Estimating Matching Games Without Individual-level Data: Multidimensional Sorting in Government Recruitment*

Qiwei He†

August, 2024 (Latest Version)

Abstract

This paper investigates the non-parametric identification of matching games without individual-level data. I build a many-to-one Non-transferable Utility job matching model with two-sided heterogeneity and demonstrate the non-parametric identification of the model using only position-level data based on instruments. I apply the framework to study matching mechanism design in the context of the National Civil Service Exam (NCSE), a primary method for recruiting entry-level government officials in China. The NCSE is a mechanism that each candidate applies to *one* position and subsequently takes a meritocratic exam determining admission. In the counterfactual analysis, I introduce a serial-dictatorship mechanism as an alternative. A strategy-proof mechanism may not be ideal when candidates differ in multiple dimensions while there is information friction in the hiring process.

Keywords: non-parametric identification, government recruitment, multi-dimensional sorting, market design

JEL Clacification: C14, C51, D47, D73, J49, H11

*I am indebted to Michael Waldman, Philipp Kircher, and Francesca Molinari for their continued advice and support. I received helpful comments and suggestions from Nikhil Agarwal, Panle Jia Barwick, Justin Bloesch, Jiahua Che, Steve Coate, Mariacristina De Nardi, Martin Hackmann, Yinghua He, Ravi Kanbur, Elena Manresa, Jesse Shapiro, Paulo Somaini, Joerg Stoye, Ashutosh Dinesh Thakur and Guo Xu. Seminar participants at I appreciate Bo Feng's help with the dataset.

†He, University of Edinburgh, qiwei.he@ed.ac.uk

1 Introduction

Recruiting bureaucrats is a necessary task for all governments that has important implications on governance, for example, the provision of public goods and services.[1] Though the literature has intensely studied self-selection into the government or public sector (e.g., Besley (2005); Besley et al. (2022)), how candidates are assigned to specific jobs is less understood. While different countries use different ways to solve the assignment problem, many governments worldwide, including those of China and India, adopt a centralized matching market framework to streamline their recruitment (Gale and Shapley, 1962; Becker, 1973). Typically, the framework involves a screening phase, such as a written exam or interview. Then the government employs a mechanism to match candidates to job openings, considering both screening outcomes and individual preferences.

Other than a candidate's intrinsic ability, there is an important second dimension that crucially affects match quality in the context of government recruitment, namely, the civic-mindedness of candidates.[2] People apply for government jobs with different intrinsic motivations. Some are committed to serve the public, while others focus more on collecting personal benefit. Moreover, even though civic-mindedness plays a pivotal role in the quality of public good and service provision, especially in societies with relatively weak institutions, it is challenging to measure it through a meritocratic exam. A natural question emerges: Given a pool of candidates and available positions, can the government match candidates with high ability and high civic-mindedness to suitable jobs when candidate civic-mindedness is unobservable in the recruitment process? In this paper, I study how matching mechanisms shape matching patterns

[1] Best, Hjort and Szakonyi (2023), Dahis, Schiavon and Scot (2023), Ashraf et al. (2020), Spenkuch, Teso and Xu (2023), Moreira and Pérez (forthcoming) and Fenizia (2022) along with others provide empirical evidence that individual government workers significantly impact government performance.

[2] Please refer to Besley and Ghatak (2018) for a discussion of pro-social preferences and incentives in bureaucracy. Dal Bó, Finan and Rossi (2013) and Ashraf et al. (2020) are examples that empirically study prosocial preferences and political selection.

in government recruitment, as well as how the resulting matching outcomes impact government performance.

Due to data privacy or security concerns, individual-level data can be unavailable in studying government recruitment. In contrast, data at a more aggregate level is typically more accessible. This data limitation presents a significant econometric challenge: can NTU matching models with two-sided heterogeneity be estimated using aggregate-level data rather than individual-level data? Since this empirical challenge exists in more general settings than government recruitment, I start by building a general NTU matching model in the labor market with two-sided heterogeneity.

I consider large matching markets where, in each market, there is a continuum of candidates and a set of firms each posting a position. Upon posting, all firms commit to market-specific position characteristics, including the wages and the hiring quota for each position. On the other side, candidates differ in their qualifications that firms care about, for example, experience or skills. After the vacancy posting, candidates make application choices. I first consider the scenario where each candidate is allowed to apply for one position motivated by the empirical application. Then I extend the main result to the scenario where each candidate can make multiple applications. After the application stage, each firm conducts a meritocratic selection based on its preference on candidate qualifications to determine which candidate is matched. I define an equilibrium for the matching market where a set of equilibrium hiring standards adjust to fill all positions.[3]

The researcher is assumed to observe only a set of firm-level characteristics and *aggregated* equilibrium allocations including the number of applicants and equilibrium hiring standard for each firm. The main challenge in identifying firm preferences and candidates' preferences/types is that the observed equilibrium allocations at the firm (or position) level are simultaneously determined by both firm and candidate preferences.

[3]Here, equilibrium hiring standards can be multi-dimensional. For a related setting where stable matching is characterized by cutoffs in large markets, see Azevedo and Leshno (2016).

The key insight is that when applying for a position, candidates care about the probability of being matched and the expected utility conditional on being matched with the firm. In equilibrium, the probability for a candidate to be matched with a firm must contain the firm's preference, represented by the candidate's measured qualification for the position relative to the equilibrium hiring standard.[4] Therefore, a candidate's application decision contains information from both sides of the market, so do the aggregated application decisions at the position level. Therefore, conditional on observing the total number of applicants, equilibrium hiring standard, and position characteristics for each position, both firm and candidate preferences/types can be non-parametrically identified given instruments that shift firm preferences and candidate preferences/types independently. Here equilibrium hiring standards provide useful information on the competitiveness of each job and the distribution of candidate qualifications.

The non-parametric identification result is established based on instruments, building on the framework developed by Newey and Powell (2003) and Berry and Haile (2014). Berry and Haile (2014) demonstrate the identification of non-parametric models of differentiated products markets given "market-level" data. There are several conceptual differences between the matching model considered here and their setting. First, in the model, two-sided preferences (both firm and candidate) matter for the allocation. Second, the number of available positions is fixed in the matching market, while the product market usually features an elastic supply. Third, based on the first and second points, the potential multiplicity of equilibrium hiring standards poses a threat to non-parametric point identification.[5] I characterize conditions under which a unique set of equilibrium hiring standards clears each matching market. Finally, accommodating the possibility of multiple applications is crucial in the context of matching market.

[4]Here, candidates must be able to correctly know firm preferences and equilibrium hiring standards on the equilibrium path.

[5]Please refer to Molinari (2020) for related discussions in the literature.

The empirical application focuses on the National Civil Service Examination (henceforth, NCSE) in China, which is one of the principal mechanisms through which the Chinese central government fills entry-level positions. It is very popular, with over 2.5 million applicants competing for approximately 37,000 jobs in the 2023 NCSE. Under the NCSE, each candidate can only apply for *one* position before the meritocratic selection that determines the matching outcome, which leads to strategic application. I compile a dataset of the NCSE for 2016 and 2017, which incorporates rich position-level information, including position characteristics, the number of applicants for each position as well as the cutoff score required to pass the exam for each position.

I setup a parametric empirical model to study the NCSE. Candidates differ in their ability and civic-mindedness. Civic-mindedness is an important second dimension that determines the candidate willingness to engage in rent-seeking activities. The exam is modeled as a noisy signal of the applicant's ability. However, candidate civic-mindedness is assumed to be not measurable through the exam process. Government preference affects the assignment of exam papers with different difficulties to positions. While candidates have preferences on a variety of position attributes, lower civic-minded candidates are assumed to derive higher utility from potentially high rent positions, as indicated by the number of past anti-corruption cases.[6]

I estimate the empirical matching model using an algorithm that builds on (Berry, Levinsohn and Pakes, 1995), with an additional market-clearing constraint in the estimation process. Variations in observed cutoff scores provide useful information in identifying the candidate ability distribution. Variations of position characteristics within and across application markets provide information regarding candidate preferences. I employ two sets of instruments. First, a set of instruments motivated by the institutional background is used to identify government preferences reflected by

[6]The interpretation that a higher number of past anti-corruption cases implies greater rent-seeking opportunity aligns with the empirical study by Wang and Dickson (2022), Sun, Zhu and Ni (2022) and Pan, Shu and Ye (2023).

the assignment of exam papers.[7] Second, I utilize functions of position characteristics that are commonly used in the empirical literature of discrete choice to identify candidate types and preferences.

In the counterfactual analysis, I consider a serial dictatorship (SD) mechanism that is similar to the mechanism currently used in the India civil service recruitment. Under SD, every candidate takes the same meritocratic exam. The candidate ranked rth based on exam scores makes a choice in the rth order until all positions are filled. Simulation results indicate that, in contrast to the NCSE, SD is more likely to match high-ability candidates to popular positions, but is also more likely to match low civic-minded candidates to potentially high rent positions. These patterns emerge since candidates can choose from all remaining positions when their turn comes to maximize their post-match utility. Since empirically evaluating government performance poses a significant challenge, I compare the NCSE and SD through a simulation. Simulation suggests that when the cost of corruption in government is high, then NCSE is a better option even though it introduces strategic application, because it is less likely to fill high rent positions with low civic-minded candidates. In the context of government recruitment, this paper makes a contribution by studying how do matching mechanisms shape matching patterns in government recruitment when candidates have multidimensional characteristics.

Related Literature. This paper contributes to the identification of many-to-one Non-transferable Utility (NTU) matching models. By considering limited preference heterogeneity, Agarwal (2015) and Diamond and Agarwal (2017) demonstrate the identification and estimation of preferences in a large matching market. He, Sinha and Sun (forthcoming) show non-parametric identification of many-to-one two-sided NTU matching markets with rich preference heterogeneity from both sides. The

[7] I use whether a position is located in a sub-provincial level (or above) city as an instrument for the use of the provincial exam paper for that position because cities at the sub-provincial level (or above) are more likely to have positions affiliated with a department at or above the provincial level. The instrument based on institutional background to identify candidate preferences is the local budgetary fiscal expenditure.

main distinction between this paper and others in the literature is that, to my knowledge, it represents the first attempt to non-parametrically identify many-to-one NTU matching models when individual-level data is not available to the researcher.[8]

The paper contributes to the understanding of political selection, in particular using exams to recruit bureaucrats.(Besley, 2005; Besley et al., 2022; Best, Hjort and Szakonyi, 2023; Dahis, Schiavon and Scot, 2023; Moreira and Pérez, forthcoming) Thakur (2021) also studies the civil service recruitment from a matching market perspective focusing on India. My paper differs from Thakur (2021) in several dimensions. First, I develop a matching game model and show non-parametric identification identified without individual data to overcome the empirical challenge. Second, the "apply first, then exam" system of the NCSE presents a unique institutional feature.[9] I analyze the equilibrium sorting pattern of the NCSE, and compare its empirical matching outcomes with those produced by a strategy-proof alternative mechanism. Third, I emphasize the importance of understanding how multidimensional talent characteristics affect government performance and examining the sorting of these traits under different matching mechanisms. Civic-mindedness considerred here echoes discussions of pro-social preferences in bureaucracy (e.g., Francois (2000); Besley and Ghatak (2003, 2018)).

The remainder of the paper is organized as follow. Section 2 presents the job match model and provides the main identification result. Section 3 applies the theoretical framework to analyze matching mechanism design in the context of National Civil Service Exam (NCSE) in China. Section 4 concludes.

[8] A growing strand of literature studies the econometrics of decision problems under consideration set formation (Cattaneo et al., 2020; Abaluck and Adams-Prassl, 2021; Barseghyan, Molinari and Thirkettle, 2021; Barseghyan et al., 2021; Agarwal and Somaini, 2022). There is a conceptual differene between the probability that different consideration sets are drawn in a job seeker's decision problem and the perceived chance of being hired in this paper. The probability of being matched with a job consists of both job preferences and its equilibrium hiring standard.

[9] A burgeoning literature studies the "civil service fever" in China. (Feng et al., 2023; Feng, He and Liu, 2023)

2 The Matching Game Model and Its Identification

This section presents an NTU job matching game model with two-sided heterogeneity for labor markets. Then I study its non-parametrication with only aggregated-level data.

The model is formulated as a job search problem but retains applicability to other matching markets of similar structure, such as school admissions. Consider a collection of markets, each indexed by t. Markets are generally defined either geographically or temporally, or both. There is a continuum of agents (i) of mass 1 within each market, i.e., a large market approximation. On the other side of the market, there is a set of firms ($j \in \{0, 1, ..., J\}$). In every market t, each firm j posts *one* vacancy or position. Each position posted by firm j in market t can hire up to λ_{jt} share of agents. I define the set of all these capacities in market t as $\lambda_t = \{\lambda_{0t}, ..., \lambda_{Jt}\}$. Let $x_t = (x_{1t}, ..x_{Jt})$, where x_{jt} represents a set of characteristics associated with the position posted by firm j in market t. These characteristics x_t can include factors such as compensation, office environment, and office location. I assume that firms commit to those characteristics (λ_t and x_t) upon posting. Furthermore, I assume that both λ_t and x_t are common knowledgte. An agent can be matched with at most one firm in the market she resides. I assume that the total number of jobs is much less than the total size of agents. Therefore, there will be some unmatched agents, but no unmatched firms.

Firm Preferences. Candidates have some qualifications (q_i) that firms have preferences on. q_i can be multidimensional, such as education level and experience. q_i is assumed to be i.i.d. drawn from a distribution F_q. While q_i is private information for each agent i, F_q is known by all firms and agents.

I model firm preferences on candidates as a selection function that determines each candidate's probability of being matched with a firm, which is denoted as $g(q_i, \bar{q}_{jt}, x^f_{jt})$. Higher q_i brings a higher probability of being matched since firms prefer candidates with higher qualifications. The selection function g can be motivated as a meritocratic

selection for candidates who applied for firm, where candidates achieving a higher performance in the selection are hired. The function g is assumed to be probabilistic, highlighting the uncertainty in the hiring process.[10]

I characterize the matching market using equilibrium hiring standards ($\bar{q}_{i,t}$) of firm j in market t. For instance, if firm j is only concerned about an agent's exam score (e.g., SAT or GRE), then $\bar{q}_{jt}$ represents the score cutoff at which firm j is indifferent between hiring or not hiring an agent in market t, given the number of candidates it demanded. In the model, however, $\bar{q}_{i,j}$ can be multidimensional and is endogenously determined by the application choices, labor demand and firm preferences in each market t. Intuitively, the dimension of $\bar{q}_{jt}$ must be smaller or equal to q_i. I assume that researchers observe equilibrium hiring standards of all firms in market t $\bar{q}_t = (\bar{q}_{1t}, ..\bar{q}_{Jt})$ for all t. Equilibrium hiring standard $\bar{q}_{i,t}$ naturally enters the selection function g because if firm j is popular and has a high equilibrium hiring standard, then the probability of being matched with j is lower.

Firm preferences can differ in firm-market level observables, $x^f_{jt} \subseteq x_{jt}$, which is a vector of firm characteristics that influences firm j's preference for candidates in market t. For example, a firm conducting business in a specific foreign country might prefer candidates with relevant language skills. Therefore, x^f_{jt} enters into the selection function g to capture preference heterogeneity.

