Feigning Fairness: The Consequences of a Male-Favoring Gender Quota

Jiawei Lyu *
University of Pittsburgh
Click here for the latest version.

September 7, 2025

Abstract

Hiring quotas can reshape who enters the workforce and how organizations perform. I study a gender quota that effectively preserved the representation of a historically dominant group: men in China's civil service. After merit-based hiring boosted women's share of new recruits to about 67%, some county tax bureaus introduced one-to-one gender quotas mandating equal hires by sex. I build a new dataset by web-scraping and digitizing position-level civil service exam records and county tax revenues, which I then link together. Using this dataset, I show that counties with more recent female hires were more likely to adopt quotas, especially in tax-collection roles. Exploiting variation in the timing and intensity of adoption across counties, I use a matched staggered difference-in-differences and event-study framework to show that quotas decrease female representation, lower candidate quality, and reduce worker productivity: tax revenue falls by 4.7%, or about $4 million per county per year. Evidence suggests these effects stem from both reduced worker quality and institutionalized gender bias, offering new insights into debates over reverse affirmative action when historically dominant groups become outnumbered.

Keywords: Civil Service, Gender Quota, Bureaucratic Performance, Tax Collection

*Department of Economics, University of Pittsburgh, 4901 Wesley W. Posvar Hall, 230 South Bouquet Street Pittsburgh, PA 15260. Email: jiaweilyu@pitt.edu. I am deeply grateful to Gabriel Tourek, Lester Lusher, Lise Vesterlund, Claire Duquennois, and Osea Giuntella for their invaluable guidance and support throughout this project. I also thank Mark Anderson, Shan Aman-Rana, Zoe Cullen, Alex Eble, Anne Fitzpatrick, Emily Nix, Claudia Olivetti, Sandra Rozo, Tianyi Wang, and participants at seminars at the University of Pittsburgh, University of Toronto, and Wilfrid Laurier University, as well as audiences at the 2024 Southern Economic Association Annual Conference, 2025 ASSA Annual Meeting, Pacific Conference for Development Economics 2025, Midwest International Economic Development Conference 2025, Canadian Economic Association 2025, WEAI Conference 2025, Society of Labor Economists Conference (SOLE) 2025 for their helpful comments and suggestions. All errors are my own.

1 Introduction

Policies to redress historical discrimination have expanded opportunities for disadvantaged groups.[1] But in some fields—such as education, nursing, and parts of the public sector—women now outnumber men, raising concerns about reverse discrimination (Cohen, 2024; Pollitt, 2023; Dominus, 2023; Nietzel, 2024). Rising in prominence worldwide are quotas that preserve men's representation, often justified by appeals to fairness or efficiency.[2] This paper asks two questions: what motivates adoption of such quotas, and how do they affect representation, candidate quality, and institutional productivity?

The effects of male-favoring quotas are ambiguous. Quotas may reduce backlash against women's rising representation and promote gender balance (Murray, 2014), but they also risk displacing more qualified women with less qualified men, lowering overall candidate quality. Voluntary adoption adds further complexity: organizations may adopt quotas because they perceive benefits, but adoption can also serve to protect incumbents' advantages, creating unintended consequences for equity and performance.

I provide the first empirical evidence on a gender quota that protects men's representation in China's civil service, examining how a "reverse affirmative action" policy reshapes representation and its implication for efficiency. I study gender quota adoption in China's county tax bureaus—the largest recruiters of civil servants and the backbone of government revenue collection. Because tax revenues provide a direct and quantifiable measure of productivity, this setting offers a unique opportunity to assess how quotas shape both representation and state capacity.[3] Historically, men dominated the county tax bureaus and remain overrepresented at senior levels, but since the 2011 introduction of the National Civil Service Examination, women have comprised about two-thirds of new recruits (See Figure A.2).

[1] Examples include affirmative action in universities and workplaces (Holzer and Neumark, 2000), gender quotas in legislatures (Krook, 2010), reservations in public sector jobs (Pande, 2003), and quotas in professions such as teaching, policing, and civil service (Chattopadhyay and Duflo, 2004). Together, these policies show how redistributive hiring and representation rules expand opportunities for disadvantaged groups.

[2] In Japan, medical schools historically limited women's admissions (Schieder, 2019); in South Korea, the public sector has set minimum male quotas (Sang-Hun, 2010); and in the United States, some colleges are considering affirmative action for boys (Cohen, 2024).

[3] County tax bureaus are administrative units below the city (prefecture) level, responsible for tax collection within their jurisdictions. In 2023 county tax bureaus hired 24,975 of 37,100 new civil servants (67.3%).

Some bureaus then began introducing one-to-one gender quotas in 2016, dividing vacancies in positions subject to the quota evenly between men and women. By 2023, roughly 29% of bureaus had implemented the policy in at least one hiring cycle. Using novel data on civil service examinations and county tax revenues, I apply a matched staggered difference-in-differences and event-study framework and find that quotas decrease female representation, lower candidate quality, and reduce productivity: tax revenue falls by 4.7%, or about $4 million per county annually.

To study these questions, I construct a novel dataset on civil service recruitment in China, covering all 91,149 postings from 2011 to 2023 scraped from the National Civil Service Administration examination portal.[4] For tax bureaus, I additionally collected the number of applicants and written exam cut-off scores from 2016 onward, and digitized interviewee lists and final admission announcements to extract names, gender, and educational background. Linking position, applicant, and admission data provides a nearly complete dataset to analyze how quotas affect newly hired civil servants. Building on prior work showing the critical role of tax personnel in shaping revenue outcomes across diverse contexts (Basri et al., 2021; Bergeron et al., 2022, 2024; Khan et al., 2019; Okunogbe and Tourek, 2024; Besley and Persson, 2014; Mitchell, 1998), I evaluate performance using county-level tax revenue reported in the China City Statistical Yearbooks for a subset of counties.[5] Tax revenue is the primary outcome by which tax bureaus are assessed, the main source of government funding, and a convenient monetary metric for testing the impact of quotas. As complementary indicators of productivity, I use administrative approvals and penalties from county tax bureau annual reports.

My findings have two parts. The first part investigates who adopt the quota. At the county level, adoption is more likely in counties with higher GDP, larger populations, a greater share of local residents, and lower urbanization rates. Hiring histories also matter: counties with a higher share of female new entrants prior to the policy are more likely to

[4]Each posting reports department and bureau names, website, position title and code, position category, job description, number of vacancies, work location, interview ratio, and qualification requirements. Some postings explicitly designate an equal number of vacancies for male and female candidates in paired positions, signaling adoption of the one-to-one gender quota. Source: http://bm.scs.gov.cn/pp/gkweb/core/web/ui/business/home/gkhome.html.

