stability and polarization of interests in job matching." *Econometrica: Journal of the Econometric Society*, 47–57.

Roth, Alvin E (1985), "The college admissions problem is not equivalent to the marriage problem." *Journal of economic Theory*, 36, 277–288.

Roth, Alvin E (1986), "On the allocation of residents to rural hospitals: a general property of two-sided matching markets." *Econometrica: Journal of the Econometric Society*, 425–427.

Roth, Alvin E (2003), "The origins, history, and design of the resident match." *Jama*, 289, 909–912.

Roth, Alvin E, Tayfun Sönmez, and M Utku Ünver (2004), "Kidney exchange." *The Quarterly journal of economics*, 119, 457–488.

Roth, Alvin E and Marilda Sotomayor (1992), "Two-sided matching." *Handbook of game theory with economic applications*, 1, 485–541.

Saks, Raven E and Abigail Wozniak (2011), "Labor reallocation over the business cycle: New evidence from internal migration." *Journal of Labor Economics*, 29, 697–739.

Singh, Divya (2019), "Safer elections, women turnout, and political outcomes." *Women Turnout, and Political Outcomes (December 17, 2019)*.

Siwach, Garima (2018), "Crimes against women in india: Evaluating the role of a gender representative police force." *Available at SSRN 3165531*.

Small, Kenneth A and Harvey S Rosen (1981), "Applied Welfare Economics with Discrete Choice Models." *Econometrica: Journal of the Econometric Society*, 105–130. JSTOR.

Sönmez, Tayfun and Tobias B Switzer (2013), "Matching with (branch-of-choice) contracts at the united states military academy." *Econometrica*, 81, 451–488.

Sonmez, Tayfun and M Bumin Yenmez (2019), "Affirmative action in india via vertical and horizontal reservations."

Sorkin, Isaac (2018), "Ranking firms using revealed preference." *The quarterly journal of economics*, 133, 1331–1393.

Sukhtankar, Sandip, Gabrielle Kruks-Wisner, and Akshay Mangla (2022), "Policing in patriarchy: An experimental evaluation of reforms to improve police responsiveness to women in india." *Science*, 377, 191–198.

Sviatschi, Maria Micaela and Iva Trako (2024), "Gender violence, enforcement, and human capital: Evidence from women's justice centers in peru." *Journal of Development Economics*, 168, 103262.

Thakur, Ashutosh (2021), "Matching in the civil service: A market design approach to public administration and development." Technical report, ECONtribute Discussion Paper.

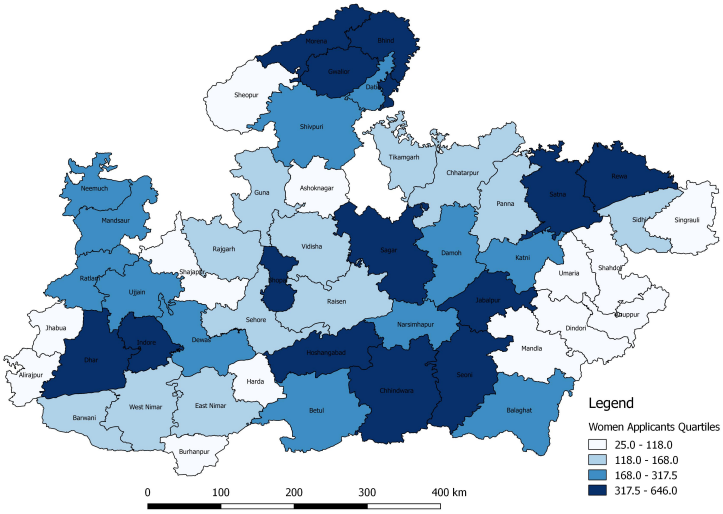
# 9 Figures and Tables

Figure 1: Reservation rules

| Vertical reservation categories | Horizontal reservation categories | Who is eligible? |
|---|---|---|
| **Unreserved category (UR) or General Category** | Open | Anybody regardless of gender and caste category- based on merit |
| | UR Female | Only UR Female |
| | UR Ex-servicemen | Only UR Ex-servicement |
| **OBC Category** | Open | All OBC candidates regardless of gender |
| | OBC Female | Only OBC Female |
| | OBC Ex-servicemen | Only OBC Ex-servicemen |
| **SC Category** | Open | All SC candidates regardless of gender |
| | SC Female | Only SC Female |
| | SC Ex-servicemen | Only SC Ex-servicemen |
| **ST Category** | Open | All ST candidates regardless of gender |
| | ST Female | Only ST Female |
| | ST Ex-servicemen | Only ST Ex-servicemen |