Example 1 *One example of the selection function g is the setting where firms evaluate candidates by using an exam that reveals a weighted sum of their individual qualifications plus noise (ϵ_{ijt}). Therefore, before taking the exam, candidate i with q_i has a probability*

$$g\left(q_i, \bar{q}_{jt}, x^f_{jt}\right) = \Pr\left((q_i)^T x^f_{jt} + e_{ijt} \geq \bar{q}_{jt}\right)$$

[10]The uncertainty may come from different sources. First, it is possible that the screening process has some built-in randomness. In other scenarios where the randomness in the hiring process is not fully controlled by the screening process, the g function may contain some additional parameters that are not part of firm preferences. The non-parametric identification result can still be applied in these scenarios if instruments that shift the randomness in the screening process that is independent of the firm preferences are available.

of being hired by firm j in market t. Here, firms assign different weights to agent qualifications based on their characteristics x_{jt}^f.

Agent Preferences. The indirect utility of agent i conditional on being matched with firm j in market t is denoted as $f\left(x_{jt}^d, w_i\right)$. Here, w_i is a vector representing individual heterogeneity in tastes or traits, and it is randomly drawn from the distribution F_w. $x_{jt}^d \subseteq x_{jt}$ is a vector of firm characteristics in market t that influences agent i's utility upon being matched with firm j. For example, x_{jt}^d may contain a market t specific wage set by firm j. F_w is known to both firms and agents but remains unknown to the researchers. Nevertheless, w_i is only observable by agent i, i.e., neither firms nor researchers can observe it.

The indirect expected utility of agent i when applying for firm j comprises agent i's perceived preference for firm j. This is expressed as

$$v_{ijt}^* = \underbrace{g^*\left(q_i, \bar{q}_{jt}^*, x_{jt}^f\right)}_{\text{perceived chance}} \quad \underbrace{f\left(x_{jt}^d, w_i\right)}_{\text{utility of being matched with } j} + \delta_{jt} + \underbrace{\epsilon_{ijt}}_{\text{i.i.d. pref. shock}} \tag{2.1}$$

Here, $g^*\left(q_i, \bar{q}_{jt}^*, x_{jt}^f\right)$ denotes agent i's perceived likelihood of being hired by firm j, with $\bar{q}_{jt}^*$ being agent i's perceived equilibrium hiring standard of firm j in market t. $\epsilon_{ijt} \sim F_\epsilon$ is an independently and identically distributed preference shock for each agent i when considering firm j, which is observable only to the candidate i. δ_{jt} is a preference shifter at firm j and market t level affecting candidate application choices. For instance, δ_{jt} can represent a firm-specific application cost. I assumed that $\delta_{jt} = x_{jt}^a \beta_j + \xi_{jt}$ with $E[\xi_{jt}|x_{jt}^a] = 0$. Here, $x_{jt}^a \subseteq x_{jt}$ is a vector of firm j's characteristics in market t that only impact agents' indirect utility at the application stage.[11] ξ_{jt} is a one-dimensional, firm-specific characteristic observable to all agents but not to the researcher. I highlight that the indirect utility v_{ijt}^* possesses information on both agent preferences for firms and agent perceived firm preferences.

[11] The main identification result remains valid under minor modifications if I relax the linearity of δ_{jt} by assuming that $\delta_{jt} = G(x_{jt}^a) + \xi_{jt}$ with G being strictly increasing.

Motivated by the empirical application, I first consider scenarios where each agent i can apply for *one* firm j in a market t. I extend the identification result to scenarios where agents can apply to multiple positions in Appendix A.2. Agent i chooses *one* firm j that offers the highest expected utility, $j = \underset{j \in J}{\text{argmax}}\ v^*_{ijt}$. As is standard in discrete choice problems, utilities are normalized relative to position $j = 0$ for each agent, denoted as $v^*_{i0t} = 0$.

Equilibrium. I define a Perfect Bayesian Equilibrium for the matching game: (1) Each agent i applies for *one* firm j that maximizes her utility, conditional on firms' preferences and equilibrium hiring standards; (2) Each firm j hires agents according to its preference subject to the number of candidates demanded λ_{jt}. (3) On the equilibrium path, agents' beliefs regarding firms' hiring policies and the set of equilibrium hiring standards ($\bar{q}_{jt}$) for all j in market t, are correct.

An important requirement imposed by the equilibrium definition is that agents must be able to form correct expectations regarding each firm j's preference ($g^* = g$) and its equilibrium hiring standard ($\bar{q}^*_{jt} = \bar{q}_{jt}$) when formulating their expected utility, i.e., $v_{ijt} = v^*_{ijt}$.[12] When agents hold correct beliefs about each firm's preference for agents and equilibrium hiring standards, their application decisions contain information regarding both agent and firm preferences. The model is an example of how non-market systems solve the allocation problem in the absence of prices. The set of equilibrium hiring standards $\bar{q}_t$ adjust to to ensure no firms have empty slots, playing a role analogous to prices in competitive equilibrium models. In equilibrium, an agent's decision on where to apply impacts all other agents' decisions through the equilibrium hiring standards.

Given the above setup, the number of agents applying for firm j in market t can

[12] Practically, a rationale for the validity of this assumption is that agents can infer the equilibrium hiring standards through historical data. In the empirical analysis, I provide evidence in the Appendix O.2.1 suggesting that realized hiring standards for the same position in China's NCSE in different years are, on average, very similar.

be expressed as

$$s_{jt} = \sigma_j(x_t, \bar{q}_t, \xi_t) = \Pr\left(j = \underset{k\in\{0,...,J\}}{\text{argmax}} \quad v_{ikt} \middle| x_t, \bar{q}_t, \xi_t\right)$$

$x_{jt} = (x^f_{jt}, x^d_{jt}, x^a_{jt})$ is a vector collecting all firm-level characteristics. I define $s_t = (s_{1t}, ..., s_{Jt})$, which is observable to researchers. Let $\sigma = (\sigma_1, ..., \sigma_J)$. Function $\sigma_j(x_t, \bar{q}_t, \xi_t)$ encompasses all the unknown objects that I introduced previously. Specifically, it includes firm preferences g, agent preferences f and β_j, as well as individual heterogeneity represented by distributions F_q, F_w and F_ϵ. Moving forward, I focus on identifying the function σ_j, upon knowing equilibrium hiring standards q_t, firm characteristics x_t and share of applicants s_t.

Non-parametric Identification. I divide the variable x^a_t into $(x^{a(1)}_t, x^{a(2)}_t)$, with $x^{a(1)}_t \in \mathbb{R}$, and a further normalization where $\beta^{(1)}_j = 1$. Therefore, I obtain $\delta_{jt} \equiv x^{a(1)}_t + \xi_{jt}$, with remaining terms incorporated into the function σ_j. The assumption of connected substitutes, as in Berry, Gandhi and Haile (2013), is then applied to the application function $\sigma_t(\delta_t, \bar{q}_t)$.

Assumption 1 ***Connected substitutes****: (1) $\sigma_k(\delta_t, \bar{q}_t)$ is non-increasing in δ_{jt} for all $j > 0$, $k \neq j$, and any $(\delta_t, \bar{q}_t)$; (2) for each $(\delta_t, \bar{q}_t) \in supp(\delta_t, \bar{q}_t)$ and any nonempty $\mathcal{K} \subseteq \{1, ..., J\}$, $\exists k \in \mathcal{K}$ and $l \notin \mathcal{K}$ such that $\sigma_l(\delta_t, \bar{q}_t)$ is strictly decreasing in δ_{kt}.*

Assumption 1 is a standard one in the discrete choice literature. It requires that firms are weak substitutes concerning the indices δ_{jt}, i.e., slightly increasing δ_{jt} must weakly decrease the desirability of other firms. Additionally, it requires enough strict substitution among firms such that there is no strict subset of firms that substitute only among themselves. Berry, Gandhi and Haile (2013) shows that a broad spectrum of demand models satisfies the assumption 1. By leveraging Assumption 1, demand can be inverted:

$$x^{a(1)}_{jt} + \xi_{jt} = \sigma^{-1}_j(s_t, \bar{q}_t), j = 1, ..., J$$

I apply the identification results in Newey and Powell (2003), as followed in Berry and Haile (2014), to the matching game. This demonstrates the non-parametric identification of agents' and firms' preferences in the matching game, contingent upon the availability of instruments.

Assumption 2 ***Instruments****: (1) For all* $j = 1, ..., J$, $E\left[\xi_{jt} \mid z_t, x_t^{a(1)}\right] = 0$ *a.s., (2) For all functions* $B(s_t, \bar{q}_t)$ *with finite expectation, if* $E\left[B(s_t, \bar{q}_t) \mid z_t, x_t^{a(1)}\right] = 0$ *a.s. then* $B(s_t, \bar{q}_t) = 0$ *a.s..*

Proposition 1 *Suppose* $\{s_t, x_t, \bar{q}_t, z_t\}$ *are observable and that Assumptions 1 and 2 hold. The function* σ_j *is identified for all* j*.*

In Theorem 1, I demonstrate that the proposed matching market model is non-parametrically identified in the absence of individual-level data. The non-parametric identification is contingent on the completeness of the instruments (Assumption 2). In particular, the instruments must move the endogenous variables (s_t and $\bar{q}_t$ here) sufficiently to ensure that any function of these variables can be differentiated from others based on the exogenous variation provided by the instruments. Satisfying the completeness condition for the proposed job match model requires that instruments should be able to move agents' preferences/types and firm preferences sufficiently.[13]

The intuition is that agents, when making application choices, must take into account the probability of being matched with a firm, which indicates that agents' application choices carry information about selection criteria of all firms. Lastly, I emphasize that the entire analysis hinges on the simultaneous observability of both the number of applicants and the equilibrium hiring standards for all firms. This is critical because the equilibrium hiring standard offers valuable information about firms and candidates.

[13] Another set of potential unknown parameters that may affect the matching outcome is the one that governs the noise in the firm preference g. In this case, it is required that instruments also shift the noise distribution sufficiently.

One threat to identification is: Holding all other model elements fixed, are there more than one set of equilibrium hiring standards that are consistent with the matching game model? If multiple equilibria exist, it is more appropriate to establish set identification. This underscores the importance of understanding the conditions under which there is a unique set of $\bar{q}_t$.

Similarly to defining σ_j, I write π_j as a function of x_t, $\bar{q}_t$ and ξ_t that represents the number of agents who apply for position j in market t and meet the equilibrium hiring standard $\bar{q}_{jt}$. $\pi_j(x_t, \bar{q}_t, \xi_t) \leq \lambda_{jt}$ must hold due to firm j's slot constraint in market t. I impose the connected substitutes assumption on the function π.

Assumption 3 *(1) $\pi_k(\bar{q}_t)$ is non-increasing in $-\bar{q}_{jt}$ for all $j > 0$, $k \neq j$, and any $(\bar{q}_t)$; (2) for each $\bar{q}_t \in supp(\bar{q}_t)$ and any nonempty $\mathcal{K} \subseteq \{1, ..., J\}$, $\exists k \in \mathcal{K}$ and $l \notin \mathcal{K}$ such that $\pi_l(\bar{q}_t)$ is strictly decreasing in $-\bar{q}_{kt}$.*

Assumption 3 shares a similar intuition as the Assumption 1 on job application function σ. In a case where there are only two firms, $j = 1, 2$. This assumption requires that if firm 1 raises its hiring standard, then firm 2 must correspondingly hire more candidates.

Lemma 1 *Given assumption 3 and holding all model components as given except for $\bar{q}_t$, there exists a unique set of equilibrium hiring standards $\{\bar{q}_t\}$ that clears the matching market conditional on applications.*

Proof. With Assumption 3, function π is injective in $\bar{q}_t$ by the main results of Berry, Gandhi and Haile (2013). ■

Whether or not the matching game model implies a function π satisfying Assumption 3 requires case-by-case scrutiny for specific functional forms. In the empirical analysis, I verify that the proposed empirical model satisfies Assumption 3.

3 Empirical Analysis: the National Civil Service Exam

In this section, I apply the theoretical framework and identification result to study the matching mechanism design of government recruitment in China.

3.1 Institutional Background and Data

The National Civil Service Exam (NCSE) is the major way for China's central government to hire government officials through meritocratic selections.[14] Besides positions in departments' Beijing headquarters, the NCSE also fills positions in their branch offices nationwide at the provincial, prefecture, and county levels. A prominent example is the National Tax Bureau of China. Although its headquarter resides in Beijing, it also maintains local branches in every county of China, with all positions filled via the NCSE.

Figure 1 presents the structure of the hiring process for a cycle of the NCSE. Typically, in early October, the SCS posts detailed job openings online. Prospective candidates can visit the application system developed by the SCS to browse available positions with corresponding restrictions.[15] Applicants are constrained to apply for only *one* position per exam, meaning that this is not a strategy-proof mechanism. Following completion of the application stage, the SCS evaluates whether the basic information of each examinee satisfy the restrictions.

The SCS typically holds its nationwide exam towards the end of November after the application stage. The exam is meritocratic, with standardized exam paper divided into two segments: an ability test (50%) and an essay (50%).[16] The ability test

[14] For a more comprehensive discussion of the civil service fever in China over the past decade, see Feng et al. (2023). Appendix O.5.1 provides a comprehensive list of all the recruiting departments through NCSE during the sample period (2016 and 2017).

[15] There are some general restrictions on who can participate in the test. For example, only Chinese citizens can register for the exam, and the minimum age requirement is 18. For most positions, applicants have to be younger than 35 years old. Beyond that, each position has position-specific requirements. For example, some positions require applicants to hold a bachelor's degree. Other restrictions include minimum work experience, political status, and college major.

[16] Certain positions, such as those in the police force and the China Insurance Regulatory Commission (CIRS), require specialized exams. In the main empirical analysis, positions requiring specialized skills are excluded.

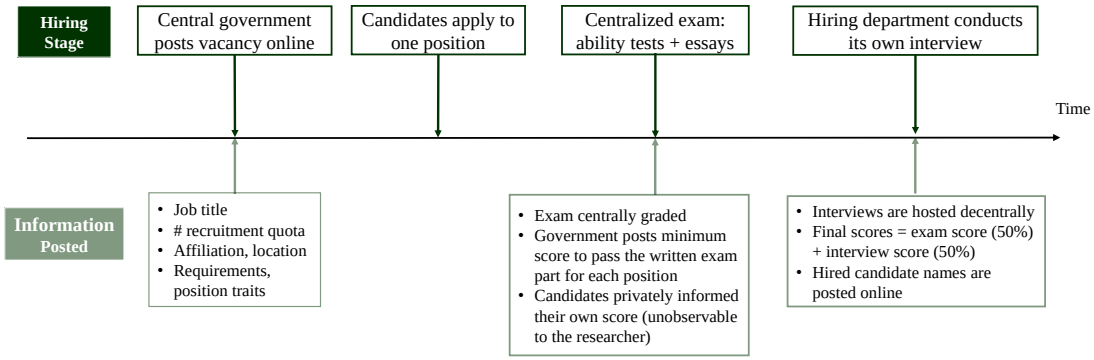


Figure 1: Stylized Structure of Hiring Processes of the NCSE

for positions affiliated to departments at the provincial level or above has roughly 10% more challenging questions compared to the exam paper for other positions. Online Appendix O.6 provides an official guide of the examination paper provided by the SCS, supporting that the exam is meritocratic. The SCS organizes the grading process after the exam. For each position, SCS ranks all applicants based on their scores and establishes a minimum score required for advancing to the interview phase.[17] The minimum qualifying score for passing the written exam phase (i.e., qualifying for the interview) for each position is publicly accessible online, and individual candidates also privately receives their own scores.[18] After the interview conducted by each department, the recruiting department makes the final hiring decision based on both the written examination score (50%) and the interview performance (50%). The list of selected candidates is posted online. Exam takers who do not pass the NCSE in a given year can retake it the following year, provided they remain eligible.