[5]Not all cities report county-level tax revenue in the yearbooks.

implement quotas. Since not all positions are subject to the quota, I also examine adoption within counties. Positions involving tax collection are significantly more likely to divide vacancies evenly between men and women than those in administrative roles.[6] These patterns are consistent with anecdotal evidence that tax departments seek to hire more men, based on the belief that men and women have different comparative advantages.

The second part examines the causal impact of the quota. To improve comparability of that have been ever treated and never-treated counties, I use Coarsened Exact Matching (CEM) (Ager et al., 2022; Aneja and Xu, 2022; Azoulay et al., 2019; Iacus et al., 2012; Voth and Xu, 2020), matching counties that adopted the policy to those that never did based on pre-treatment characteristics from the 2010 Census. Treatment intensity is measured as the average share of positions subject to the quota in a county after adoption. On the matched sample, I then estimate effects using a staggered difference-in-differences and event-study framework, exploiting variation in adoption timing across counties.

I begin by studying the effectiveness of the quota. The event study shows that counties implementing the policy experience a decline in the share of female new recruits—about 15 percentage points in the current year and 5 points cumulatively. The number of applicants also falls by 30.7%, with suggestive evidence of a disproportionate drop among women. The 50% cap on female hires may further exclude more qualified women while admitting less qualified men. Consistent with this, cut-off scores fall for male-targeted positions but rise slightly for female-targeted positions, suggesting that the quota displaces marginal rather than top female new recruits, in line with the decline in female applications.

Replacing qualified women with less qualified men could undermine tax collection efficiency, especially since quotas are more often applied to tax-related positions. A one–percentage point increase in the share of quota positions reduces tax revenue by 0.085%. Given that the average share of quota positions in treated counties is 55%, the policy is estimated to reduce tax revenue by 4.7%—equivalent to about $4 million annually per county.

As robustness checks, I show that the results are not driven by local economic conditions or hiring activity: county GDP slightly rises after quota adoption, and new hiring itself

[6]Tax-related positions include responsibilities such as revenue collection, taxpayer services, and auditing; administrative roles include human resources and general office work.

has a small positive effect on revenue. Adjusted event-study estimates confirm robustness to heterogeneous treatment effects (De Chaisemartin and d'Haultfoeuille, 2020; Callaway and Sant'Anna, 2021; Sun and Abraham, 2021; Roth and Sant'Anna, 2023). Alternative productivity measures, such as administrative approvals, show consistent declines. The findings remain robust when using an IV based on even-numbered vacancies, controlling for spillovers from neighboring counties, and applying Conley spatial adjustments to standard errors. Together, these checks reinforce the conclusion that the productivity decline is a direct consequence of quota adoption.

The decline in productivity I document operates through two main channels: a direct effect from reduced worker quality and an indirect effect from institutionalized gender bias. Directly, using written exam cut-off scores as a proxy for candidate quality, I show that a one standard deviation increase in the cut-off score is associated with a 1.5% increase in tax revenue, consistent with revenue losses arising from the admission of less qualified men. Indirectly, institutionalized gender bias undermines productivity through two explanations. The first is patronage networks: in counties with higher corruption, revenue declines following quota adoption are substantially larger, suggesting the policy provides cover for patronage hiring and strengthens corrupt male ties. Supporting this interpretation, I document the prevalence of alcohol-consumption culture and weak regulatory enforcement, which facilitate such practices. The second is resource allocation: tax departments often perceive men as better suited for certain roles or leadership positions, and as the male share of recruits rises, women face fewer opportunities and weaker incentives, further discouraging effort. I also rule out gender composition as an alternative mechanism, pointing instead to reduced worker quality and institutionalized gender bias as key explanations for the productivity decline.

This study contributes to the literature on gender quotas by examining a rare but policy-relevant case: a voluntarily adopted quota favoring men in a context where they are historically advantaged. While most gender quotas aim to increase women's representation (Eckbo et al., 2022; Bhavnani and Lee, 2021; Francis-Tan and Tannuri-Pianto, 2018), quotas for men remain largely unexplored. Yet as women make up a growing share of the public-sector workforce—particularly in education and health care globally (Shi et al., 2019)—debates

over gender balance have prompted the introduction of quotas to preserve male representation. Recent evidence from Schaede and Mankki (2024) shows that abolishing a 40% male quota for teachers in Finland led to long-term losses for primary students, highlighting the potential efficiency gains of such policies. This paper is the first to document both the adoption and impact of a one-to-one gender quota mandating equal hiring of men and women. Unlike traditional threshold-based quotas, this policy enforces strict gender parity, offering new insight into the motivations and consequences of nontraditional quota regimes

More broadly, this paper contributes to the large literature on affirmative action, representation, and efficiency. Recent studies find that affirmative action can enhance both representation and productivity (Black et al., 2023; De Sousa and Niederle, 2022), though much of the focus is on contexts where such policies are prohibited (Schaede and Mankki, 2024; Bleemer, 2022; Ellison and Pathak, 2021; Arcidiacono et al., 2014; Chan and Eyster, 2003). Other work points to potential inefficiencies, mismatch effects (Oliveira et al., 2024; Antonovics and Backes, 2014), or backlash that undermines efficiency (Ip et al., 2020; Leibbrandt et al., 2018). While many studies suggest representation gains can offset efficiency losses, few examine quotas that improve neither. This paper fills that gap by analyzing a voluntary quota that favors a group not historically underrepresented. In doing so, it highlights the unintended consequences of quota design and contributes new evidence to debates over the trade-offs and limits of affirmative action.

Another strand of related literature examines the determinants of bureaucratic performance, emphasizing the role of worker quality in shaping productivity (Fenizia, 2022; Limodio, 2021; Colonnelli et al., 2020; Besley et al., 2017; Dal Bó et al., 2017). This paper contributes to that literature by showing how gender quota policies affect the quality of new hires. Related studies also examine how promotion incentives influence civil servant performance (Bianchi et al., 2023; Bertrand et al., 2020), but few have analyzed how gender quotas affect hiring outcomes—especially in the context of civil service exams in developing countries (Dahis et al., 2023; Aman-Rana, 2023; De Janvry et al., 2023; Chen et al., 2020). This paper is the first to study the effects of a voluntary one-to-one gender quota in China's national civil service exam. More broadly, it sheds light on how such policies influence tax enforcement capacity, a key pillar of fiscal capacity in developing countries (Okunogbe and

Tourek, 2024; Bergeron et al., 2024).

The paper is structured as follows: The next section details civil servant recruitment and tax bureaus. I describe the data and sample in Section 3. Section 4 presents the first set of findings, showing the determinants of adopting the quota. I explain the empirical strategy for estimating the quota's impact in Section 5, before analyzing its effects on gender representation, recruitment outcomes, and economic performance in Section 6. Section 7 discusses mechanisms, and robustness checks are presented in Section 6.4. The final section concludes.