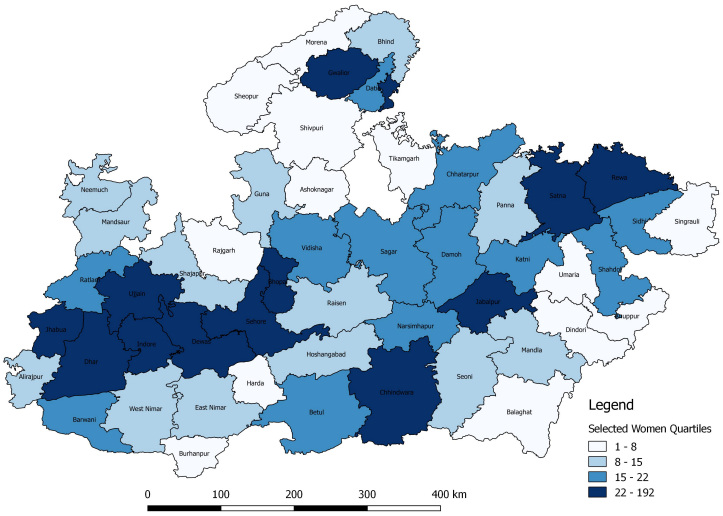
This Table summarizes the different categories defined by the affirmative action policy. These are discussed in Section 2.3.

Figure 2: Spatial Distribution of Women Applicants



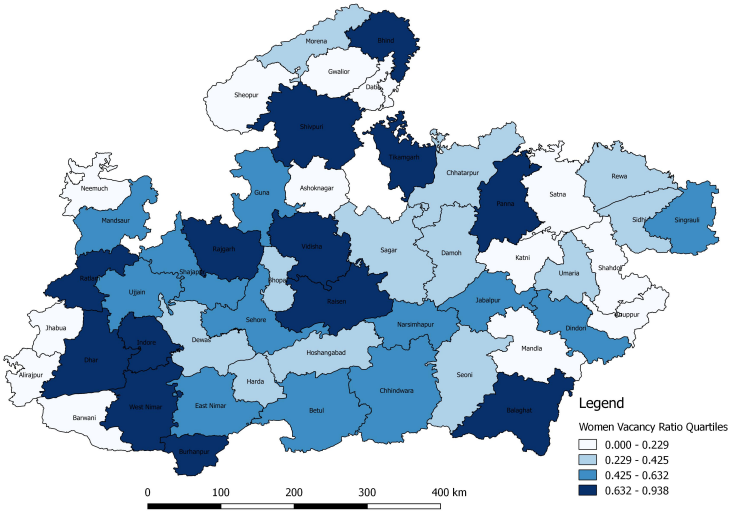

This map shows the districts of Madhya Pradesh with four quartile classes based on the number of women applicants in the 2016 recruitment round from each district. Out-of-state applicants are not included here.