The NCSE in China provides an intriguing context to study matching mechanism design in government recruitment. First, it has an formalized application and recruitment system. Second, it posts many positions with diverse characteristics across

[17]There is a national minimum score that any applicant needs to pass in order to be hired, which is typically very low.

[18]It's noteworthy that if certain positions remain unfilled after the exam and the interview, the SCS offers those who failed for other positions a secondary opportunity to apply. I do not focus on this part of the process because of a data constraint. It is not a particular concern to the study since unfilled positions are often considered to be non-desirable.

locations, providing rich regional variation. Third, in China, an economy with corruption opportunities, the distinct motivations of government officials play a significant role for governance.

Dataset. The job posting information for the NCSE comes directly from the SCS, which details job characteristics and basic requirements for each position. The exam results for NCSE 2016 and 2017 were hand-collected from a prominent company that provides exam preparation materials. The data of The NCSE results is reliable, as the company uses it to help prospective exam candidates in their position choices. For each position, I obtain the number of applicants and the minimum score necessary for passing the written exam phase for each position. I merged the NCSE results data with the job posting data. To determine the location of each posted position, I manually matched the position information to the corresponding prefecture location.[19]

Table 1: Summary Statistics of the NCSE Dataset

Variables	NCSE 2016		NCSE 2017	
	Mean	Std. Dev.	Mean	Std. Dev.
Recruitment quota	1.776	1.480	1.736	1.205
# of applicants	84.015	146.397	87.300	157.770
Cutoff score	113.349	16.212	114.910	21.793
App-to-Recruit Ratio	55.581	97.743	58.146	122.898
Positions	14,146		14,502	

Notes: I only include positions that have a positive number of recruitment quota, # of applicants, cutoff score, App-to-Recruit Ratio (applicants to recruitment quota ratio, or $\frac{\#\text{ applied for } j}{\#\text{ demanded by } j}$), $\frac{\text{interview }\#}{\text{enrollment }\#}$ and a identified position location.

Table 1 presents summary statistics of the collected NCSE dataset. The statistics reveal that the NCSE attracts substantial attention, as the average number of

[19]Two data limitations prevent me from constructing a county level measure: 1, the county location of many NCSE positions cannot be identified; 2, several variables important in the later analysis, for example the measure of non-private sector wage, are unavailable in the county level statistics year book.

applicants per position exceeds 85. Even accounting for the number of enrollments, there are still on average more than 55 applicants competing for each available opening. However, the large standard deviation indicates significant variation in applicant numbers across positions, suggesting significant differences in their popularity.

3.2 The Empirical Model

Though the identification result is non-parametric, non-parametrically estimating a discrete choice model without individual-level data presents a significant challenge. I specify a parametric model to capture key features of the NCSE building on the framework proposed in Section 2.

I omit the market subscript to detail the model for a specific market. *Agent* i is characterized by ability a_i and civic-mindedness c_i, which are private information of each agent. a_i is i.i.d. drawn from a cumulative distribution function (CDF) $F_a = \mathrm{N}(\mu_a, 1)$. I assume that the mean ability level of candidates in a given market is a linear function of the GDP per capita of that province, i.e., $\mu_a = \theta_\mu GDP_m + \bar{\theta}_\mu$. This captures the education differences between matching markets. In Appendix O.2.2, I show that provinces with a higher GDP per capita have higher average cutoff scores for passing the written exam phase. Civic-mindedness , c_i, is assumed to follow normal distribution $F_c = \mathrm{N}(\mu_c, \sigma_c^2)$. The distributions of ability and civic-mindedness are assumed to be independent. In Appendix A.4, I present survey evidence from China that supports the independence of ability and civic-mindedness among individuals interested in government employment. It is also supported by experimental evidence in the literature.(Hanna and Wang, 2017; Barfort et al., 2019).

The centralized exam is modeled as a noisy signal of the ability level of each exam taker, which produces an exam score $\text{score}_i = a_i + \eta_i$ for each agent i. η_i is i.i.d. drawn from a normal distribution $\mathrm{N}(\mu_\eta, 1)$. Since the province exam paper is designed to be harder than the prefecture exam paper. I account for the exam difficulty differences by nomalizing the mean noise for the prefecture exam paper and introducing a mean noise for province exam paper μ_η. Given the exam structure, an

agent with ability a_i has a probability $\Phi(a_i - \bar{q}_j - \mu_\eta \mathbb{1}\text{Prov. Exam})$, of being matched with position j. This corresponds to the selection function (g) in equation (2.1) and represents the preference of the government. In equilibrium all exam takers accurately anticipate equilibrium passing scores. Appendix O.2.1 shows empirical evidence that this is plausible in the NCSE because equilibrium passing scores for the same position across different years tend to be similar.

The utility that an agent i derives from matching with position j is $u(x_j, m_j|c_i) = \theta_x' x_j + c_i m_j$. The vector x_j consists of the administrative level of the position's director, whether position j is high status and powerful and a measure of wage for position j. Appendix A.3.2 provides a detailed definition regarding which positions are considered as high status and power. For example, tax departments generally have higher status and are more powerful than other departments, and law enforcement positions hold more power than entry-level secretary positions. The measure of wage is constructed using the average urban non-private units of the prefecture where position j is located. Appendix O.5.5 offers more details.

I assume that candidates with lower civic-mindedness ($c_i > 0$) derive greater utility from being matched with positions with high potential rent-seeking opportunities, as this allows them to potentially extract more rents. Conversely, candidates with higher civic-mindedness derive less utility from positions with potential rent-seeking opportunities. Since these positions are known to have a pervasive rent-seeking atmosphere, high civic-mindedness candidates can feel uncomfortable working there. Especially because high rent positions may have many low civic-minded incumbent with whom those of a high civic-mindedness might prefer not to work. I use city-level cumulative number of past corruption cases (m_j) prior to the application deadline to capture candidates' perceived rent-seeking opportunities for each position. I choose convicted officials at relatively high administrative levels for constructing m_j because these cases attract significant public attention, making them more likely to be known

by exam takers.[20] Empirical evidence by Sun, Zhu and Ni (2022), Wang and Dickson (2022) and Pan, Shu and Ye (2023) supports that people in China interpret more corruption cases as revealing the potential rent-seeking opportunity.

The expected payoff that agent i derives when applying for position j is

$$U_{ij}^{e}(a_i, c_i, \bar{q}_j) = \underbrace{\overbrace{\Phi(a_i - \bar{q}_j + \mu_\eta \mathbb{1}_{\text{Prov. Exam}})}^{\text{prob of being matched with } j} \overbrace{u(x_j, m_j | c_i)}^{\text{utility from } j} + \delta_j}_{\text{Deterministic part } U_{ij}(a_i, c_i)} + \epsilon_{ij} \tag{3.1}$$

ϵ_{ij} follows an Extreme Value Type 1 distribution. δ_j is a position-specific utility affecting only the application decision and can be interpreted as an application cost. Similar as Equation 2.1, δ_j is assumed to follow $\delta_j = \theta_\delta X_j^\delta + \xi_j$ where X_j^δ represents position characteristics that affect decision-making regarding applications but do not directly affect the utility derived from obtaining a position.[21] For instance, it's assumed that being matched with a position that requires all candidates to possess two years of work experience does not deliver additional utility to the candidate upon a match. $\theta_\delta X_j^\delta$ is the cost of achieving a specific application requirement. Here, I assume that candidates, regardless of their abilities and civic-mindedness, incur the same cost to meet a specific qualification requirement. ξ_j is unobserved position-level characteristics, which influences exam takers' position choices at the application level.

The number of candidates that a position j demands is denoted as λ_j. I assume that the government's labor demand is exogenously given since the headcount of the government is centrally controlled by the State Commission Office for Public Sector Reform (SCOPSR).[22]

Since there is no available information about the population that does not participate in the NCSE, I normalize all positions with more labor demand than ap-

[20] Empirically, m_j is constructed by counting the total number of convicted government officials at the Deputy department level or higher in a prefecture. Cases typically involve (vice) Mayors or (deputy) departmental heads at the provincial level.

[21] See Appendix O.5 for detailed descriptions.

[22] Several other studies on government employment have also assumed that the supply of government jobs is exogenously determined, for example Postel-Vinay and Turon (2007), Burdett (2012) and Bradley, Postel-Vinay and Turon (2017)

plicants to be the 'outside' option ($\lambda_j \geq s_j$), i.e., anyone can get it once he/she applies. After normalization, the probability that an agent i chooses position j is $Pr_{ij}(a_i, c_i, \bar{q}_j) = \frac{\exp\left(U_{ij}(a_i, c_i, \bar{q}_j)\right)}{1 + \sum_j \exp\left(U_{ij}(a, c_i, \bar{q}_j)\right)}$.

The interview process is excluded from the analysis due to a data limitation, which I argue is not a big concern since most of the selection takes place during the examination stage. Specifically, the central government requires that if a position aims to hire one candidate, typically three candidates are permitted to participate in the interview process. Nevertheless, more than 18 candidates are competing for one interview slot according to Table 1. In the analysis, I assume that candidates face random selection during the interview process.[23]

Integrating over types, the share of candidates who apply for position j is

$$s_j = \int_{\mathbb{R}} \int_{\mathbb{R}} \frac{\exp\left(U_{ij}(a_i, c_i, \bar{q}_j)\right)}{1 + \sum_j \exp\left(U_{ij}(a_i, c_i, \bar{q}_j)\right)} dF_c(c_i) dF_a(a_i) \tag{3.2}$$

s_j denotes the observed share of applicants for position j in an application market. Given the application decisions and exam results, the central administration chooses cutoff scores $\bar{q}_j$s to fill all posted positions. Specifically, for any positions j, the following equality must hold when position j is selective:

$$\lambda_j = \int_{\mathbb{R}} \underbrace{\Phi(a_i - \bar{q}_j + \mu_\eta \mathbb{1}_{\text{Prov. Exam}})}_{\text{prob of agent with } a_i \text{ passing } \bar{q}_j} \underbrace{\left[\int_{\mathbb{R}} \frac{\exp\left(U_{ij}(a_i, c_i, \bar{q}_j)\right)}{1 + \sum_j \exp\left(U_{ij}(a_i, c_i, \bar{q}_j)\right)} dF_c(c_i) \right]}_{\text{number of applicants for position } j} dF(a_i) \tag{3.3}$$

I establish that the set of equilibrium cutoff scores is unique by invoking Lemma 1. Please refer to Online Appendix O.1.1 for the calculation.

Proposition 2 *Fixing a set of utility parameters and number of candidates demanded by each position j $\{\lambda_j\}$, I define the RHS of equation (3.3) as $\pi_j(\bar{q}_1, \bar{q}_2, ..., \bar{q}_j)$ $\forall j > 0$*

[23] The analysis here can be better understood as an exercise that explores the talent pool composition prior to the interview process, with the meritocratic exam serving as a mechanism that imposes constraints on subsequent interviews. One may also wonder whether the interview can be a useful tool to screen for civic-mindedness. Based on the organization of interviews in the NCSE, it is unlikely that the interviewers have enough time and flexibility to elicit candidate civic-mindedness. Appendix A.3.1 provides detailed information supporting this point.

and show that $\frac{\partial \pi_j}{\partial \bar{q}_j} < 0$ *for all* j *and* $\frac{\partial \pi_j}{\partial \bar{q}_{j'}} > 0$ *for all pairs of* j *and* j'*. Consequently, there exists a unique set of* $\bar{q}_j$ *that clears matching markets according to Lemma 1.*

The NCSE restricts each agent to apply for *one* position before the exam, leading to strategic application choices. Candidates do not always choose a position maximizing the utility conditional on being matched. Moreover, the difference in measurability leads to a unique sorting pattern in equilibrium. Low civic-mined candidates apply more frequently to high rent jobs pushes up the equilibrium cutoff scores for those positions, further discouraging candidates with higher civic-minded from applying. Consequently, in equilibrium, high rent positions are more likely to match with low civic-minded candidates.

3.3 Estimation and Empirical Matching Pattern

The sample used for structural estimation is constructed using the 2016 and 2017 National Civil Service Exam (NCSE) samples, with several exclusions. Please refer to Appendix A.3.3 for more details about the exclusions.

I define each province as an application market t, reflecting exam takers' home preferences. I assume that all candidates who choose a position from application market t after evaluating all positions in market t.[24] I exclude positions that require specialized skills because those positions attract a group of candidates with specific expertise in the required skills, making them less likely to apply for other more generic positions.[25] Additionally, I omit positions in *Hainan*, *Xinjiang*, and *Tibet* due to their unique administrative structures.[26]

[24] I exclude positions from the four municipalities: *Beijing*, *Shanghai*, *Tianjin*, and *Chongqing*. Though municipalities in China are provincial-level administrative units, they cover significantly smaller areas compared to typical provinces. Consequently, there is little variation in prefecture-level characteristics within these municipalities, making it challenging to treat any of them as a distinct application market.

[25] Positions requiring specialized skills include police-related positions for railways, translating positions requiring special language skills, and positions requiring meteorology, sailing, computer science, and medical knowledge.

[26] Another reason is that the "China City Statistical Year Book" provides limited coverage for the three regions. For instance, while Xinjiang comprises 14 prefectures, the Year Book only includes data for 2 of them for almost all statistics.

I aggregate positions within a prefecture by department types, the administrative level of the director of that position and exam types.[27] I calculate the average position characteristics, average cutoff scores and average number of applicants for the constructed representative position. For example, if the State Agency of Tax opens 5 positions in a prefecture for the 2016 NCSE year and all positions are evaluated using the prefecture exam, then I combine them into one position. The combined position is assigned position-level characteristics equal to the average of all 5 positions.

The reasons for aggregating positions at the prefecture $\times$ department types $\times$ exam types level are twofold. First, the average number of positions within an application market is 349 without aggregation. It is unlikely that candidates evaluate so many positions when applying. Such a large consideration set also poses difficulty in estimation. Additionally, I have only prefecture variations in average wage and corruption cases. Since a main goal of the empirical exercise is to estimate the sorting pattern of civic-mindedness among NCSE exam takers, aggregating the positions at the prefecture $\times$ department types $\times$ exam types level does not represent a significant loss in terms of the information. After aggregation, I obtain 48 application markets and 1,853 combined positions, indicating that there are on average 38 aggregated positions in an application market for agents to choose from. Table A.1 presents the summary statistics of the sample constructed for the estimation. For detailed summary statistics of the positions requirements (X_j^δ), please refer to Appendix A.3.4

I classify parameters of interest into two sets, i.e., non-linear parameters (NLP) $\Theta = \{\theta_x, \theta_\mu, \bar{\theta}_\mu, \mu_c, \sigma_c, \eta_p\}$ and linear parameters (LP) θ'_δ. I modify the nested-fixed point algorithm originally proposed by Berry, Levinsohn and Pakes (1995) by imposing a slot constraint on the minimization problem due to the fixed number of slots for each position j. I derive a Generalized Method of Moments (GMM) estimator from the

[27] Because of the large number of departments participating in the NCSE and the fact that not all departments recruit every year, I categorize all departments into 9 department types based on their functional similarities. For the empirical analysis, I use department types instead of departments. For further details, please refer to Online Appendix O.5.

population moment condition $E[Z_j \xi_j] = 0$. I employ a standard two-step procedure for GMM estimation. Please refer to Appendix O.3.1 for additional details of the estimation.