2 Institutional Background

2.1 Civil Servants Recruitment

As the largest and primary recruitment channel for civil servants in China, the National Civil Service Examination (NCSE) attracts millions of college graduates each year (Dahis et al., 2023; Aman-Rana, 2023). The stability and benefits of civil servant positions makes it one of the most competitive examinations in China (Feng et al., 2024; Li et al., 2024). In 2024, the number of applicants for the NCSE exceeded 3 million, competing for 39,600 available positions. A key factor behind the large applicant pool for the National Civil Service Examination (NCSE) is its standardized and transparent recruitment process. Each year, each county-level bureau independently determines its annual hiring plan and submits the plan for approval by upper-level taxation bureaus. And the National Civil Service Administration consolidates and publicly posts these vacancies on the official examination portal. Listings include detailed information on each position—such as location, job responsibilities, number of openings, and eligibility criteria (e.g., education, major, work experience). They also indicate the ratio of shortlisted candidates to final hires. Notably, the "Notes'' section may specify additional conditions, including gender preferences.[7]

Each applicant is permitted to apply for only one position per examination cycle. Then

[7]Before 2016, some postings explicitly required or preferred male candidates, while none favored women. Other conditions may include recent graduation status, housing provisions, minimum service duration, or local Hukou registration.

all qualified candidates must take a standardized national written examination. The number of candidates advancing to the interview stage is determined by the predefined shortlisting ratio. For example, if a position offers two vacancies with a 3:1 ratio, six candidates will be shortlisted for interviews. The written exam cut-off score is determined by the lowest score among the shortlisted candidates—in this case, the sixth-highest score. Consequently, cut-off scores vary across positions and examination years. Final admission is based on the composite score derived from both the written exam and interview. The recruiting department then selects the highest-ranking candidates to fill the available positions. Upon completion of the hiring process, departments publicly disclose information about successful applicants—including their names, gender, and educational backgrounds—on official government websites.

The State Taxation Administration is the largest recruiter in the NCSE, hiring about 12,500 individuals annually and reaching 67% of all new civil servants in 2023 (see Figure A.1).[8] Most positions are entry-level and fall into two broad categories: tax-related and non-tax-related roles, reflecting the division between operational and administrative functions.[9] Since 2016, some counties have gradually implemented a one-to-one gender quota policy at the position level. The policy is subtly enforced by explicitly listing gender requirements in job postings, which appear in matched pairs—one designated for men, the other for women. Before its adoption, county-level tax departments filled positions based solely on civil service exam scores, regardless of gender. Under the quota, certain positions are now gender-targeted, with equal vacancies allocated to male and female candidates. Male-designated positions are open only to men, and female-designated positions only to women. These roles are filled by the highest scorers within each gender group, producing "outcome equality'' through a 50/50 gender split among new hires for those positions.

[8]The sharp decline in 2019 reflects the merger of local and state taxation administrations, which increased staffing levels in tax bureaus by 79 percent and temporarily reduced the need for new hires.

[9]Tax-related roles involve core responsibilities such as tax collection, taxpayer services, and enforcement. Non-tax roles include internal bureau functions such as personnel, budgeting, auditing, and disciplinary oversight.

2.2 Taxation Bureaus

Tax capacity - an important measure of a tax bureau's productivity—tends to grow with economic development (Okunogbe and Tourek, 2024). In 2022, China's tax-to-GDP ratio was 20.05%, well below the OECD average of 34.04% and the U.S. ratio of 27.61%. Unlike the United States, where taxation occurs at the federal, state, and local levels, China's State Taxation Administration operates through a centralized four-tiered system: national headquarters, provincial, municipal, and county levels. County-level bureaus, as the lowest tier, are responsible for direct tax collection and revenue administration.[10] These frontline units collect roughly 85% of total government revenue and manage core functions such as taxpayer registration, liability assessment, compliance enforcement, and service provision for local individuals and businesses. While tax payments are now primarily processed through checks, bank cards, and authorized electronic systems, county bureaus remain central to overseeing withholding, ensuring timely payments, and handling audits, disputes, and refunds. Much of this work depends on personnel capacity, making the quality of tax staff critical to the effectiveness of these operations.

China's taxation system is one of its largest government institutions, employing an average of 500,000 civil servants over the past decade. County-level bureaus typically range in size from 100 to 150 staff members.[11] Since the 1950s, tax bureau personnel were primarily drawn from government-assigned college graduates, military academies, and selected candidates from the labor market. Beginning in 1994, recruitment increasingly relied on standardized examinations, culminating in the 2006 National Civil Service Examination Law, which institutionalized a merit-based and transparent selection process.

Despite these reforms, taxation bureaus have remained male-dominated. Figure A.2 shows consistently high male representation in a sample city over time. While the civil service exam has expanded access for women—who now comprise over 65% of new entrants—men still dominate both overall staffing and leadership roles.

[10]The headquarters oversees 31 provincial bureaus, which manage 535 municipal bureaus and, in turn, 2,924 county bureaus. Each county bureau supervises multiple offices across towns.

[11]Estimated using the number of certificate holders published on the Administrative Law Enforcement Platform (See Figure A.3). Certificate holders account for roughly 75% of total staff, providing a reliable proxy for broader gender patterns.

3 Data

I assemble a comprehensive dataset to study the one-to-one gender quota policy, combining recruitment data, economic outcome data, tax department personnel data, and additional sources for mechanism analysis. The recruitment data capture the adoption of gender quotas, the number of applicants, and candidate quality. The economic data include county-level tax revenue and the number of administrative cases. Internal employee records from tax departments, along with auxiliary data such as corruption indicators, enable analysis of the policy's underlying mechanisms.

3.1 Recruitment Data

Position Information and Gender Quota. The dataset includes 91,149 civil service examination postings from 2011 to 2023, scraped from the official Civil Service Examination website, covering positions in county-level taxation bureaus. Each posting provides detailed information, including the department and bureau names, website, position title and code, position category, job description, number of vacancies, work location, interview ratio, and all qualification requirements. Gender quota adoption is identified based on whether a county's postings include paired gender-targeted positions.

Applicants and admission data. I also web-scrapped the number of applicants and the written exam cut-off scores for each position since 2016. While individual-level applicant data are not available, I collected the names of all candidates who passed the written exam, as well as the names, genders, and educational backgrounds of final admitted candidates from public announcements. By linking the position postings, applicant counts, and admission records, I construct a comprehensive and near-complete dataset for the civil service examination. This allows me to analyze how the quota policy affects the quality of newly hired civil servants.