Figure 3: Spatial Distribution of Selected Women Candidates



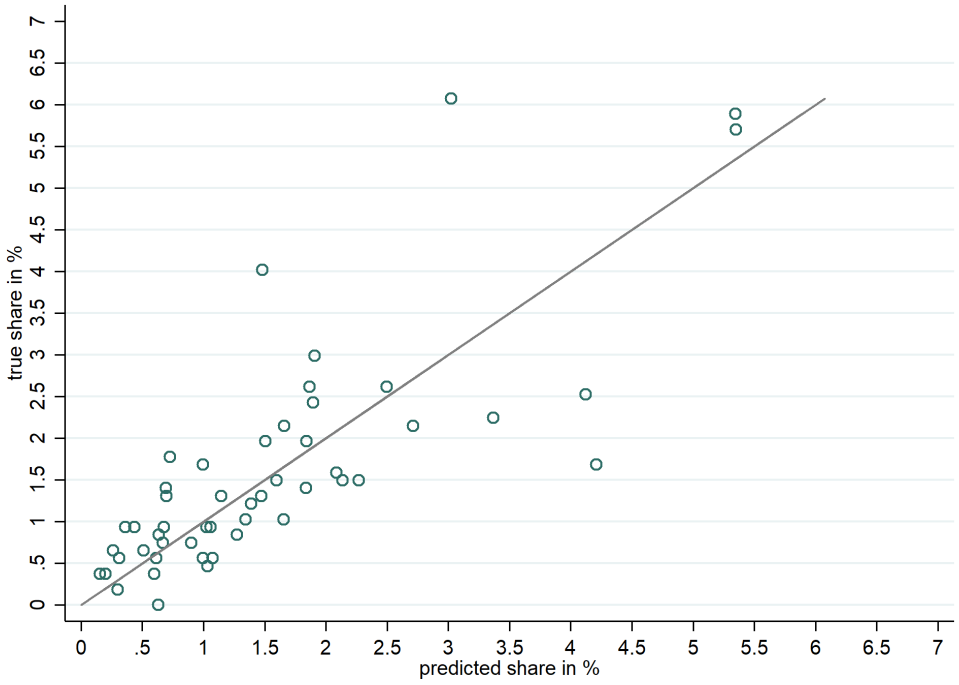

This map shows the districts of Madhya Pradesh with four quartile classes based on the number of selected women candidates in the 2016 recruitment round from each district.

Figure 4: Spatial Distribution of the Proportion of Vacant Women Positions




This map shows the districts of Madhya Pradesh with four quartile classes based on the proportion of seats that remain vacant after the assignment. This indicates the extent of unfulfilled demand of each district.

Figure 5: Out-of-sample Predictions




This figure compares the out-of-sample predicted shares at the district level with true shares from a model similar to Column 3 Table 4, but with age as the only individual level taste-shifter since the prediction is for the sample of women recruited in 2012-2013 recruitment rounds. Special armed forces indicator was excluded from the estimation due to no position of such type in the 2012-13 recruitment rounds. Bhopal has been omitted from this graph from readability (predicted share=20.3% and true share=23.4%). R-squared= 0.9

Table 1: Percentage of Candidates Allotted Their Top Ranked Preferences

| Recruitment Rounds | Top 1 (in %) | | Top 3 (in %) | | Top 5 (in %) | | No. of Candidates | |
|---|---|---|---|---|---|---|---|---|
| | Women | Men | Women | Men | Women | Men | Women | Men |
| 2012 | 56.52 | 36.7 | 94.2 | 59.82 | 97.83 | 70.49 | 138 | 3962 |
| 2013-I | 95.26 | 24 | 96.39 | 40.29 | 96.39 | 46.67 | 443 | 6458 |
| 2013-II | 68.95 | 34.2 | 78.9 | 55.44 | 78.9 | 65.1 | 583 | 4713 |
| 2016 | 50.55 | 17.03 | 75.52 | 30.82 | 87.01 | 41.93 | 1001 | 8484 |

This Table shows the percentage of candidates who got assigned one of their top K choices. This is based on observed data on the order of the assigned police unit in the rank-ordered list submitted by the candidates. Rank-ordered lists are not observed.

Table 2: Determinants of Vacancies as a Proportion of Initial Demand of Police Units for Women Constables

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| Civil Police Unit Indicator | -0.1328 | -0.4516** | -0.4974*** | -0.6022*** | -0.8459*** |
| | (0.156) | (0.184) | (0.180) | (0.216) | (0.222) |
| Initial Vacancy | 0.0052*** | 0.0078*** | 0.0080*** | 0.0091*** | 0.0103*** |
| | (0.001) | (0.001) | (0.001) | (0.002) | (0.002) |
| Amenity Index | | -0.0525** | -0.0592*** | -0.0708** | -0.0472 |
| | | (0.024) | (0.021) | (0.029) | (0.030) |
| Assaults per 1000 women | | | 0.7356 | | |
| | | | (0.565) | | |
| Women Applicants within 100km | | | | | -0.0021*** |
| | | | | | (0.001) |
| Zone Fixed Effects | No | No | No | Yes | Yes |
| Observations | 55 | 55 | 54 | 55 | 55 |
| Adjusted $R^2$ | 0.24 | 0.30 | 0.32 | 0.23 | 0.35 |