The insight of the non-parametric identification result shown in Theorem 1 is that instruments sufficiently shifting both candidates' preferences/types and government preferences are needed. I use three sets of instruments in the estimation.

The government preference is represented by the parameter η_p, which captures the difference between the mean noise of two exam (province or prefecture). The first instrument I use to identify the government preference is the assignment of exam papers to positions. The assignment is determined based on the level of the director of that position, i.e., long before the labor demand is planned. The second instrument used to identify the government preference is whether a position located in a sub-provincial level (or above) city. Since sub-provincial level (or above) cities are much more likely to have positions affiliated with a department at or above the provincial level, positions in those cities are more likely to use a province exam paper. Second, I use a set of prefecture-level variables to help identify candidate preferences for prefecture characteristics. I use the average prefecture characteristics for all other positions in the same application market as instruments. Additionally, I utilize the prefecture budgetary fiscal expenditure, which is mean-independent of any position-level unobserved characteristics or preference shocks since the government budgetary expenditure is planned at the beginning of the fiscal year before the exam. The third set of instruments is based on position characteristics, which assist in identifying candidate preferences for position characteristics. Specifically, I employ the average position characteristics for all other positions in the same application market.

Table 2 presents the estimated NLP (Θ) from the preferred specification. Consistent with the empirical patterns shown in Appendix O.2.2, application markets with a higher average GDP per capita attract candidates of correspondingly higher ability on average ($\theta_\mu > 0$). The province exam paper is found to be generally more

Table 2: Non-linear Parameter Estimates

Non-linear Para	Estimates and Standard Errors		
	$\bar{\theta}_\mu$ (Ability Dist Cons)	θ_μ (Ability Dist Slope)	μ_c (Civic Dist Mean)
Estimate	-7.345	5.972	2.450
Std. Error	(2.961)	(3.176)	(1.893)
	σ_C (Civic Dist Var)	η_{prov} (Exam Mean)	θ_x (Status / Power)
Estimate	0.580	-1.714	6.962
Std. Error	(0.332)	(0.704)	(4.991)
	θ_x (Level)	θ_x (Wage)	
Estimate	7.704	0.729	
Std. Error	(6.064)	(0.640)	

Notes: Please refer to Table O.3.1 in the Appendix O.3.2 for LP estimates θ_δ. Appendix O.3.1 present more details about the estimation. A Sargen - Hansen test (or Sargan's J test) is performed to check the validity of the instruments. I find that the p-value is larger than the significance level (0.05), meaning that the null hypothesis is not rejected and the instruments are considered valid. In Appendix O.3.3, I report estimates coming from alternative specifications.

challenging than the prefecture exam paper ($\eta_{prov} < 0$). Estimates on candidates' preferences ($\theta_x > 0$) suggest that positions characteristics, including status (powerful), administrative levels, and wages, are all desirable. Low civic-minded candidates express stronger preferences for positions with higher past anti-corruption cases than high civic-minded candidates, though candidates on average prefer positions with higher past anti-corruption cases ($\mu_c > 0$). Table O.3.1 in the Appendix O.3.2 reports the estimated LP (θ_δ). Linear parameters are interpreted as costs associated with requirements for the candidates, suggesting that these requirements may deter candidates from applying for the position. In the later analysis, it is sufficient to use the $\{\delta_j\}$ obtained in the estimation process to capture position-specific characteristics that shifts the application decision for each position j.

Using the set of estimates, I obtain the model predicted average ability (A_j) and civic-mindedness (C_j) of all matched candidates for each position j by the following

two equations.

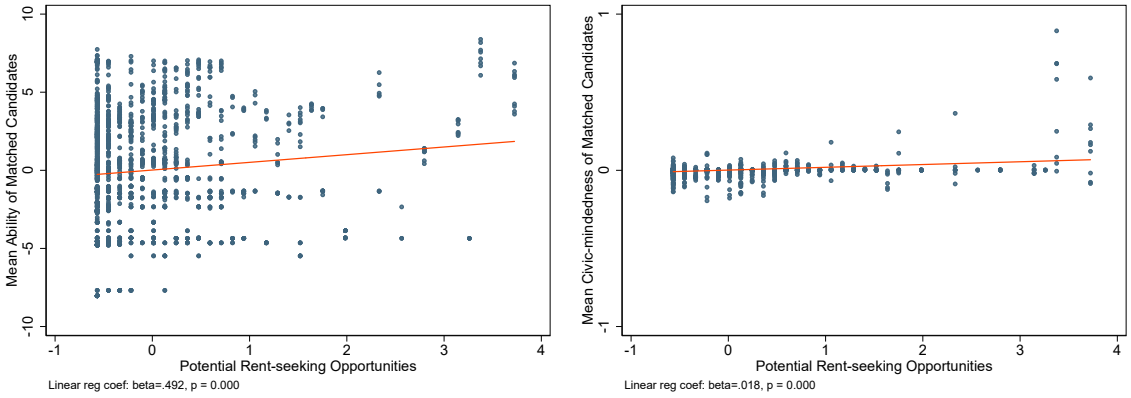
$$A_j = \frac{1}{\lambda_j} \int \int a_i \Phi(a_i - \bar{q}_j + \mu_\eta \mathbb{1}_{\text{Prov. Exam}}) Pr_{ij}(a_i, c_i) dF_c dF_a$$

$$C_j = \frac{1}{\lambda_j} \int \int c_i \Phi(a_i - \bar{q}_j + \mu_\eta \mathbb{1}_{\text{Prov. Exam}}) Pr_{ij}(a_i, c_i) dF_c dF_a$$

Here, $Pr_{ij}(a_i, c_i)$ represents the estimated choice probability derived from the empirical model and its estimates. Within the integrals, $\Phi(a_i - \bar{q}_j + \mu_\eta \mathbb{1}_{\text{Prov. Exam}}) Pr_{ij}(a_i, c_i)$ denotes the likelihood that a candidate of type a_i and c_i applies to and subsequently matches with position j.

3.4 Empirical Sorting Pattern

Upon estimating the position-level average civic-mindedness and ability induced by the NCSE, I study the empirical matching pattern.



Notes: The y-axis of the left figure represents the estimated average ability of all matched candidates ($\{A_j\}$) residualized using year fixed effects. I drop prefectures with past corruption cases that are in the 99th percentile or higher in the plot to prevent the influence of outliers. The y-axis of the right figure represents the estimated average civic-mindedness of all matched candidates ($\{C_j\}$) residualized by year fixed effects. For each figure, I draw a linear fit prediction.

Figure 2: Multi-dimensional Sorting: Potential Rent-seeking Opportunities

The left panel of Figure 2 indicates that high rent positions are more likely to be matched with high ability candidates on average. This empirical pattern emerges because positions with higher potential rent-seeking opportunities are preferred by

candidates on average. Since the higher the C_j, the lower the average civic-mindedness of position j, the right panel reveals that high rent positions hire candidates with lower civic-mindedness on average, though the magnitude is not very large. The empirical matching pattern shown in right panel is consistent with the equilibrium characteristics discussed at the end of Section 3.2.

The magnitude of the correlation between civic-mindedness and past corruption cases is not very large, because the NCSE restricts candidates to apply for only one position before the exam, i.e., the potential competition and exam uncertainty refrain candidates with lower civic-mindedness to apply for positions that they prefer the most.

In Figure A.1, I present additional empirical evidence of sorting: high wage positions tend to match with candidates of higher ability. However, candidates, whether having lower or higher civic-mindedness, show similar likelihoods of being matched with these positions. This is expected because candidates with different civic-mindedness have similar preferences for wages.

3.5 Counterfactual Analysis

I consider a serial dictatorship mechanism as the alternative matching mechanism, where the priority is given to candidates who obtain a higher score in the exam. In practice, a similar structure for matching candidates to civil service positions is utilized by the Indian Administrative Service (IAS) (Thakur, 2021).

Mechanism 1 *A serial dictatorship mechanism with meritocratic priority ranking allocates positions to candidates following 5 steps: (1), Government posts vacancies for all candidates to consider. (2) Candidates' preference shocks for all positions are realized. (3) The central government organizes an exam for all candidates interested in getting a position. The exam is a noisy signal of candidate ability but candidate civic-mindedness remains unknown to the government. (4) Government obtains a ranking for all candidates according to the scores. (5) The candidate with the rth rank in the exam chooses in the rth place until all positions are allocated.*

The mechanism outlined above eliminates the exam uncertainly in the application and is strategy-proof because (1) obtaining a higher score increases the chance for a candidate to be matched with his/her most preferred position and (2) no candidates can obtain a more preferred position through untruthful reporting at the stage of choosing a position. An alternative way to design the mechanism is to ask all candidates to submit a full rank order list for all positions in the application market to which they belong before the government announces candidate scores and obtains a ranking for all candidates according to it. This alternative arrangement delivers the same matching outcome as mechanism 1 in the context of this paper when only one exam paper is used.

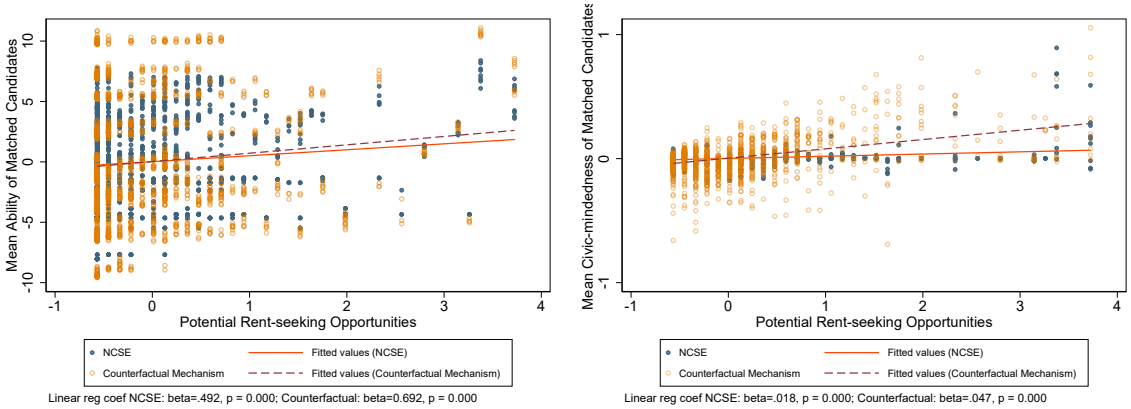
In the counterfactual analysis, I simulate 10,000 candidates each with an ability and a civic-mindedness in each application market. For each candidate, I simulate an exam score by adding a noise term to their intrinsic ability-level. Consistent with the empirical model, the exam noise is drawn from a standard Normal distribution. In the counterfactual analysis, I assume that there is only one exam paper that is used for all candidates.[28] A candidate's indirect utility for applying for a position is now different from what's specified in Equation (3.1) because, when the counterfactual mechanism 1 is implemented, candidates have already learned their ranking among all candidates in the application market. Specifically, candidate i's utility of being matched with position j is now $U^c_{ij}(a_i, c_i) = u(x_j, m_j|c_i) + \delta_j + \epsilon_{ij}$

Here, $u(x_j, m_j|c_i)$ and δ_j are part of the estimation results. I simulate preference shock ϵ_{ij} for each candidate-position pair. Collecting the matched candidates in each position, I calculate the average ability level and civic-mindedness level for each position j. I repeat the simulation 100 times to obtain the average matching pattern in the counterfactual analysis. I denote the simulated average ability and civic-mindedness of all matched candidates through the counterfactual mechanism 1 as $\{A^s_j\}$ and $\{C^s_j\}$ respectively.

[28]Since candidates differ in a single dimension (i.e., ability) from the government's perspective, it is sufficient to have only one exam for all candidates in the counterfactual analysis.

Sorting Pattern: Alternative Mechanism versus the NCSE

Given that the application occurs after the revelation of candidates' relative exam rankings, high ability candidates are more likely to be matched with a desirable position under the alternative mechanism compared to the NCSE. Therefore, candidates with high ability are more likely to be matched with generally desirable positions (high x_j). Consequently, the average ability of candidates for desirable positions is higher, while the average ability of candidates for less-desirable positions is lower compared to that generated by the NCSE. This reasoning extends to the analysis of civic-mindedness. Candidates with lower civic-mindedness are more likely to be matched with positions they prefer, i.e., positions with high corruption opportunities. As a result, compared to the NCSE, the counterfactual mechanism promotes a higher degree of sorting.



Notes: The y-axis of the left figure represents the average ability of matched candidates residualized by year fixed effects. I drop prefectures with past corruption cases that are in the 99th percentile or higher in the plot to prevent the influence of outliers. The y-axis of the right figure represents the average civic-mindedness of matched candidates for each position residualized by year fixed effects. The dots denote the same matching outcome estimates shown in Figure 2. The circles denote simulated matching outcomes based on the mechanism 1. For each figure and each matching outcome, I draw a linear fit prediction. The fit prediction for the NCSE is a solid line while the fit prediction for the counterfactual mechanism is a dashed line.

Figure 3: Multi-dimensional Sorting (Counterfactual) : Potential Rent-seeking Opportunities

Figures 3 depict the counterfactual sorting pattern in contrast to the estimated

sorting pattern derived from NCSE to illustrate the point. The left panel of Figure 3 demonstrates that under the alternative mechanism, the correlation between the average ability of matched candidates and potential rent-seeking opportunities is more pronounced than the correlation under the NCSE. Specifically, the linear regression coefficient exhibits an increase of about 40%. The correlation between the average civic-mindedness of matched candidates and potential rent-seeking opportunities appears much stronger (a 161% increase) than the correlation derived from NCSE. Figure A.2 shows that high ability candidates are more likely to be matched with positions with high average wages under the alternative mechanism comparing to the NCSE. Furthermore, it also shows that the average civic-mindedness of matched candidates and the prefecture's average wages under the counterfactual matching mechanism remains very close to 0. This aligns with expectations given that candidates prefer high wages regardless of their civic-mindedness.