3.2 Economic Outcome Data

Tax revenue data. To measure the work performance measure of taxation bureaus, I collect county-level tax revenue. The Chinese government, however, does not officially publish the county-level tax revenue data. To address this, I access data from the City Statistical

Yearbook and digitized information on total tax revenue, corporate tax revenue, individual tax revenue, and value-added tax revenue.[12] Not all City Yearbook publish the tax revenue, I collect all City Yearbooks and record all available data.[13] This dataset covers 894 counties across 21 provinces from 2011 to 2022, resulting in 8,188 county-year observations.

Number of administrative cases. An alternative proxy for worker productivity is the number of administrative cases. Each county-level bureau publishes an annual report detailing the number of administrative approvals and penalties. I manually collected this information from official bureau websites, available starting in 2019.

3.3 Other Data

In addition to the core datasets used for the main analysis, I incorporate several supplementary sources. These include data on the number of certificate holders and the gender composition of leadership across multiple provinces to describe the structure of tax departments. I also use confidential individual-level promotion data from one city to explore potential mechanisms. Further, I draw on China's Corruption Investigations Dataset (Wang, 2020) and county-level characteristics from the 2020 Census.

4 What motivates quota adoption?

To understand the voluntary adoption of the one-to-one gender quota policy, I first examine whether pre-determined county characteristics predict adoption. Specifically, I regress a county's adoption decision on its socioeconomic profile and institutional context. By controlling for province fixed effects, I compare counties ever adopted the policy to those never adopted within the same county. As shown in Panel A of Figure 1, counties with higher GDP, larger populations, a greater share of local residents, and lower urbanization rates are more likely to adopt the policy. Adoption is also more common in counties with higher levels of corruption. This section characterizes the patterns and determinants of voluntary adoption, while endogeneity concerns are addressed in the next section.

[12] The validiy of tax revenue data are following the First Digit Law Test, see Figure A.4

[13] Provinces with a non-missing rate above 50% tend to have higher population densities (See Figure A.5).

In addition to county characteristics, I examine the correlation between policy adoption and recruitment patterns. The one-to-one gender quota explicitly aims to balance the gender ratio among new hires. Thus, counties with a higher share of female entrants in recent years may be more likely to adopt the policy. To avoid reverse causality, I restrict the analysis to pre-adoption years in counties that eventually adopted the policy and include counties that never adopted. Controlling for year fixed effects, I find that a higher cumulative female share among new hires in previous years is positively associated with subsequent quota adoption in Panel B of Figure 1. Recruitment size, measured by the cumulative number of new hires in previous years, is also positively associated with quota adoption.

A natural follow-up question is whether counties with higher GDP, larger populations, a greater share of local residents, and lower urbanization rates also have a higher cumulative female share among recent new hires—potentially explaining their greater likelihood of adopting the quota. Figure A.6 shows that the female ratio among prior new hires tends to be higher in wealthier, more urbanized counties with more migrants, more women, and lower corruption—i.e., in more contemporary or progressive areas. This may reflect either greater female applicant selection into civil service or less discrimination in hiring. When combined with Figure 1, the data suggest that while female ratio of recent new entrants partially explain the correlation between county characteristics and quota adoption, they do not fully account for it. Although counties with higher GDP and a higher female share of prior new hires are more likely to adopt the quota, counties with higher GDP per capita and less male ratio do not adopt it. Conversely, some of the adopters are less urbanized, have fewer migrants, exhibit higher corruption, and have lower prior female shares - suggesting that other institutional or culture characteristics also drive adoption.

Anecdotal evidence suggests that some taxation bureaus perceive men and women as having different comparative advantages. In an interview with a city-level tax bureau leader, the official noted challenges in recruiting male candidates, citing greater opportunities for men in the private sector and a lower inclination toward public sector jobs. The leader also expressed a view that men may be better suited for roles involving frequent social interaction, such as business travel, networking events, or client entertainment.

To examine institutional preferences related to gender roles, I analyze quota implementa-

tion at the position level, as not all positions within a county are subject to the policy. Most positions are entry-level roles with equivalent salaries. I define a unique position as one that shares the same county and identical criteria—education, work experience, language skills, Hukou status, and any other special requirements. Position tasks are broadly categorized as either tax-related or non-tax-related. Controlling for position and year fixed effects, I find that positions focused solely on tax-related tasks—such as tax collection, taxpayer services, and enforcement—are significantly more likely to be gender-targeted. In contrast, positions limited to non-tax tasks, such as administration or human resources, are less likely to adopt the quota. Because each position typically recruits only one or a few individuals per year, it is difficult to directly assess gender composition by task. However, task type may correlate with required academic majors, which in turn influence gender composition (See Figure A.7). To address this, I categorize major requirements into four groups—Arts and Humanities, STEM, Business, and Law—and control for them in the regression. The results remain robust, suggesting that quota implementation reflects underlying preferences of tax bureau officials regarding gender roles (See Figure 1).

The rationale for the one-to-one gender quota may reflect concerns about gender discrimination. An alternative explanation is that it responds to perceived threats to male-dominated power structures (see Figure A.8), though further evidence is needed to explore this dimension in depth.

5 Empirical Approach

To assess the impact of the one-to-one gender quota, I exploit its staggered adoption across counties. Using job postings, I identify whether each position is designated for men, women, or expresses a gender preference. Positions with no gender specification are coded as non-quota, while those explicitly designated for men or women are classified as quota positions. In county-level tax bureaus, male and female positions are always posted in equal numbers, consistent with the one-to-one gender quota. A county is defined as adopting the quota in a given year if at least one pair of gender-targeted positions is posted (see Figure A.9).[14]

[14] At minimum, one male and one female position must be posted.

Since counties can decide annually whether to implement the quota or hire at all, I define both the adoption year and all subsequent years as treated, assuming ongoing effects from newly hired civil servants. Not all positions in a given year are gender-targeted. Figure A.10 shows the timing and geographic variation of quota adoption from 2016 to 2023. To capture treatment intensity, I calculate the average share of gender-targeted positions among total postings across treated years. Across counties, the share of gender-targeted positions ranges from 8% to 100%, with an average of 55% among treated counties (see Appendix Figure A.11). The stability of this share over time supports the use of a binary treatment indicator.

Since adoption is voluntarily determined at the county level, a key concern is potential endogeneity. To address this, I use Coarsened Exact Matching (CEM) to match each treated county with counties that never adopted the quota but share similar pre-treatment characteristics.[15] This method generates a counterfactual group that is comparable to treated counties in the joint distribution of observable traits (Ager et al., 2022; Aneja and Xu, 2022; Azoulay et al., 2019; Iacus et al., 2012; Voth and Xu, 2020). The matching variables include tertiary sector share of GDP, population size, male ratio, and urban Hukou ratio from the 2010 Census, allowing me to matching around 62% of observations. Additional characteristics are used in robustness checks.