Dependent variable is the ratio of vacancy to original postings for women at the level of police units. Type is an indicator variable for civil police units. Out of the 11 other units, initial vacancy was known only for 4 of them. Amenity index represents the first principal component of district level amenities. Factor loadings are in Table 7. Robust std errors are in parentheses.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3: Summary Statistics of the Estimation Samples

| Variable | Obs | Mean or Proportion | Std. Dev. | Min | Max |
|---|---|---|---|---|---|
| **Recruitment Round 2016: Women** | | | | | |
| Age | 994 | 21.74 | 2.64 | 18 | 32 |
| Above High-school Education | 994 | 0.30 | 0.46 | 0 | 1 |
| Distance in km. | 994 | 123.60 | 102.74 | 5 | 625.37 |
| Rural | 994 | 0.69 | 0.46 | 0 | 1 |
| Married | 994 | 0.03 | 0.18 | 0 | 1 |
| Flexible Posting | 994 | 0.03 | 0.17 | 0 | 1 |
| **Recruitment Round 2016: Men** | | | | | |
| Age | 6,303 | 22.72 | 3.13 | 18 | 42 |
| Above High-school Education | 6,303 | 0.29 | 0.45 | 0 | 1 |
| Distance in km. | 6,303 | 204.80 | 163.14 | 0 | 802.84 |
| Rural | 6,303 | 0.76 | 0.42 | 0 | 1 |
| Married | 6,303 | 0.06 | 0.24 | 0 | 1 |
| Flexible Posting | 6,303 | 0.03 | 0.17 | 0 | 1 |
| **Recruitment Round 2012: Women** | | | | | |
| Age | 134 | 22.13 | 2.86 | 18 | 31 |
| Distance in km. | 134 | 145.39 | 107.70 | 36.56 | 476.91 |
| Flexible Posting | 134 | 0.07 | 0.26 | 0 | 1 |
| **Recruitment Round 2013 I: Women** | | | | | |
| Age | 404 | 23.00 | 2.76 | 18 | 33 |
| Distance in km. | 404 | 126.89 | 99.90 | 36.56 | 625.37 |
| Flexible Posting | 404 | 0.14 | 0.35 | 0 | 1 |
| **Recruitment Round 2013 II: Women** | | | | | |
| Age | 532 | 22.28 | 2.65 | 18 | 32 |
| Distance in km. | 532 | 162.53 | 136.62 | 36.56 | 788.91 |
| Flexible Posting | 532 | 0.17 | 0.38 | 0 | 1 |

This Table presents the summary statistics for the estimation samples used for the main model and the robustness checks.

Table 4: Conditional Logit Location Choice Model: Preferences for Distance and Amenities

| | (1) Women | (2) Women | (3) Women | (4) Men | (5) Men | (6) Men |
|---|---|---|---|---|---|---|
| Distance in km. | -0.01643*** | -0.01872*** | -0.01992*** | -0.00940*** | -0.01120*** | -0.01119*** |
| | (0.001) | (0.001) | (0.001) | (0.000) | (0.000) | (0.000) |
| Above HS × Distance | 0.00281** | 0.00346*** | 0.00383*** | -0.00152*** | -0.00088** | -0.00084** |
| | (0.001) | (0.001) | (0.001) | (0.000) | (0.000) | (0.000) |
| Married × distance | 0.00534** | 0.00445* | 0.00422* | -0.00173** | -0.00141* | -0.00143* |
| | (0.002) | (0.002) | (0.002) | (0.001) | (0.001) | (0.001) |
| Rural × distance | -0.00237* | -0.00220* | -0.00202* | 0.00021 | 0.00124*** | 0.00125*** |
| | (0.001) | (0.001) | (0.001) | (0.000) | (0.000) | (0.000) |
| Assaults per 1000 women | -3.68271*** | -3.79331*** | | -0.12071 | 0.68968 | |
| | (0.696) | (0.912) | | (0.266) | (0.373) | |
| Index of amenities | 0.29667*** | 0.20312*** | | 0.33619*** | 0.10774*** | |
| | (0.042) | (0.046) | | (0.017) | (0.020) | |
| Wage | 0.00034* | 0.00048** | | 0.00014** | 0.00042*** | |
| | (0.000) | (0.000) | | (0.000) | (0.000) | |
| Main City Dummy | -1.15564*** | -0.99716*** | -0.13704 | -1.52640*** | -0.35188** | 0.16489 |
| | (0.253) | (0.297) | (0.427) | (0.103) | (0.126) | (0.182) |
| Flexible Post | -1.85981*** | -1.65474*** | -1.57338*** | -1.74772*** | -1.48345*** | -1.49980*** |
| | (0.343) | (0.349) | (0.365) | (0.221) | (0.225) | (0.234) |
| Main City × Flexible Post | 1.77039*** | 1.47711*** | 1.30813** | 0.68442** | 0.45528 | 0.38324 |
| | (0.371) | (0.380) | (0.398) | (0.237) | (0.241) | (0.249) |
| Armed Forces Post | -2.61098*** | -2.72006*** | -2.93545*** | -5.13281*** | -5.18463*** | -5.27647*** |
| | (0.403) | (0.404) | (0.411) | (0.096) | (0.108) | (0.109) |
| Railway Post | -1.35650*** | -1.35241*** | -1.39148*** | -1.04596*** | -1.05670*** | -1.11855*** |
| | (0.206) | (0.207) | (0.210) | (0.087) | (0.087) | (0.088) |
| Intercepts | None | Zone | District | None | Zone | District |
| Individuals | 994 | 994 | 994 | 6303 | 6303 | 6303 |
| LR chi2 | 1330.55 | 1299.05 | 1368.08 | 6784.17 | 6564.17 | 6850.35 |
| Log likelihood | -2305.54 | -2244.64 | -2176.34 | -15985.71 | -15523.40 | -15329.84 |
| Average MRS (Rs./km) | 49 | 40 | 44 | 67 | 27 | 19 |