Consequence on Government Performance

Since government jobs in China are desirable and government employees in China have very low quit rate (less than 0.015%), I explore the potential implications of these mechanisms on government performance by conducting a simulation exercise.[29] The goal is to underscore the significance of multi-dimensional sorting — specifically, ability and civic-mindedness sorting, especially when the latter is unobservable.[30]

Empirically assessing government performance is notoriously hard, especially for assessing the contribution of newly recruited candidates to departmental performance. Therefore, I posit a stylized government's public goods and services production function (equation 3.4), through which I can compare the cost and benefit of each mech-

[29] See the following news report for the low quit rate for government employees in China. http://union.china.com.cn/kx/txt/2018-01/30/content_40208798.htm

[30] I refer to Thakur (2021) for discussions on designing matching mechanisms that balance the trade-off between respecting candidate preferences and optimizing government performance in light of ability sorting.

anism in a transparent way.

$$Y = \int_j \left(W_j^a \times A_j^{\vartheta_a} - W_j^c \times \underbrace{\mathbb{1}_{\{C_j \geq 0\}} C_j^{\vartheta_c}}_{\text{low civic-mindedness is costly}} \right) dj \tag{3.4}$$

Y represents total government output, which can be public goods or services. $A_j^{\vartheta_a}$ represents how average ability of a position contributes to the government services or public goods. A position having candidates with higher average ability leads to higher government output. $\mathbb{1}_{\{C_j \geq 0\}} C_j^{\vartheta_c}$ in Equation 3.4 represents the cost of having low civic-minded government workers.[31] Employing candidates with lower civic-mindedness decreases the performance of position j because they may participate in rent-seeking activities. ϑ_a signifies the importance of ability in government production while ϑ_c signifies the importance of civic-mindedness. ϑ_c can be conceptualized as the cost of corruption or the effectiveness of government agencies against rent-seeking activities or the quality of other institutional arrangements to combat corruption. The *larger* the parameter ϑ_c, the *larger* the cost of corruption but the *lower* the effectiveness of government agencies against rent-seeking activities. I assign an equal ability weight ($W_j^a = 1, \forall j$) but assign different importance of civic-mindedness on positions $W_j^c \neq 1$ based on potential rent-seeking opportunities to capture that low civic-minded candidates incur greater costs when matched to high rent positions.

Empirically, I fix $\vartheta_a > 1$ and varying ϑ_c.[32] Figure 4 presents the result. The solid line depicts government output at various ϑ_c values under the NCSE, while the dashed line illustrates the same under the alternative mechanism. The analysis indicates that the NCSE outperforms the alternative mechanism when the cost of hiring low civic-minded candidates (ϑ_c) is larger than a threshold. Conversely, when

[31] I assume that high civic-minded candidates cannot bring additional benefit to the government. The assumption is made because of two primary reasons: (1) the introduction of an additional parameter complicates the analysis; and (2) majority of the positions are filled with candidates with low civic-mindedness on average (or $C_j > 0$).

[32] When only employee ability matters for government performance, then the alternative mechanism yields a better governance outcome if government performance is a convex function of position-level average ability ($\vartheta_a > 1$).

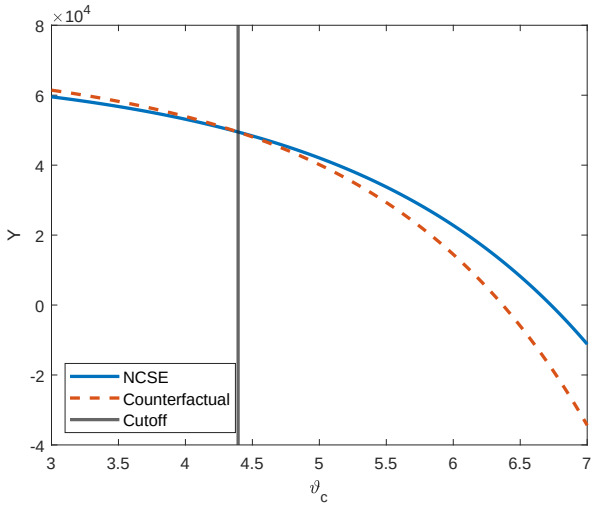


Figure 4: Implications on Government Performance: A Comparison

Notes: The Y-axis represents the government output Y. The solid line plots the government output given NCSE while the dashed line plots the government output given the alternative mechanism. I fix $\vartheta_a = 1.5$ in the simulation to focus on how the change of ϑ_c affect Y. The x-axis represents the importance of civic-mindedness in government production ϑ_c. A higher ϑ_c implies a more important impact of candidate civic-mindedness relative to the ability. The vertical line indicates the threshold value where the Ys induced by two mechanisms are equal. Position weights W_j^c is constructed based on the number of potential rent-seeking opportunities (past corruption cases).

the relative importance of civic-mindedness (ϑ_c) falls below this threshold, the alternative mechanism is superior. Different values of ϑ_a yield a similar pattern, with the primary distinction being a shift in the threshold. This pattern emerges because the alternative mechanism is more likely to match lower civic-mindedness with high rent jobs than the NCSE, inducing a significant cost to government production when civic-mindedness carries significant impact on government performance (a large ϑ_c relative to ϑ_a). In Online Appendix O.4, I demonstrate that the pattern displayed in Figure 4 is robust when I employ only weights on ability based on non-private sector wage or with both weights.

Alternatively, this exercise could be construed as a revealed preference analysis, offering insights into the underlying government objectives given the mechanism of NCSE and a strategy-proof mechanism. Since the Chinese government has designed and adopted the current NCSE mechanism for government recruitment, the above

simulation exercise suggests that it is very likely that low civic-minded government officials pose significant threats to the Chinese government.

4 Conclusion

This paper develops a new approach to the non-parametric identification of matching models when individual-level data is unavailable. I apply the model and identification results to study matching mechanism design in government recruitment, highlighting multi-dimensional sorting and information friction in the matching process. A central takeaway is that matching mechanisms promoting sorting may be less desirable when there are important, yet hard-to-detect, worker attributes relevant to job performance. The findings have broader implications for labor market design, particularly in situations where attributes like work ethics are pivotal for on the job performance but remain undetectable in standard hiring processes.

There are several directions in which the design of matching mechanism in government recruitment presented in this paper could be further developed. First, this paper primarily focuses on analyzing the written exam part of the National Civil Service Examination given data limitation. However, it is important to recognize that other methods, such as interviews, could potentially serve as effective tools for identifying candidate civic-mindedness in other contexts. Should such additional screening tools become available, they may change the sorting patterns in government recruitment. Moreover, this paper focuses on a fixed pool of candidates interested in those positions. One promising direction to improve our understanding about government recruitment includes integrating the self-selection process into government jobs with the mechanisms that determine the assignment of agents to their respective roles. I leave these extensions to future research.

References

Abaluck, Jason and Abi Adams-Prassl. 2021. "What do consumers consider before they choose? Identification from asymmetric demand responses." The Quarterly Journal of Economics 136(3):1611–1663.

Agarwal, Nikhil. 2015. "An empirical model of the medical match." American Economic Review 105(7):1939–1978.

Agarwal, Nikhil and Paulo J Somaini. 2022. Demand Analysis under Latent Choice Constraints. Technical report National Bureau of Economic Research.

Ashraf, Nava, Oriana Bandiera, Edward Davenport and Scott S Lee. 2020. "Losing prosociality in the quest for talent? Sorting, selection, and productivity in the delivery of public services." American Economic Review 110(5):1355–94.

Azevedo, Eduardo M and Jacob D Leshno. 2016. "A supply and demand framework for two-sided matching markets." Journal of Political Economy 124(5):1235–1268.

Barfort, Sebastian, Nikolaj A Harmon, Frederik Hjorth and Asmus Leth Olsen. 2019. "Sustaining honesty in public service: The role of selection." American Economic Journal: Economic Policy 11(4):96–123.

Barseghyan, Levon, Francesca Molinari and Matthew Thirkettle. 2021. "Discrete choice under risk with limited consideration." American Economic Review 111(6):1972–2006.

Barseghyan, Levon, Maura Coughlin, Francesca Molinari and Joshua C Teitelbaum. 2021. "Heterogeneous choice sets and preferences." Econometrica 89(5):2015–2048.

Becker, Gary S. 1973. "A theory of marriage: Part I." Journal of Political economy 81(4):813–846.

Berry, Steven, Amit Gandhi and Philip Haile. 2013. "Connected substitutes and invertibility of demand." Econometrica 81(5):2087–2111.

Berry, Steven, James Levinsohn and Ariel Pakes. 1995. "Automobile prices in market equilibrium." Econometrica: Journal of the Econometric Society pp. 841–890.

Berry, Steven T and Philip A Haile. 2014. "Identification in differentiated products markets using market level data." Econometrica 82(5):1749–1797.

Besley, Timothy. 2005. "Political selection." Journal of Economic Perspectives 19(3):43–60.

Besley, Timothy and Maitreesh Ghatak. 2003. "Incentives, choice, and accountability in the provision of public services." Oxford review of economic policy 19(2):235–249.

Besley, Timothy and Maitreesh Ghatak. 2018. "Prosocial motivation and incentives." Annual Review of Economics 10:411–438.

Besley, Timothy, Robin Burgess, Adnan Khan and Guo Xu. 2022. "Bureaucracy and development." Annual Review of Economics 14:397–424.

Best, Michael Carlos, Jonas Hjort and David Szakonyi. 2023. "Individuals and organizations as sources of state effectiveness." American Economic Review 113(8):2121–2167.

Bradley, Jake, Fabien Postel-Vinay and Hélene Turon. 2017. "Public sector wage policy and labor market equilibrium: a structural model." Journal of the European Economic Association 15(6):1214–1257.

Burdett, Ken. 2012. "Towards a theory of the labor market with a public sector." Labour economics 19(1):68–75.

Cattaneo, Matias D, Xinwei Ma, Yusufcan Masatlioglu and Elchin Suleymanov. 2020. "A random attention model." Journal of Political Economy 128(7):2796–2836.

Conlon, Christopher and Jeff Gortmaker. 2020. "Best practices for differentiated products demand estimation with pyblp." The RAND Journal of Economics 51(4):1108–1161.

Dahis, Ricardo, Laura Schiavon and Thiago Scot. 2023. "Selecting top bureaucrats: Admission exams and performance in brazil." Review of Economics and Statistics pp. 1–47.

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(3):1169–1218.

Deng, Yongheng, Shang-Jin Wei, Jing Wu and Rongjie Zhang. 2023. Secrets in Asset Purchases: Estimating the Unofficial Incomes of Officials. Technical report National Bureau of Economic Research.

Diamond, William and Nikhil Agarwal. 2017. "Latent indices in assortative matching models." Quantitative Economics 8(3):685–728.

Feng, Bo, He Qiwei, Jin Xin and Xu Xu. 2023. "State Building and Civil Servant Fever in Contemporary China." Working Paper .

Feng, Bo, Qiwei He and Hanzhang Liu. 2023. "Self-Selection into Government Employment in China: Individual Traits and (Mis)Beliefs." working paper .

Fenizia, Alessandra. 2022. "Managers and productivity in the public sector." Econometrica 90(3):1063–1084.

Francois, Patrick. 2000. "'Public service motivation'as an argument for government provision." Journal of Public Economics 78(3):275–299.

Gale, David and Lloyd S Shapley. 1962. "College admissions and the stability of marriage." The American Mathematical Monthly 69(1):9–15.

Hanna, Rema and Shing-Yi Wang. 2017. "Dishonesty and selection into public service: Evidence from India." American Economic Journal: Economic Policy 9(3):262–90.

He, YingHua, Shruti Sinha and Xiaoting Sun. forthcoming. "Identification and Estimation in Many-to-one Two-sided Matching without Transfers." Econometrica .

Heiss, Florian and Viktor Winschel. 2008. "Likelihood approximation by numerical integration on sparse grids." journal of Econometrics 144(1):62–80.

Molinari, Francesca. 2020. "Microeconometrics with partial identification." Handbook of econometrics 7:355–486.

Moreira, Diana and Santiago Pérez. forthcoming. "Civil Service Exams and Organizational Performance: Evidence from the Pendleton Act." American Economic Journal: Applied Economics .

Newey, Whitney K and James L Powell. 2003. "Instrumental variable estimation of nonparametric models." Econometrica 71(5):1565–1578.

Pan, Yating, Zhan Shu and Zhipeng Ye. 2023. "Exploring the dynamics of corruption perceptions in sustained anti-corruption campaigns: A survey experiment in China." Humanities and Social Sciences Communications 10(1):1–9.

Postel-Vinay, Fabien and Hélène Turon. 2007. "The public pay gap in Britain: small differences that (don't?) matter." The Economic Journal 117(523):1460–1503.

Spenkuch, Jörg L, Edoardo Teso and Guo Xu. 2023. "Ideology and performance in public organizations." Econometrica 91(4):1171–1203.

Sun, Zongfeng, Lin Zhu and Xing Ni. 2022. "How Does Anti-Corruption Information Affect Public Perceptions of Corruption in China?" China Review 22(2):113–143.

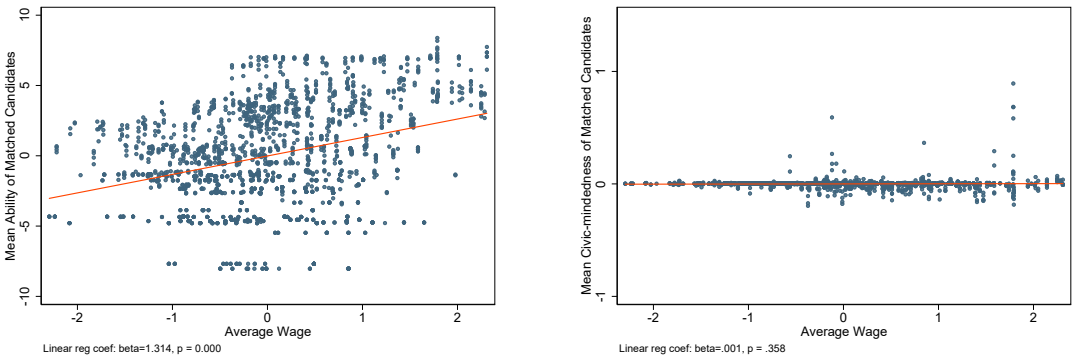
Thakur, Ashutosh. 2021. Matching in the civil service: A market design approach to public administration and development. Technical report ECONtribute Discussion Paper.

Varadhan, Ravi and Christophe Roland. 2008. "Simple and globally convergent methods for accelerating the convergence of any EM algorithm." Scandinavian Journal of Statistics 35(2):335–353.

Wang, Yuhua and Bruce J Dickson. 2022. "How corruption investigations undermine regime support: evidence from China." Political Science Research and Methods 10(1):33–48.

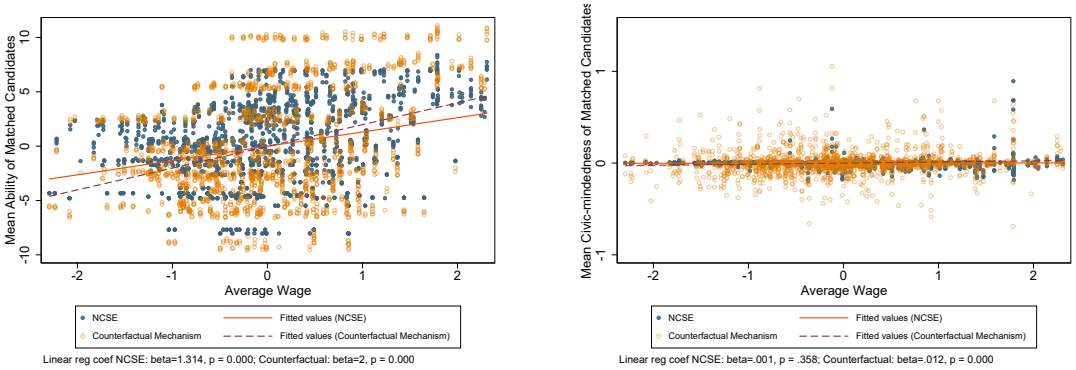
Appendices

A.1 Figures and Tables



Notes: The y-axis of the left figure represents the estimated average ability of all matched candidates ($\{A_j\}$) residualized using year fixed effects. The Y-axis of the right figure represents the estimated average civic-mindedness of all matched candidates ($\{C_j\}$) residualized by year fixed effects. For each figure, I draw a linear fit prediction plot.