Figure 2 demonstrates the validity of the matched control group. Before matching, counties with higher GDP, larger populations, a higher ratio of local residents, and a lower urban hukou ratio were more likely to adopt the one-to-one gender quota policy within the same province. After matching, there are no significant differences in these characteristics between counties that adopted the policy and those that did not. Although corruption levels remain imbalanced, I conduct robustness checks that include corruption as one of the matching variables and treat corruption as a potential mechanism to be explored in section 7.2.1. To further address concerns about endogeneity, I implement an instrumental variable approach using the share of even-numbered vacancies as an instrument for treatment intensity.

Treated counties implemented the one-to-one gender quota policy at different times be-

[15] Iacus et al. (2011) show that CEM more effectively reduces imbalance, model dependence, and estimation error than other matching methods, such as propensity score matching and Mahalanobis matching.

tween 2016 and 2023. Although approximately 56% of these counties pause the policy for one or more years—either by stop hiring or suspending the quota — the initial implementation still leads to an accumulation of new hires influenced by the quota. Therefore, I assume that counties adopt the quota beginning in the initiation year and remain treated thereafter. Since the quota policy is implemented at the position level and each county hires for multiple positions, the initiation of the quota may not accurately capture the treatment intensity. To better measure this, I calculate the proportion of gender-targeted positions relative to the total number of positions within the same county. Consistent with the timing of policy adoption, I use the average of this proportion across all post-treatment years as an indicator of treatment intensity.

I then apply a staggered difference-in-differences approach using counties adopted quota and unadopted in the same year, incorporating county fixed effects and year fixed effects:

$$Y_{ct} = \beta Treat_{ct} + \delta_t + \gamma_c + \Theta_c \cdot t + \epsilon_{ct} \tag{1}$$

$$Y_{ct} = \beta CTreat_{ct} + \delta_t + \gamma_c + \Theta_c \cdot t + \epsilon_{ct} \tag{2}$$

where Y_{ct} represents the outcome of interests within county c in year t, including female ratio of new hires, average standardized cut-off score, number of applicants, county-level tax revenue, and number of administrative cases. $Treat_{ct}$ equals 1 when county c initiates the gender quota from year t onwards. $CTreat_{ct}$indicates the treatment intensity measuring the mean share of positions requiring gender in county c from year t onwards. δ_t accounts for the year fixed effect, γ_c denotes the county fixed effect, and ϵ_{ct} represents the error term. I also incorporated the city year trend $\Theta_c \cdot t$ to account for city-level policies that may exhibit variations over time. For example, the Value Added Tax Reform, trialed from 2012 to 2016 in selected cities, led to an increase in local tax revenue (Feng et al., 2024). To estimate the impact of adopting the one-to-one gender quota, I turn to an event study to examine the effects of both the policy's introduction and intensity (Miller, 2023):

$$Y_{ct} = \sum_{k=-a}^{b} \theta_k I\{t = t_c^* + k\} + \delta_t + \gamma_c + \Theta_c \cdot t + \epsilon_{ct} \tag{3}$$

The window of event study varies across different outcomes because of differences in data availability. Specifically, the analysis spans 5 years before and after the event for female ratio, cut-off scores, and number of applicants; 7 years before and 3 years after for tax revenue; and 5 years before and 4 years after for the number of administrative cases. $I\{t = t_c^* + k\}$ is an indicator variable for event time t_c^*, meaning that the quota policy took place k periods before or after this observation's calendar time. A separate term is included for each event time. The key coefficients of this specification are the θ_k. The coefficients after the event (θ_k for $k \geq 0$) capture the dynamic effects of the gender quota. The terms before the event (θ_k for $k < 0$) provide a falsification test. Similarly, δ_t accounts for the year fixed effect, γ_c denotes the county fixed effect, and ϵ_{ct} represents the error term. $\Theta_c \cdot t$ indicates the city year trend.

6 Impacts of the One-to-one Quota

6.1 First Stage Result

To investigate the impact of the one-to-one gender quota, the first step is to examine whether it changed the female share of new hires. Figure 3 presents the event study estimates of the quota's adoption on the female hiring ratio following the main specification. Since data on new hires is available only from 2016, the event study is limited to a shorter window, covering five years before and after the treatment. The increase observed in the year prior to treatment aligns with the findings in the determinants section – counties with a higher female share among prior new hires are more likely to adopt the gender quota. The cumulative female share of new hires since 2016 exhibits a parallel trend. I find one percentage increase of gender-targeted position share led to a 0.25% decrease in the female share of new hires, relative to a baseline average of 67.7%. Among all new hires since 2016, one percentage increase of gender quota intensity corresponds to 0.08% decrease in the female share. Given that the average share of gender-targeted positions among treated counties exceeds 55%, the adoption of the quota is associated with a 13.8 percentage point reduction in the female share of new hires in the current year and a 4.4 percentage point reduction in the cumulative female

share of new hires since 2016. Similar patterns emerge in the basic regression estimates for both the introduction of the treatment and its intensity (See Figure A.12).

The significant drop in the female share of new hires suggests that the quota effectively achieved its intended goal. Another notable pattern is the spike in the current-year female share just before treatment, consistent with the finding that treated counties had higher female representation among prior new hires.

6.2 Recruitment Outcome

The intended shift in gender composition may influence applicants' behavior and quality, by excluding more qualified women while admitting less qualified men. I then analyze the effects on candidates' quality and application change in the following section.

Written Exam Cut-off Score. Although the quota achieved gender parity among new entrants, a key concern for recruitment efficiency is whether taxation bureaus can still select the most qualified candidates. To assess this, I examine changes in candidate quality, proxied by the written exam cut-off score—defined as the lowest score among applicants shortlisted for interviews.

Analyzing county-level averages, I find that the introduction of the gender quota reduces the cut-off score by 0.124 standard deviations. Furthermore, each one percentage point increase in the share of gender-targeted positions is associated with a 0.223 standard deviation decline in the cut-off score (see Panel A of Table A.1). Panel A in Figure 4 presents the corresponding event study, both before and after matching method, showing no pre-trend.

Since the one-to-one gender quota restricts further increases in female representation while expanding opportunities for men, I next examine whether it differentially affects candidate quality by gender. The county-level average cut-off score for each gender before the quota is based on gender-neutral positions, while post-quota scores are calculated separately using only male-targeted and female-targeted positions. Panel B of Figure 4 shows significant gender differences in cut-off scores. On average, male scores decline by 0.772 standard deviations, while female scores increase by 0.116 standard deviations(See Panel B of Table A.1). The event study reveals that male candidate quality drops sharply following the quota's introduction and remains low, whereas female candidate quality rises modestly and

stabilizes at pre-quota levels.