The Table presents estimated preferences for the sample of women and men candidates in the 2016 Recruitment Round. There are four types of postings: Civil Police Units (omitted category), Armed Forces, Railways, and Special Branches (Flexible Post dummy). Main city indicator represents Bhopal, Indore, Gwalior, and Jabalpur. Amenity index represents the first principal component of district level amenities. Factor loadings are in Table 7. Std. errors are in parentheses.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 5: Second stage of BLP for Columns 3 and 6 from Table 4

| | (1) Women | (2) Men |
|---|---|---|
| Assaults per 1000 women | -0.39865 | 0.25874 |
| | (2.783) | (1.076) |
| Index of amenities | 0.09493 | 0.00879 |
| | (0.068) | (0.020) |
| Wage | 0.00045* | 0.00059** |
| | (0.0002) | (0.0002) |
| Zone Intercepts (Zone 1 Omitted): | | |
| Zone 2 | 1.72938*** | 0.48819*** |
| | (0.188) | (0.099) |
| Zone 3 | -0.39200*** | -0.92837*** |
| | (0.080) | (0.057) |
| Zone 4 | 0.53357*** | -0.40579*** |
| | (0.154) | (0.087) |
| Zone 5 | 1.28610*** | 0.20808* |
| | (0.217) | (0.098) |
| Zone 6 | 2.15683*** | 1.13246*** |
| | (0.048) | (0.037) |
| Zone 7 | 0.98053*** | 0.09471 |
| | (0.155) | (0.080) |
| Zone 8 | -0.06624 | -0.16353*** |
| | (0.081) | (0.048) |
| Zone 9 | 0.28619 | -0.21394 |
| | (0.315) | (0.118) |
| Zone 10 | -0.10769 | -0.57978*** |
| | (0.219) | (0.092) |
| Zone 11 | 2.51648*** | 1.01650*** |
| | (0.096) | (0.060) |
| Adjusted $R^2$ | 0.786 | 0.711 |
| Districts | 51 | 51 |

Dependent variable is the vector of estimated intercepts from the first stage conditional logit model. Robust std errors clustered at the zone level are in parentheses.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 6: Percentage of Candidates Allotted Their Top Ranked Preferences Based on Predicted Rank Order from Column 3 Table 4

| | Original Vacancy | Adjusted Vacancy |
|---|---|---|
| Top 1 | 76.40 | 63.13 |
| Top 3 | 88.76 | 79.98 |
| Top 5 | 92.85 | 87.95 |
| Top 10 | 98.16 | 95.51 |
| Mean Preference Order | 1.85 | 2.57 |

This Table shows the percentage of candidates who will be assigned one of their top K choices by the DA mechanism based on simulated preferences for women candidates in 2016. Original vacancy for women is observed in for all but four units. Adjusted vacancy is based on a flat vacancy rate of 55% as it equalized the number of selected women candidates with positions filled. Note that this doesn't account for the affirmative action policy. Predicted rank order is based on a specification similar to Column 3 of Table 4

Table 7: Factor Loading for Variables in District Amenities Index

| Variable | Factor Loading |
|---|---|
| log(workers) | 0.3265 |
| Big Hospital Dummy | 0.3555 |
| Number of CBSE Schools | 0.3794 |
| Proportion of Institutional Deliveries | 0.2693 |
| Proportion of HHs with Tap Water Access | 0.361 |
| Proportion of HHs with a TV | 0.4062 |
| Literacy Rate | 0.301 |
| Proportion of Urban Population | 0.4058 |

This Table presents factor loadings for the district amenities index variable used in the police unit choice model.