Figure A.1: Multi-dimensional Sorting: Average Wages



Notes: The y-axis of the left figure represents the average ability of matched candidates for each position residualized by year fixed effects. The y-axis of the right figure represents the average civic-mindedness of matched candidates for each position residualized using year fixed effects. The dots are matching outcomes in Figure 2. The circles denote simulated matching outcomes based on mechanism 1. The fit prediction for the alternative mechanism is a dashed line.

Figure A.2: Multi-dimensional Sorting: Average Wages

Table A.1: Summary Statistics of the Sample Used for Estimation

	Mean	Std. Dev.
Panel A		
# of application markets (t)	48	—
# of positions per application markets (t)	38.604	—
Share of j using province exam paper	0.080	—
Panel B		
# of candidates demanded	4.318	2.166
# of applicants	91.192	94.941
Cutoff score	120.965	7.857
Panel C		
High Status / Powerful	0.323	0.43
Level	1.492	.609
Non-Private unit wage	57354.89	10743.84
Past corruption cases	4.889	8.623
Observations	1,853	

Notes: Panel A shows summary statistics for the application markets. Panel B summarizes three NCSE results that are used in the estimation. Panel C presents summary statistics for variables that matters for the utility conditional being matched with a positions. For summary statistics of the position requirements (X_δ), please refer to Table A.1 in the Appendix A.3.4.

A.2 Multiple Applications

In this subsection, I discuss how to extend the main identification result to the case where agents can make multiple application choices.

I assume that each agent i is allowed to apply to a total of $\bar{J}_a$ different positions. Moreover, in the application stage, I assume that each agent applies sequentially to firms according to the descending order of the expected utility that he/she derives.

Specifically, agent i initially applies to one firm, denoted as j_1, which provides the highest utility. Following this, the agent applies to another firm, denoted as j_2, offering the highest utility among the remaining firms, and continues this process sequentially until reaching the limit $\bar{J}_a$.[33] The application process concludes after agents have submitted all their applications.

After the application stage, neither the firms nor the researchers can ascertain the order in which agents applied to each firm. However, I can simulate a rank order list of every agent. Assuming that each agent's choice of j_a is unique with probability 1, the share of agents who apply to firm j can be expressed as:

$$s_{jt} = \sigma_j(\chi_t) = Pr_1(j_1 = j) + Pr_2(j_2 = j) + ... + Pr_{\bar{a}}(\bar{J}_a = j)$$

$s_t = (s_{1t}, ..., s_{Jt})$ is the vector that collects the number of agents who apply for firm j in market t empirically. $\sigma(\chi_t) = (\sigma_1(\chi_t), ..., \sigma_J(\chi_t))$ is a vector that collects the model predicted number of agents who apply for firm j in market t. $Pr_a(j_a = j)$ is the share of candidates that choose firm j in their ath application. For example, if $\bar{J}_a = 2$, then I can write the probability $Pr_1(j)$ and $Pr_2(j)$ more precisely:

$$Pr_1(j) = Pr\left(\underset{j\in\mathcal{J}}{\text{argmax}}\ v_{ijt}|\chi_{jt}\right), \quad Pr_2(j) = (1 - Pr_1(j)) \sum_{k\neq j} Pr\left(\underset{j\in\mathcal{J}/k}{\text{argmax}}\ v_{ijt}|\chi_{jt}\right)$$

While $Pr_1(j)$ is the same that is shown in the analysis before, $Pr_2(j)$ is the conditional probability that agent i sends the second application to firm j while sending the first application to a firm $k \neq j$.

Similarly, I can formulate the slot constraint function $\pi_j(x_t, \bar{q}_t, \xi_t)$ in this scenario. The feasibility of this approach stems from two key factors. First, each agent's preference order on applied firms can be simulated due to the application structure. Second, the large market approximation implies that there is a deterministic share of agents

[33] Notice here that a job seeker does not have an incentive to apply for a position that does not derives the highest expected utility in his/her ath application conditional on a set of equilibrium hiring standards because of the assumption that each applicant is infinitesimally small.

applying for a position is successfully hired. The uniqueness of the set of equilibrium hiring standards arises similarly from the Assumption 3.

Finally, the identification result follows closely from proposition 1 given the completeness assumption on instruments (Assumption 2). I provide an illustration of a simple random utility model that satisfies the connected substitutes assumption, wherein applicants are allowed to apply to two firms.

Example 2 *In the example, each agent can apply for two firms. The utility that an agent i's expected utility from applying for firm j equals $\delta_{jt} + u(q_i, x_{jt}) + \epsilon_{ijt}$. Here, q_i is i.i.d. drawn from a known distribution F_q while ϵ_{ijt} is i.i.d. drawn from EVT1.*

According to the above setup, I can write $Pr_1(j)$ and $Pr_2(j)$ as

$$Pr_1(j) = \frac{\exp(\delta_{jt} + u(q_i, x_{jt}))}{\sum_{j \in \mathcal{J}} \exp(\delta_{jt} + u(q_i, x_{jt}))} \qquad Pr_2(j) = \left(1 - \frac{\exp(\delta_{jt} + u(q_i, x_{jt}))}{\sum_{j \in \mathcal{J}} \delta_{jt} + u(q_i, x_{jt})}\right) \sum_{k \neq j} \frac{\exp(\delta_{jt} + u(q_i, x_{jt}))}{\sum_{j \in \mathcal{J}/k} \exp(\delta_{jt} + u(q_i, x_{jt}))}$$

The total number of applications received by firm j can be written as

$$s_j = \int_{\mathbb{R}} \int_{\mathbb{R}} \overbrace{\frac{\exp\left(U_{ij}(q_i, c_i)\right)}{1 + \sum_j \exp\left(U_{ij}(q_i, c_i)\right)}}^{j \text{ is the 1st choice}} dF_c(c_i) dF_q(q_i) + \int_{\mathbb{R}} \int_{\mathbb{R}} \underbrace{\left(1 - \frac{\exp\left(U_{ij}(q_i, c_i)\right)}{1 + \sum_j \exp\left(U_{ij}(q_i, c_i)\right)}\right)}_{j \text{ is not the 1st choice}} \sum_{k \neq j} \underbrace{\frac{\exp\left(U_{ij}(q_i, c_i)\right)}{1 + \sum_{j \in \mathcal{J}/k} \exp\left(U_{ij}(q_i, c_i)\right)}}_{j \text{ is the 2nd choice}} dF_c(c_i) dF_q(q_i) \quad \text{(A.1)}$$

Differentiating s_{jt} w.r.t. δ_{jt} for all j yields the Jacobian matrix. Three results can be demonstrated: 1, $\frac{\partial s_{jt}}{\partial \delta_{jt}} > 0$; 2, $\frac{\partial s_{jt}}{\partial \delta_{kt}} < 0 \ \forall k \neq j$; 3, the Jacobian matrix satisfies the Diagonal Dominance Condition through conventional arguments thereby implying its non-singularity. Consequently, the connected substitutes assumption is satisfied. This means that the set of $\{\delta_{jt}\}$ can be obtained by inverting the observed number of applications. Following this, one can rely on instruments to identify both firm and candidate preferences and types.

A natural upper limit of the number of applications each agent can submit is $J - 1$, or $\bar{J}_a \leq J - 1$. This is because if all agents send applications to all firms, then

the number of applicants is no longer informative. Formally, the upper limit $\bar{J}_a$ is restrained by the Assumption 1, because inversion result will become invalid as the number applications each candidate is allowed to submit increases.

A.3 Details of the NCSE Dataset

A.3.1 Interview process

During the interview process, each candidate will face many interviewers. The entire process has been recorded starting from NCSE 2016. In the interview process, interviewers will ask each candidate 4-5 questions, and the whole process typically lasts less than 30 minutes. Results of the interview will typically be announced 1-7 days after the interview. The questions asked during these interviews are typically not designed by the hiring department itself but are designed by another professional department within the government, although there is some flexibility.

In sum, due to the rigid process and the limited time, interviewers and interviewees do not have a chance to interact closely, making it hard for the interviewer to elicit a candidate's civic-mindedness during the process.

A.3.2 High Status and Powerful Positions

I classify positions from the following departments as high status and powerful positions: State Administration of Taxation, China Maritime Safety Administration, National Energy Administration, National Materials and Equipment Reserve, Administration of Customs, and Ministry of Natural Resources. Furthermore, I exclude positions tasked with publicity, data analysis, secretarial and assistant roles, statistical, and computer-related tasks from the list of high status and powerful positions.

A.3.3 Sample Construction

The initial sample adjustment involves the exclusion of positions from four municipalities—*Beijing*, *Shanghai*, *Tianjin*, and *Chongqing*. This equates to a reduction of

3,991 positions, or 13.9% of the original 28,648 positions.[34]

For the structural analysis, the assumption is that each agent assesses every position in the applicant market. In order to formulate a more homogeneous group for analysis, positions demanding unique skills are omitted. Agents applying for these positions are presumably disinclined to consider alternative positions due to their specialized skills. The discarded positions include policy-related railway positions, translator roles necessitating specific language skills, and positions requiring expertise in meteorology, sailing, computing, or medicine. In total, 6,225 positions demanding unique skills were dropped, which constituted 22% of the original 28,648 positions.

The third adjustment to the sample involved the exclusion of positions from *Hainan* province, *Xinjiang*, and *Tibet*. This resulted in a reduction of 2,082 positions, or 7.3% of the total positions.

After these sample adjustments, a refined sample consisting of 16,763 positions was obtained. Given the presence of 48 application markets in the sample, each application market, on average, contained 349 positions. For each application market t, non-competitive positions—those with a recruitment quota equal to or greater than the number of applicants—are assigned. Five application markets feature exclusively competitive positions. For these markets, a position with the lowest cutoff score for passing the written exam phase is normalized to be the outside option.

A.3.4 Summary Statistics of Position Requirements

I present summary statistics for position restrictions contained in the X_δ after I aggregate positions at the prefecture $\times$ department types $\times$ exam types.

Except for the degree requirement variable, all other variables can be understood as the share of positions within a prefecture $\times$ department types $\times$ exam types cell that have the specific requirement. For example, the average share of positions within

[34]Majority of positions affiliated with the central government headquarters are in Beijing. Positions tied to the central government but located outside Beijing are also omitted due to their negligible presence in the dataset (approximately 0.1%).

a prefecture × department types × exam types cell of having a minimum work experience requirement is 18.3%. Fresh college grad is a position requirement indicating that only the fresh college graduate of the year can apply. English Exam minimum score is a position requirement that only agents who pass the College English Test (CET-4 or CET-6) with a certain minimum score. Poor region is a position characteristic indicating that the position locating in a poor area. Placed based policy is a position characteristic indicating that the position enjoys a place based policy.

Table A.1: Summary Statistics of the Sample Used for Estimation (X_δ)

Variable Name	Mean	Std. Dev.
Education: degree requirement	0.560	0.546
Minimum Work Experience	0.180	0.285
Political Status	0.075	0.221
Fresh College Graduate	0.354	0.369
English Exam Minimum Score	0.368	0.456
College major: humanity	0.172	0.263
College major: Social Science	0.633	0.368
College major: Science	0.401	0.367
College major: Finance	0.506	0.380
Minimum serving requirement	0.008	0.270
Poor region	0.107	0.251
Placed based policy	0.169	0.315
Observations	1,853	

The degree requirement is coded differently. Before the data aggregation at the prefecture × department types × exam types level, it equals one if the position has no minimal degree requirement. It equals 2 if it requires at least a bachelor's degree. It equals 3 if it requires at least a master's degree. Therefore, the degree requirement

after the data aggregation is a composite measure of the degree requirement within a prefecture × department types × exam types cell. A higher degree requirement measure implies a higher average degree requirement for positions within a prefecture × department types × exam types cell.

A.4 Correlation Between a_i and c_i

In the empirical model, candidate ability (a_i) and civic-mindedness (c_i) are assumed to be independently distributed. I present survey evidence supporting this assumption. The survey sample is drawn from Feng, He and Liu (2023), which surveys college students from Liaoning and Guangdong provinces in China.

To measure civic-mindedness, I construct two measures. First, the 'Dictator game' from the Feng, He and Liu (2023) survey assesses respondents' pro-social preference. This game involves respondents dividing a windfall amount of money between themselves and a trusted charity organization. The scenario is presented twice with a different amount each time. From this, I derive a 'willingness to donate' variable, calculated as the average proportion of money respondents choose to donate. Second, I use behavioral indicators, aggregating responses to whether participants engaged in various activities over the past three years. These activities include blood donation, disaster- or poverty-relief donations, online advocacy, discussion, or message forwarding for social issues, and face-to-face participation in social causes.

The primary dependent variable proxying each respondent's ability is their self-reported status as a student in an elite Chinese university, i.e., those universities as part of Project '211' or Project '985'.