Given that tax officials perceive men and women as better suited for different roles—and that tax-related positions are more likely to adopt the quota than non-tax-related ones—it is important to examine how the policy affects candidates across task types. I analyze changes in cut-off scores at the position level, controlling for unique position fixed effects (i.e., positions with identical criteria). As shown in Table 2, cut-off scores decline when a position becomes male-targeted and increase slightly when it becomes female-targeted. See Figure A.13 and A.14 for event studies. Moreover, the decline in male candidate quality is more pronounced in positions with tax-related tasks. In other words, when taxation bureaus apply the quota more frequently to tax-related positions—based on the belief that men are better suited for such roles—the resulting efficiency loss is greater in these positions, which are critical to tax collection.

Number of Applicants. Even if applicant behavior remains unchanged, the decline in candidate quality may result from admitting less qualified men while excluding moderately qualified women. However, gender restrictions in job postings may also influence applicants' behavior. Although individual-level applicant data are unavailable, I observe the number of applicants per position. Using an event study framework, Figure 5 shows that the number of applicants declines significantly following the adoption of the gender quota. Specifically, as shown in Panel A of Table A.2, the introduction of the gender quota reduces the number of applicants by 45, relative to a baseline average of 292. A one-percentage-point increase in the share of gender-targeted positions decreases the number of applicants by 89, representing 30.67% of the average.

A key challenge in analyzing applicant behavior by gender is the lack of data on the gender composition of the applicant pool. I observe the number of male and female applicants for gender-targeted positions, but not for positions prior to the implementation of gender quota. To approximate this, I infer the gender ratio of applicants for gender-neutral positions from the pool of shortlisted candidates, under the assumption that men and women have similar score distributions. Although the gender of shortlisted applicants is not directly available, names are published. I use a Python-based Chinese name gender prediction tool to estimate

the gender of interviewees and, by extension, the gender distribution of applicants.[16] Panel B of Table A.2 provides suggestive evidence that women are more discouraged from applying when a county adopts the gender quota policy.

If cut-off scores were linearly correlated with the number of applicants, a decline in female applicants would be expected to lower the cut-off score for female-targeted positions. However, despite the drop in female applicant numbers, the cut-off score for these positions increases, suggesting two possible explanations. First, although the total number of female applicants declines, their chances of being hired may also fall—some may be replaced by less qualified male applicants. If the reduction in applicant volume does not offset the decline in hiring probability, the cut-off score could rise despite fewer applicants. Second, the women discouraged from applying may not be among the most competitive candidates, meaning their absence has little effect on the cut-off score among top-performing applicants.

Since the number of vacancies in a county may vary across years, I control for vacancies in all regressions examining both candidate quality and applicant volume. A higher number of vacancies is associated with lower cut-off scores and more applicants. This also suggests that cut-off scores do not linearly correlate with applicant volume, while it depends on which segment of applicants opts out of applying.

While the one-to-one gender quota policy achieves gender parity among new entrants, it reduces recruitment efficiency by admitting less qualified men and crowding out moderately qualified women. This inefficiency is more pronounced in positions where taxation bureaus perceive men to perform better.

6.3 Economic Impacts

Given the well-documented role of personnel in revenue generation, recruitment inefficiencies may translate into productivity losses. In this section, I use county-level tax revenue as a primary measure of taxation bureau productivity to show the economic impacts.

As shown in Column (1) of Table 1, a one-percentage-point increase in the share of gender-targeted positions leads to a 0.085% decline in total tax revenue. Given a mean share of 55%, the one-to-one gender quota is estimated to reduce total tax revenue by approximately

[16]The "ngender" package, which has higher accuracy for predicting male names than female names.

4.7%. With average county-level tax revenue exceeding 6 billion RMB, this translates to an estimated loss of 30 million RMB—or roughly 4 million USD. I also controlled the pre-treatment matching variables and year fixed effect in column (3), and the result is still robust and even has a larger magnitude. The sample shrinks because some control variables have missing values. The event study in Figure 6 shows no pre-trends and a significant decline in total tax revenue following the gender quota, showing the treatment intensity effect. Column (2) and (4) suggest the effect of gender quota introduction. The ever-treated variable also includes counties that adopt the quota in only a small share of positions, making it difficult to isolate the true effect of policy introduction, as the estimated impact may reflect a mix of partial and minimal implementations (See Figure A.15 for event study results). I also show the effect of treatment intensity before matching method in Figure A.16 for robustness check.

6.3.1 Number of Administrative Cases

An alternative measure for personnel productivity is the annual number of administrative cases the tax workers processed. The administrative cases include administrative approvals and penalties. Administrative approvals—such as requests for invoice printing, tax deferral, or quota adjustments—reflect routine interactions and staff workload. In contrast, penalties are issued only for taxpayer noncompliance, are often waived for minor first-time violations, and depend more on taxpayer behavior than on bureaucratic effort.

Figure 7 presents the effects of the gender quota on both administrative approvals and penalties. The annual number of penalties remains unchanged, consistent with the view that penalties are driven more by taxpayer behavior than by bureaucratic input. In contrast, administrative approvals decline significantly after the adoption of the gender quota. Although estimates from three to four years before adoption fall slightly below zero, they show no clear pre-trend, suggesting the post-quota decline is not driven by pre-existing differences. The data for number of administrative cases covers only 2019 to 2023, restricting the pre-treatment window to at most five years. The results remain robust when restricting the sample to counties with complete tax data, despite the shorter time window (See Figure A.17).

6.4 Robustness Check

I also present a series of robustness checks across different samples and specifications, addressing several potential concerns.

First, to rule out the possibility that declines in tax revenue and administrative cases are driven by reduced local economic activity, I regress county-level GDP on the one-to-one gender quota policy. As shown in Figure A.19, the policy does not reduce GDP and is associated with a slight increase, mitigating concerns that revenue losses reflect broader economic downturns. Second, since the quota is implemented only when counties are hiring, a potential concern is that hiring itself—rather than the quota—drives the observed effects, possibly through changes in workforce composition. To address this, I use the same specification with "ever hired" as the treatment variable. The event study in Figure A.20 shows that hiring has a slightly positive effect on tax revenue, ruling out hiring as a confounding factor. Column (1) of Table A.3 further shows that the main results—on applicant numbers, candidate quality, and tax revenue—remain robust when the sample is restricted to counties that engaged in hiring. In addition, since some counties discontinue the quota after initial adoption, I also restrict the sample to counties that consistently applied the policy once adopted. Column (2) of Table A.3 shows an even larger and statistically significant effect, reinforcing the robustness of the main findings. The findings also remain robust when clustering standard errors using Conley spatial adjustment (see Table A.4) and when restricting the sample to match that used for controlling pre-treatment matching variables and year fixed effects (See Table A.5).