Table A.1 presents the estimated correlation between self-reported civic-mindedness and ability, controlling for respondent demographics like age and gender. I find no significant correlation between measured civic-mindedness and ability levels, even when controlling for self-reported intentions to be employed by the government. This survey-based empirical evidence supports the assumption in the empirical model that

Table A.1: Correlation Between Ability and Civic-mindedness: Survey Evidence

	(1)	(2)	(3)	(4)
	Willingness to Donate		Past Prosocial Actions	
Elite University	-0.00894	-0.00970	0.0470	0.0573
	(0.0269)	(0.0269)	(0.117)	(0.116)
Interest in Government Jobs	No	Yes	No	Yes
Age	Yes	Yes	Yes	Yes
Gender	Yes	Yes	Yes	Yes
Obs	419	419	419	419

Notes: willingness to donate is a self-reported willingness to donate. Past prosocial actions include blood donation, advocating for the society online/offline, volunteer experience. Elite university is a variable that takes value 1 if the survey respondent self reported as a student from an elite university (a member of project 211 or project 985) in China.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

candidate ability (a_i) and civic-mindedness (c_i) are independently distributed

Online Appendices: Not for Publication

O.1 Omitted Proofs

O.1.1 Proof for proposition 2

Proof. Pick an arbitrary j,

$$\pi_j = \int\int \Phi(a_i - \bar{q}_j + \mu_\eta \mathbb{1}_{\text{Prov. Exam}}) \left[\frac{\exp\left(U_{ij}(a_i, c_i, \bar{q}_j)\right)}{1 + \sum_j \exp\left(U_{ij}(a_i, c_i, \bar{q}_j)\right)} \right] dF_c dF_a$$

Then

$$\frac{\partial \pi_j}{\partial \bar{q}_j} = \int\int -\Phi'(a_i - \bar{q}_j + \mu_\eta \mathbb{1}_{\text{Prov. Exam}}) \left[p \frac{\exp\left(U_{ij}(a_i, c_i, \bar{q}_j)\right)}{1 + \sum_j \exp\left(U_{ij}(a_i, c_i, \bar{q}_j)\right)} \right]$$
$$+ \Phi(a_i - \bar{q}_j + \mu_\eta \mathbb{1}_{\text{Prov. Exam}}) \left[p \frac{\partial \frac{\exp\left(U_{ij}(a_i, c_i, \bar{q}_j)\right)}{1 + \sum_j \exp\left(U_{ij}(a_i, c_i, \bar{q}_j)\right)}}{\partial \bar{q}_j} \right] dF_c dF_c$$

The first term is negative. I only need to show that $\frac{\partial \frac{\exp\left(U_{ij}(a_i, c_i, \bar{q}_j)\right)}{1 + \sum_j \exp\left(U_{ij}(a_i, c_i, \bar{q}_j)\right)}}{\partial \bar{q}_j} < 0.$

$$\frac{\partial \frac{\exp\left(U_{ij}(a_i, c_i, \bar{q}_j)\right)}{1 + \sum_j \exp\left(U_{ij}(a_i, c_i, \bar{q}_j)\right)}}{\partial \bar{q}_j} =$$
$$\frac{-\exp\left(U_{ij}(a_i, c_i, \bar{q}_j)\right) \Phi'(a_i - \bar{q}_j + \mu_\eta \mathbb{1}_{\text{Prov. Exam}}) \left(1 + \sum_j \exp\left(U_{ij}(a_i, c_i, \bar{q}_j)\right)\right) + \exp\left(U_{ij}(a_i, c_i, \bar{q}_j)\right)^2 \Phi'(a_i - \bar{q}_j + \mu_\eta \mathbb{1}_{\text{Prov. Exam}})}{\left(1 + \sum_j \exp\left(U_{ij}(a_i, c_i, \bar{q}_j)\right)\right)^2}$$
$$= \frac{-\exp\left(U_{ij}(a_i, c_i, \bar{q}_j)\right) \Phi'(a_i - \bar{q}_j + \mu_\eta \mathbb{1}_{\text{Prov. Exam}}) \left(1 + \sum_{k \neq j} \exp\left(U_{ik}(a_i, c_i, \bar{q}_k)\right)\right)}{\left(1 + \sum_j \exp\left(U_{ij}(a_i, c_i, \bar{q}_j)\right)\right)^2} < 0$$

Now I've shown that $\frac{\partial \pi_j}{\partial \bar{q}_j} < 0$ for all j. The case for $\frac{\partial \pi_j}{\partial \bar{q}_{j'}} > 0$ is analogous.

$$\frac{\partial \pi_j}{\partial \bar{q}_k} = \int\int \Phi(a_i - \bar{q}_j + \mu_\eta \mathbb{1}_{\text{Prov. Exam}}) \left[\frac{\partial \frac{\exp\left(U_{ij}(a_i, c_i, \bar{q}_j)\right)}{1 + \sum_j \exp\left(U_{ij}(a_i, c_i, \bar{q}_j)\right)}}{\partial \bar{q}_k} \right] dF_c dF_a$$

Notice the term in the bracket is negative.

$$\frac{\partial \frac{\exp\left(U_{ij}(a_i, c_i, \bar{q}_j)\right)}{1+\sum_j \exp\left(U_{ij}(a_i, c_i, \bar{q}_j)\right)}}{\partial \bar{q}_k} = \frac{-\exp\left(U_{ij}(a_i, c_i, \bar{q}_j)\right) \Phi'(a_i - \bar{q}_k + \mu_\eta \mathbb{1}_{\text{Prov. Exam}}) \exp\left(U_{ik}(a_i, c_i, \bar{q}_k)\right)}{\left(1+\sum_j \exp\left(U_{ij}(a_i, c_i, \bar{q}_j)\right)\right)^2} < 0$$

■

O.2 NCSE: Additional Empirical Patterns

O.2.1 Hiring Standard: Cross-year Comparison

I demonstrate that the equilibrium passing scores for the same position, despite having openings in different years, exhibit a considerable degree of similarity. This observation substantiates the assumption in the matching game model that candidates have rational expectations regarding the equilibrium passing score of each position, as they can infer the passing score for a specific position by leveraging information from past exams.

Supplemental position comparison data is collected from the same company that offers exam training to NCSE takers. In order to assist exam takers in making informed application decisions, the company's website provides historical cutoff scores required for passing the written exam phase for the same position from past exams (if available) for direct comparison. Applicants can conduct a search for positions they are interested in and subsequently access the cutoff scores from the preceding three years of exams. I web scraped the data from the company's website and have accordingly compiled a set of identical positions. Altogether, there are 1657 positions that had openings for the 2017 and 2015 NCSE, as well as 2854 positions that had openings for the 2017 and 2015 NCSE. To account for the differences in exam papers, I have standardized the scores for each year.

In Figure O.2.1, the left (right) graph depicts the distributions of equilibrium passing scores for positions that had openings in both 2017 and 2016 (2015). The distributions of scores exhibit close resemblance to each other. T-tests suggest that

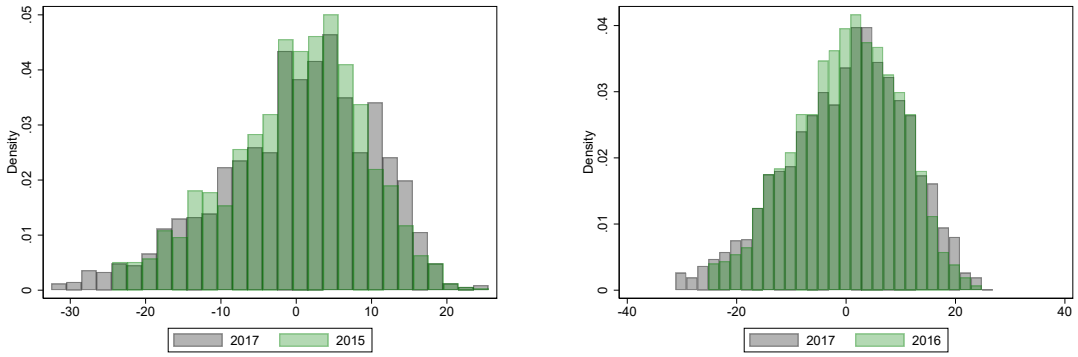


Figure O.2.1: Equilibrium Passing Score ($\bar{q}_t$) of a Same Positions in Different Years

Notes: In the left (right) panel, I plot the histogram of equilibrium passing score ($\bar{q}_t$) of a same positions in 2017 and 2015 (2016). I drop positions that require special skills. T-tests suggest that scores are not statistically significantly different for any of the two comparison groups.

the mean equilibrium scores across any two pairs of exams do not display a statistically significant difference.

O.2.2 Empirical Evidence of Sorting: Province-level Evidence

I present empirical evidence of sorting at the provincial level. I show that provinces with a higher GDP per capita tend to have higher average equilibrium passing scores. In line with the primary structural estimation, I exclude positions that are in four municipalities and those that require special skills. Furthermore, positions from *Hainan*, *Xinjiang*, and *Tibet* are also excluded.

In the left panel of Figure O.2.2, I plot the average equilibrium hiring standards for each province against its corresponding provincial GDP per capita. The figure shows a clear upward trend, indicating that provinces with higher GDP per capita generally have higher equilibrium hiring standards on average. These findings corroborate the empirical specification that provinces with a higher GDP per capita have a candidate pool with a higher average ability.

One explanation is that provinces with higher GDP per capita attract more applicants than other provinces. Therefore, cutoff scores in those provinces are naturally higher than that in other provinces. In the right panel of Figure O.2.2, I show that the correlation between provincial average applicants to recruitment quota ratio (App-

to-recruit ratio) and provincial GDP per capita is much less pronounced than the correlation between average cutoff scores and the provincial GDP per capita. This evidence further supports that provinces with a higher GDP per capita have a candidate pool with a higher average ability.

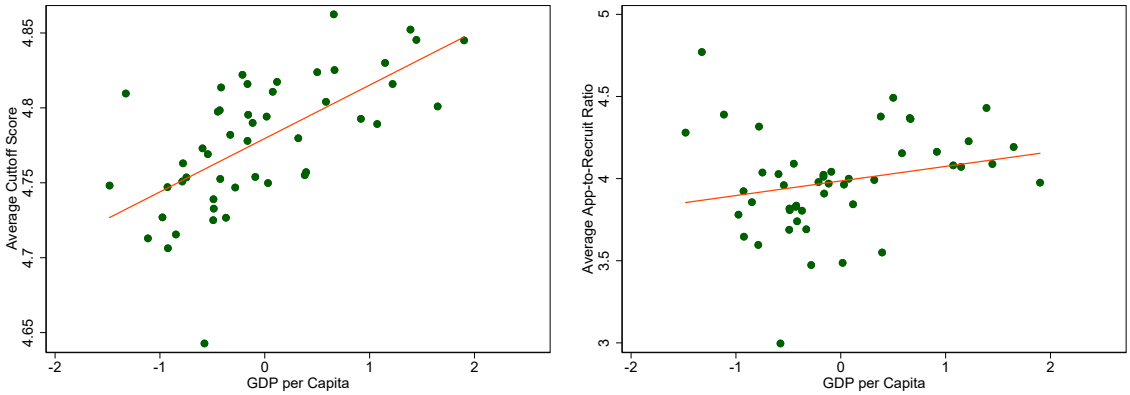


Figure O.2.2: Empirical Sorting: Province GDP per Capita

Notes: In the left panel, the Y-axis represents the logged province average cutoff score (equilibrium passing score $\bar{q}_t$) in 2016 and 2017 NCSE. In the right panel, The Y-axis represents the logged province average App-to-recruit Ratio in 2016 and 2017 NCSE. X-axis represents the province GDP per capita. I exclude positions that are in municipalities and require special skills. I drop positions from *Hainan*, *Xinjiang* and *Tibet*.

Several other plausible explanations exist as to why provinces with higher GDP per capita possess a candidate pool with superior average ability levels. First, provinces with higher GDP per capita tend to host more prestigious universities, which draw talents from across the country. Secondly, regions with a higher GDP per capita are likely to offer better amenities, attracting candidates from other provinces to apply.

O.3 Estimation: Details and Additional Results

O.3.1 Additional Details of the Estimation

Below I outline the algorithm that I use in the estimation.

1. An initial set of non-linear parameters ($\hat{\Theta}$) is guessed in the outer loop.

2. The application equation (3.2) is inverted for each application market (t) to get $\hat{\delta}_j$, which is a function of Θ.

3. According to $\hat{\delta}_j = \theta_\delta X_\delta + \xi_j$ the parameter$\hat{\theta}_\delta$is estimated, and the residual $\hat{\xi}_j\left(\hat{\Theta}\right)$is obtained.

4. Instruments (Z_j) are interacted with the residual $\hat{\xi}_j\left(\hat{\Theta}\right)$ to form moments, denoted as $\Omega(\hat{\Theta}) = \frac{1}{J}\sum_j \left[Z_j \times \hat{\xi}_j\left(\hat{\Theta}\right)\right]$.

5. In the outer loop, the GMM objective function is minimized,

$$\Theta^* = \text{argmin}_\Theta \left[\Omega(\hat{\Theta})\right] W \left[\Omega(\hat{\Theta})\right]'$$

   subject to the slot constraint:

$$\frac{1}{J}\sum_j \left[\int_{\mathbb{R}} \Phi(a_i - \bar{q}_j + \mu_\eta \mathbb{1}_{\text{Prov. Exam}}) \left[\int_{\mathbb{R}} \frac{\exp\left(U_{ij}(a_i, c_i, \bar{q}_j)\right)}{1 + \sum_j \exp\left(U_{ij}(a_i, c_i, \bar{q}_j)\right)} dF_c(c_i)\right] dF(a_i) - \lambda_j\right] = 0$$

6. In the first step, the weighting matrix W is set as $Z'Z$. In the second step, I re-estimate the model using the optimal weighting matrix.

7. Given the set of estimates Θ^*, I estimate $\hat{\delta}_j^*$ for each position j and obtain estimates $\hat{\theta}_\delta^*$ by estimating linear regression $\hat{\delta}_j^* = \theta_\delta X_\delta + \xi_j$.

A key step of the estimation algorithm is inverting the number of applicants to obtain the δ_j for each position j using a contraction mapping. As recommended by Conlon and Gortmaker (2020), I use the SQUAREM algorithm of Varadhan and Roland (2008) to solve the contraction mapping. I use sparse grids generated by the method proposed in Heiss and Winschel (2008) to ensure that the numerical integration is accurate.[35] Further, I use 10 randomly drawn starting values to help confirm that the estimation procedure identifies a global minimum of the objective function.

[35] I download the code from the website http://www.sparse-grids.de/ to get sparse grids.

O.3.2 Linear Parameter Estimates (θ_δ)

In this subsection, I present linear parameter estimates obtained in the structural estimation. I regress the inverted set of $\{\delta_j\}$ on a set of position requirements X_δ to obtain the set of linear parameters θ_δ

Table O.3.1: Linear Parameter Estimates θ_δ

Linear Para.	Estimates (Standard Errors)			
	Education	Work exp	Political status	Grad Only
Estimate	-0.227	-0.765	-0.513	-0.540
Std. Error	(0.525)	(0.226)	(0.157)	(0.343)
	English Min	Humanity	Social science	Science
Estimate	-0.349	0.603	0.400	0.464
Std. Error	(0.393)	(0.211)	(0.492)	(0.365)
	Finance	Min 5 year	Poor	Place policy
Estimate	-0.324	0.052	-0.218	-0.495
Std. Error	(0.425)	(189)	(0.183)	(0.140)
	Constant			
Estimate	1.628			
Std. Error	(0.671)			

I find that requirements such as a higher degree, longer work experience, attaining political status, being a fresh graduate of that year, passing an English exam, having a major in finance and willingness to work in relatively undeveloped areas or areas with special treatments are associated with a higher cost for candidates. Contrarily, positions requiring a humanities, social science, or science major do not seem to discourage candidates from applying, indicating that a significant number of candidates possess one of these majors. The set of parameter estimates are intuitive since all position requirements other than the college major requirements appear to be costly to candidates. One exception is the commitment to the position for more than five years, which is in-precisely estimated.

O.3.3 Parameter Estimates: Alternative Specification

In this subsection, I present parameter estimates with alternative specifications. Specifically, I classify positions from the State Administration of Taxation as high status and powerful positions. Furthermore, I exclude positions tasked with secretarial and assistant roles, statistical, and computer-related tasks from the list of high status and powerful positions.

Table O.3.2: Non-linear Parameter Estimates: Other Results

Non-linear Para	Estimates and Standard Errors		
	$\bar{\theta}_\mu$ (Ability Dist Cons)	θ_μ (Ability Dist Slope)	μ_c (Civic Dist Mean)
Estimate	-3.055	2.409	1.769
Std. Error	(0.001)	(0.000)	(0.001)
	σ_C (Civic Dist Var)	η_{prov} (Exam Mean)	θ_x (Power)
Estimate	0.465	-0.938	1.422
Std. Error	(0.000)	(0.000)	(0.000)
	θ_x (Level)	θ_x (Wage)	
Estimate	1.417	1.377	
Std. Error	(0.001)	(0.000)	

Notes: Please refer to Table O.3.3 for linear parameter estimates θ_δ.