To address concerns about the validity of the TWFE estimator, Figure A.18 also reports adjusted event-study estimates that are robust to treatment effect heterogeneity across time or groups (De Chaisemartin and d'Haultfoeuille, 2020; Callaway and Sant'Anna, 2021; Sun and Abraham, 2021; Roth and Sant'Anna, 2023). To complement the matching strategy, I implement an instrumental variable approach. Since the one-to-one gender quota is implemented through paired postings, adopting counties tend to post an even number of gender-targeted vacancies. I use the share of even-numbered vacancies as an instrument for treatment intensity and estimate a 2SLS regression, reported in Table 3. Column (3) shows

that a one-percentage-point increase in the gender-targeted share reduces tax revenue by 0.138%. The instrument is strongly correlated with treatment intensity, with a first-stage F-statistic of 386.93, indicating strong relevance. The 2SLS estimate is substantially larger in magnitude than the OLS estimate reported in column (1), suggesting that OLS understates the true negative effect of the quota. This pattern can be attributed to the positive selection, where unobserved factors correlated with both quota adoption and stronger tax performance attenuate the estimated effect.

A potential limitation of the instrumental variable approach is the assumption that the share of even-numbered vacancies is exogenous to tax revenue outcomes. Anecdotal evidence suggests that some taxation bureaus may strategically delay or reschedule vacancies to create an even number and enable quota adoption. In the regression sample, the average share of even-numbered vacancies rises from approximately 50% before treatment to 70% after treatment. While the even-numbered structure is not directly linked to tax revenue, it is possible that underlying factors influencing vacancy planning may also affect fiscal outcomes. To address this concern, I restrict the sample to counties that consistently post an even number of vacancies. Within this group, counties that do not adopt the quota may be less aware of or responsive to the pairing requirement, helping mitigate concerns about endogeneity in the instrument. The results in column (4) to (6) are still robust.

Another concern is the potential for spillover effects, where the adoption of the gender quota in one county may influence outcomes in neighboring counties. If untreated counties are indirectly affected by nearby adoption, the estimated treatment effect could be biased. To address this, I conduct robustness checks that control for the average share of gender-targeted positions in neighboring counties within the same city. As shown in Table A.6, the main results—on tax revenue, cut-off scores, and number of applicants—remain unchanged, suggesting that spillover effects do not drive the findings. The event study also shows no pre-trend and the effect is robust (See Figure A.21).

7 Channels

The results indicate that the one-to-one gender quota reduces institutional productivity, as reflected in both tax revenue and the number of administrative approvals. The quota's implementation coincided with a decline in female representation and may have undermined merit-based selection, leading to the hiring of less qualified male candidates (see Section 6). To explore the underlying mechanisms, I examine two potential channels through which the increased male share and decreased candidate quality may have contributed to this decline.

First, the female representation and candidate quality may directly influence work productivity: a lower female ratio lowers the productivity, or lower candidate quality lowers the productivity. Second, conservative gender norms may reinforce existing male-dominated networks, place greater emphasis on men's roles, and ultimately reduce productivity—particularly in institutions where weaker capacity is reflected in issues such as corruption. Counties may aim to attract more men either because they believed men were better suited for certain roles, or because they sought to preserve existing male-dominated networks and maintain institutional control.

In the following, I document limited scope for representation and meritocracy and two pieces of evidence that underscore the effect of gender norms as a driving force. While the direct effect of gender composition on tax revenue appears limited, the candidate quality is positively related to both tax revenue and number of administrative approvals. To access the influence of gender norms, I first examine the heterogeneity in corruption, which is more prevalent in male-dominated systems (Dollar et al., 2001; Swamy et al., 2001; Brollo and Troiano, 2016). Counties with higher levels of corruption experience larger negative effects following the quota's implementation. In such settings, the quota may be co-opted to reinforce patronage among male civil servants. Supporting this interpretation, I provide additional evidence on alcohol consumption culture and regulatory laxity, both of which facilitate corruption. I then disentangle the effects of candidate quality and corruption to demonstrate how corruption amplifies the negative impact. Finally, I offer suggestive evidence that tax departments' perception of men as better suited for certain roles or senior positions may also undermine female workers' motivation and incentives.

7.1 Gender Representation and Meritocracy

7.1.1 Meritocracy and Performance

One recruitment outcome is a decline in the written exam cut-off score, which serves as a proxy for candidate quality. If lower-quality candidates reduce tax revenue, the observed decline may reflect a loss in tax capacity linked to personnel ability. Figure 8 shows a positive correlation between cut-off scores and productivity before treatment, measured by tax revenue and the annual number of administrative approvals, using the main regression specification after matching. Higher cut-off scores are associated with both higher tax revenue and more approvals. Table A.7 indicates that a one standard deviation increase in the cut-off score corresponds to a 1.5% increase in tax revenue.

7.1.2 Gender Composition and Performance

The one-to-one gender quota may affect the gender composition of the taxation bureaus, which may also affect the workers' productivity. I first examine whether the female ratio of new entrants – both current year new entrants and cumulative new entrants – affect the tax revenue. Table A.11 shows the female ratio of cumulative new entrants show positive relationship with the tax revenue, though the results for other kinds of female ratio are not significant.

7.2 Gender Norm and Institutional Capacity

7.2.1 Male-Dominated Networks and Corruption

Analysis of quota adoption patterns reveals that counties with higher levels of corruption are more likely to adopt the policy, even if they had a lower share of female new entrants prior to implementation. Unobserved factors associated with corruption—such as cultural norms or patronage networks—may simultaneously influence voluntary quota adoption and contribute to declines in tax revenue. Prior work also finds a negative correlation between female representation and corruption (Decarolis et al., 2023). To explore this, I examine the effect of the gender quota by corruption level. Figure 9 shows that counties with a

higher number of corruption cases from 2010 to 2015 experience larger negative impacts on productivity, as measured by both tax revenue and administrative cases. The decline in candidate quality, reflected in lower cut-off scores, is also more pronounced in highly corrupt counties compared to less corrupt ones (see Table A.8 and Figure A.22).

Alcohol Consumption Culture. In China, corruption is often linked to alcohol consumption, as drinking facilitates the building of *guanxi* through informal interactions and serves as a vehicle for gift-giving and unofficial negotiations. Recognizing this, Xi Jinping's anti-corruption campaign explicitly targeted lavish banquets, where alcohol played a central role. Given this context, I also explore whether local drinking culture may be associated with the impact of the gender quota policy. I collect the city list known for strong drinking culture, ranked by weighted city-level alcohol industry GDP.[17] Table A.10 shows that counties located in these cities experience a larger negative effect of the gender quota on productivity.