In Table O.3.2 and O.3.3, I show non-linear and linear estimates using the alternative definition of powerful and high status positions. Estimates reported in Table O.3.2 (for non-linear parameters Θ) and O.3.3(for linear parameters θ_δ) closely resemble those reported in Table 2 (for non-linear parameters Θ) and Table O.3.1 (for linear parameters θ_δ) based on the preferred specification.

O.4 Consequence on Gov Performance: Position Weighting

Here, I present more simulation results following the discussions at the end of Section 3.5. Specifically, I show simulation results when employing different weighting functions.

In the left panel of Figure O.4.1, I use the prefecture average non-private sector

Table O.3.3: Linear Parameter Estimates θ_δ: Other Results

Linear Para.	Estimates (Standard Errors)			
	Education	Work exp	Political status	Grad Only
Estimate	-0.141	-0.772	-0.708	-0.128
Std. Error	(0.336)	(0.145)	(0.100)	(0.219)
	English Min	Humanity	Social science	Science
Estimate	-0.594	0.702	0.311	0.258
Std. Error	(0.251)	(0.135)	(0.315)	(0.234)
	Finance	Min 5 year	Poor	Place policy
Estimate	-0.296	0.053	0.432	-1.056
Std. Error	(0.272)	(0.121)	(0.117)	(0.154)
	Constant			
Estimate	1.657			
Std. Error	(0.430)			

wage to construct the weighting function W_j in the aggregation. That is to say, positions having a higher average non-private sector wage is more important for the government performance when staffed with candidates possessing higher ability. The simulation exercise indicates that the cutoff value of ϑ_c when the NCSE outperforms the alternative mechanism is different than that is shown in Figure 4. Similarly, I use both weights on ability and civic-mindedness when plotting the right panel of Figure O.4.1. The simulation exercise indicates that the cutoff value of ϑ_c when the NCSE outperforms the alternative mechanism is smaller than that is shown in the left Figure.

O.5 Additional Details of the NCSE Data Cleaning

NCSE is managed by the central government, which fills positions for three types of government organizations: central party organs, central state administrations (along with their local branches), and public administrations governed by civil servant law (along with their local branches). NCSE positions have many categorical characteristics. For example, more than 20 different departments are hiring during 2016 and

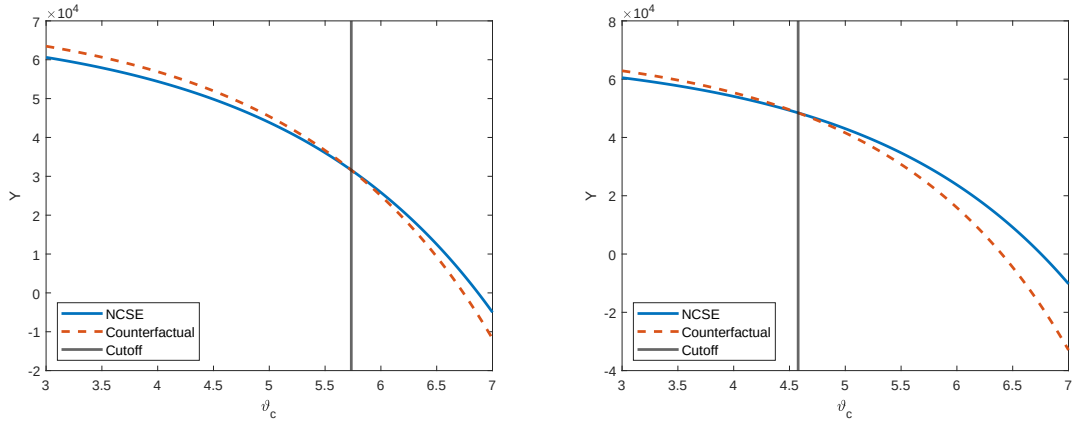


Figure O.4.1: Government Performance: Alternative Weighting

Notes: The Y-axis represents the government output Y. The solid line plots the government output given NCSE while the dashed line plots the government output given the alternative mechanism. I fix $\vartheta_a = 1.5$ in the simulation to focus on how the change of ϑ_c affect Y. The x-axis represents the importance of civic-mindedness in government production ϑ_c. A higher ϑ_c implies a more important impact of candidate civic-mindedness relative to the ability. In the left panel, positions with higher non-private unit wages are assigned a larger weight (W_j^a) in the simulation. In the right panel, positions with higher wages and higher potential rent-seeking opportunities are assigned larger weights (W_j^a and W_j^c) in the simulation.

2017. Moreover, because each position may require different combinations of majors in school, the set of majors is also large. Therefore, I make some categorization changes to reduce the dimension of some of position characteristics by combining.

O.5.1 Department

Numerous government organizations engage in recruitment via the NCSE. However, not every department conducts hiring annually. To prevent certain departments from exhibiting insufficient observations, I categorize these into nine distinct types, reflective of functional similarities. These categories encompass audit, tax, transportation, resource, custom, prison, postal, finance, and statistics departments. The specific categorizations are explained in table O.5.1

O.5.2 Education

The NCSE enforces two distinct requirements on educational background: the degree requirement and the major requirement.

Department Types	**List of Departments**
Audit	State Auditing Administration
Tax	State Administration of Taxation
Transportation	Ministry of Transport, Civil Aviation Administration of China, China State Railway Group Co., Ltd., Pearl River Administration of Navigational Affairs, Yangtze River Administration of Navigational Affairs, China Maritime Safety Administration
Resource	Ministry of Water Resources, National Bureau of Surveying and Mapping (now in Ministry of Natural Resources), China Earthquake Administration, China Meteorological Administration, State Oceanic Administration, National Energy Administration, State Administration of Work Safety, Administration of National Materials and Equipment Reserve, Ministry of Natural Resources
Custom	Entry-Exit Inspection and Quarantine Bureau (now in Administration of Customs), Administration of Customs
Prison	Qincheng Prison, Yancheng Prison
Postal	State Post Bureau
Finance	China Banking Regulatory Commission, China Insurance Regulatory Commission, Securities Regulatory Commission
Statistics	National Bureau of Statistics of China

Table O.5.1: Departments by Type

Almost every position necessitates a minimum degree requirement. For instance, applicants are expected to possess at least a bachelor's degree. I exclude positions necessitating a PhD degree due to the limited number.

Contrary to the degree requirement, the major requirement varies significantly across positions. Certain positions require a very specific college major; for instance, some positions within the China National Meteorological Administration mandate a major in meteorology. I exclude positions that necessitate specialized skills, specifically positions related to railway policy, translators requiring particular language proficiency, and positions demanding expertise in meteorology, sailing, computer science, or medicine. The remaining majors are grouped into four broad categories: humanities, social sciences, STEM, and finance, with each category comprising several distinct majors.

O.5.3 Political Status

There are three types of political status requirements in NCSE 2016 and 2017. The majority of positions do not impose any requirement concerning political status. The second type necessitates applicants to be CCP members. The third type requires applicants to either be CCP members or members of the Chinese Communist Youth League. I combine the second and third types to circumvent the small sample issue.

O.5.4 Work experience

Work experience in this context refers to the minimum number of years of Grassroots Work experience (*ji ceng gong zuo jing li*). A majority of positions do not impose any work experience prerequisite. A minority of positions require applicants to possess a minimum of two years of work experience. A few positions necessitate at least one, two, or five years of work experience. To optimize data utilization, I generate a dummy variable representing the minimum work experience requirement. The dummy variable equals 1 (0) if the position possesses (does not possess) a minimum work experience requirement.

O.5.5 Information Regarding Compensation

The NCSE database lacks wage information for each posted position. Legal compensation for government officials within China's government lacks transparency. The legal compensation of China's government workers consists of four parts: base salary, stipends, subsidies and bonuses. The base salary, consisting of headship pay and rank pay, is transparent following the announcement by the State Council since 2006 https://www.gov.cn/zhengce/content/2019-03/19/content_5375052.htm. The basic salary is the same for everyone given the same position and level. However, departments in different places have their own stipends, subsidies and bonuses plan, making the overall compensation non-transparent.

There is no reliable source offering a comprehensive description, let alone potential illegal returns (grey income and other rent-seeking opportunities) associated with

government positions. Deng et al. (2023) provide estimates regarding the unofficial incomes of officials in China. Feng, He and Liu (2023) examines the consequence of this non-transparency on the self-selection into the government recruitment using an information experiment.

I argue that the absence of compensation data does not significantly impact the analysis of this paper for three primary reasons. First, government workers' legal compensation is determined by governmental regulations, implying that it cannot adjust flexibly to the forces of demand and supply as witnessed in competitive labor markets. As a result, there's no need for concerns about wages being set endogenously in response to changes in labor supply. Second, given the absence of compensation details in job postings, candidates do not have access to position-specific wage information during the application process. Lastly, the patterns in position choice inherently reflect job seekers' holistic preferences for each available position.

I use a measure called average urban non-private units. Urban non-private units refer to all non-private legal entities in urban areas. These specifically include state-owned economies, urban collective economies, joint-venture economies, stock economies, economies invested by foreign, Hong Kong, Macao, and Taiwan investors, other economic entities, and their affiliated institutions. The employees of urban non-private units do not include those employed in township enterprises, private enterprises, individual urban workers, retired or resigned personnel, as well as others who are not counted within the scope of the workforce as per relevant regulations.

O.6 NCSE 2016: Official Guide of the Written Exam

In this section, I translate the official guide of the 2016 NCSE for the written exam. I only include the Public Subject Written Examination because it is relevant for the analysis.

The content for the Administrative Professional Test (ability test) is uniform across all position types. However, the complexity of the examination questions in

the ability test is greater at the provincial level or above, compared to the prefecture level or below. Specifically, the 2016 examination paper for the provincial level or above contains 135 questions, while that for the prefecture level or below consists of 130 questions. The province or above exam has 15 harder questions compared to the prefecture exam.

The question for the policy essay varies between the provincial exam and the prefecture exam, with the former's question being more challenging. Nonetheless, the primary set of skills each examination paper aims to assess remains largely analogous. The only distinction lies in that the provincial examination paper additionally evaluates analytical and problem posing skills, whereas the prefectural examination paper gauges implementation skills instead.

In the subsequent sections, I provide the fully translated examination syllabus from the NCSE 2016 for reference.

O.6.1 Content of the Public Subject Written Examination

The public subject written examination for civil servants recruitment in 2016 by central authorities and their directly affiliated institutions comprises two parts: the Administrative Professional Ability Test and the Policy Essay. Both parts adopt a closed-book examination format.

The Administrative Professional Ability Test is an objective test with a time limit of 120 minutes and a full score of 100. The Policy Essay is a subjective test with a time limit of 180 minutes and a full score of 100.

O.6.2 Answering Requirements

Administrative Professional Ability Test. Applicants must bring examination stationery, including pens or signing pens with black ink, 2B pencils, and erasers. Applicants must fill in the admission ticket number at the designated place with a 2B pencil and answer on the answer card. Answers made elsewhere are invalid.

Policy Essay. Applicants must bring examination stationery, including pens or

signing pens with black ink, 2B pencils, and erasers. Applicants must fill in the admission ticket number at the designated place with a 2B pencil, and answer on the answer card at the designated place with a pen or signing pen. Answers made elsewhere or made in pencil are invalid.

O.6.3 Introduction to the Administrative Professional Ability Test

Test Content. The Administrative Professional Ability Test mainly tests basic qualities and abilities closely related to civil service, suitable for objective written examination, including verbal comprehension and expression, quantitative reasoning, judgment reasoning, data analysis, and common sense judgment.

Verbal comprehension and expression mainly test applicants' ability to think and communicate in language, quickly and accurately understand and grasp the connotation of text materials, and find main information and important details based on materials. Common types of questions include reading comprehension, logical fill-in-the-blank, and sentence expression.

Quantitative reasoning mainly tests applicants' understanding of the quantitative relationship between things and their ability to solve quantitative problems, which mainly involves the analysis, reasoning, judgment, and operation of data relationships. Common types of questions include numerical reasoning and mathematical operations.

Judgment reasoning mainly tests applicants' ability to analyze and reason about various relationships, involving understanding, comparison, combination, deduction, and induction of figures, language concepts, object relationships, and text materials. Common types of questions include graphical reasoning, definition judgment, analogy reasoning, and logical judgment.

Data analysis mainly tests applicants' ability to comprehensively understand and analyze various forms of texts, charts, and other materials. This part usually consists of statistical charts, numbers, and text materials.

Common sense judgment mainly tests applicants' basic knowledge and ability

to use this knowledge for analysis and judgment, focusing on the understanding of national conditions and social conditions, basic qualities of comprehensive management, and involves politics, economics, law, history, culture, geography, environment, nature, science and technology, etc.

Introduction to Question Types. The Administrative Professional Ability Test involves various types of questions, which will be combined according to the purpose of the examination, the situation of the applicants, in terms of types, quantity, and difficulty. The following are introductions to some commonly used question types: Verbal comprehension and expression, Quantitative reasoning, Judgment reasoning, Data analysis and Common sense judgment. I omit example questions here.

O.6.4 Introduction to Policy Essay

Policy Essay is an examination subject testing the basic abilities required for office work. The Policy Essay exam paper consists of precautions, given materials, and answering requirements. Policy Essay examination is set up with two types of exam papers according to different requirements for provincial level and above (including sub-provincial level) management positions, prefecture (place) and below management positions, and administrative law enforcement positions.

Provincial level and above (including sub-provincial level) management positions' Policy Essay examination mainly tests applicants' reading comprehension ability, comprehensive analysis ability, problem posing and solving ability, and textual expression ability.

Reading comprehension ability - requires a comprehensive grasp of the content of the given materials, accurately understanding the meaning of the given materials, accurately refining the viewpoints contained in the facts, and revealing the essential problems reflected.

Comprehensive analysis ability - requires analysis and induction of the entire or part of the content, viewpoints, or problems of the given materials, multi-angle thinking about the content of the materials, and making reasonable inferences or evalua-

tions.

Problem posing and solving ability - requires using one's own practical experience or life experience, on the basis of understanding and analyzing the given materials, to identify and define problems, make assessments or trade-offs, and propose solutions or measures.

Textual expression ability - requires proficient use of the designated language, using explanation, statement, argumentation, etc., to accurately, appropriately, concisely, and fluently express thoughts and viewpoints.

prefecture (place) and below management positions and administrative law enforcement positions' Policy Essay examination mainly tests applicants' reading comprehension ability, implementation ability, problem-solving ability, and textual expression ability.

Reading comprehension ability - requires understanding the main content of the given materials, grasping the relationship between different parts of the given materials, and making appropriate explanations for the viewpoints and facts involved in the given materials.

Implementation ability - requires accurately understanding work goals and organizational intentions, adhering to the principle of lawful administration, and timely and effectively completing tasks according to objective actual conditions.

Problem-solving ability - requires using one's own knowledge and experience to make correct analysis and judgment on specific problems, and propose practical measures or methods.

Textual expression ability - requires proficient use of the designated language, and accurate and reasonable explanation, statement, or interpretation of events and viewpoints.