Space for Corruption. The heterogeneity analysis across tax categories shows that the gender quota has the largest effect on "other tax'' – excluding individual tax, corporate tax and VAT, is the largest (See Figure A.23). Other taxes are less legalized, including environmental protection tax, property tax, urban land use tax, land appreciation tax, resource tax, and tobacco leaf tax, which allow more discretion and flexibility in collection. This finding reinforces the interpretation that the observed reduction in tax revenue is linked to changes in personnel ability or behavior.

Corruption and Meritocracy. To disentangle the effects of candidate quality and corruption, I examine how cut-off scores affect tax revenue across counties with different levels of corruption, using only pre-treatment data to isolate the relationship from the influence of the gender quota. Table A.9 presents the results. Column (1) shows a positive relationship between score and tax revenue, consistent with findings in earlier sections. Since the corruption data are cross-sectional, I control for city fixed effects and county-level covariates rather than county fixed effects. Column (2) shows no significant direct relationship between corruption level and tax revenue. However, column (3) shows that, after controlling for the cut-off score, corruption is positively associated with tax revenue. This suggests that

[17] http://www.gtkjgh.org.cn/zhongcheng/news/show/id/14520.html

corruption may also reflect other factors that stimulate revenue generation. To examine potential non-linear relationships between corruption and tax revenue, column (4) indicates that the positive effect of candidate quality on tax revenue is weaker in more corrupt counties, suggesting that corruption may hinder the productivity gains from higher-ability hires. Given that more corrupt counties are also more likely to adopt the quota—and experience larger declines in candidate quality—this interaction may amplify the negative impact on tax revenue by both reducing average ability and weakening its marginal returns.

7.2.2 Gender Discrimination and Role Allocation

Position Allocation Panel B of Figure 1 shows that positions related to tax collection are more likely to adopt the quota, likely aiming to recruit more men. To provide further evidence on gender-based position allocation, I webscraped names and genders of both current civil servants and leadership personnel to examine gender composition. Since tax departments do not publicly disclose full staff rosters, I instead collected name lists of administrative enforcement inspection certificate holders in all counties of Anhui Province. These certificates are required for tax workers to conduct inspections and enforce regulations. As shown in Figure A.24, most counties have a male majority among certificate holders. Additionally, Figure A.8, based on data from four provinces, shows that female representation among leadership positions is extremely low. This evidence suggests that tax departments may perceive men as better suited for certain roles or senior positions.

Gender Difference of Promotion Speed To understand the work incentives for current workers, I use a confidential individual-level employee roster from one city in Anhui province, which includes each employee's position level and the date they attained that level. I construct a measure of promotion speed by dividing the number of years typically required to reach the current level by the actual number of years it took the employee to be promoted. A higher value indicates higher promotion speed. I exclude individuals at the lowest and highest ranks, as their promotion paths may be mechanically biased (see Figure A.25). I then calculate the male ratio within each employee's cohort, defined by a five-year age bin, both including and excluding the individual. As shown in Table A.12, and controlling for county fixed effects, age, and age squared, I find that a higher male ra-

tio within the same cohort is associated with slower promotion overall—for both men and women. However, the negative effect is significantly stronger for women, suggesting that being surrounded by more male peers may hinder women's career progression. Although the analysis is based on a single city and may not be fully generalizable, the results offer suggestive evidence that a higher male ratio in the cohort can reduce promotion speed for women, potentially weakening their work incentives in a system where promotion is a key career reward.

8 Conclusion

This paper provides the first empirical evidence on a one-to-one gender quota that favors men in a public sector setting where women dominate new hires but remain underrepresented. Unlike the existing literature, which focuses on mandatory, top-down gender quotas, I examine the adoption and consequences of a voluntarily implemented policy. Drawing on comprehensive novel data from China's civil service examination and taxation bureaus, I show that counties with a higher share of female new entrants prior to the policy are more likely to implement the quota. Within counties, tax-related positions are significantly more likely to be designated as gender-targeted compared to non-tax-related positions. I use the Coarsened Exact Matching (CEM) method to construct a balanced sample. Using matched samples, I apply a staggered difference-in-differences and event-study framework, leveraging variation in policy adoption timing across counties. While the quota policy successfully imposes numerical gender parity among new hires, it reduces the overall quality of recruited candidates, discourages applicant participation—particularly among women—and significantly lowers bureaucratic productivity, as measured by tax revenue and administrative performance. An exploration of potential mechanisms—including meritocracy, corruption, and peer gender composition—provides suggestive evidence that the gender quota reduces productivity by lowering both worker ability and incentives.

While the broader implications for social welfare warrant further discussion, this paper provides empirical evidence that a voluntarily adopted quota can reduce recruitment efficiency and institutional productivity among adopters. The analysis contributes to ongoing

debates about the need for reverse affirmative action when historically advantaged groups become outnumbered, offering important policy insights. Beyond the public sector, similar dilemmas arise in higher education and private organizations—particularly when decision-makers still come from historically advantaged groups. These institutions must grapple with whether to adopt policies that preserve representation for groups no longer underrepresented. This debate is not only crucial for advancing gender equity but also for ensuring continued progress for other historically disadvantaged groups whose upward mobility may be constrained by such policies.

References

Ager, Philipp, Leonardo Bursztyn, Lukas Leucht, and Hans-Joachim Voth, "Killer incentives: Rivalry, performance and risk-taking among German fighter pilots, 1939–45," *The Review of economic studies*, 2022, *89* (5), 2257–2292.

Aman-Rana, Shan, "Does Ability Matter for Discretionary Promotions in Bureaucracies? Evidence from Pakistan," *Evidence from Pakistan (November 1, 2023)*, 2023.

Aneja, Abhay and Guo Xu, "The costs of employment segregation: Evidence from the federal government under Woodrow Wilson," *The Quarterly Journal of Economics*, 2022, *137* (2), 911–958.

Antonovics, Kate and Ben Backes, "The effect of banning affirmative action on college admissions policies and student quality," *Journal of Human Resources*, 2014, *49* (2), 295–322.

Arcidiacono, Peter, Esteban Aucejo, Patrick Coate, and V Joseph Hotz, "Affirmative action and university fit: Evidence from Proposition 209," *IZA Journal of Labor Economics*, 2014, *3*, 1–29.

Azoulay, Pierre, Christian Fons-Rosen, and Joshua S Graff Zivin, "Does science advance one funeral at a time?," *American Economic Review*, 2019, *109* (8), 2889–2920.

Basri, M Chatib, Mayara Felix, Rema Hanna, and Benjamin A Olken, "Tax administration versus tax rates: evidence from corporate taxation in Indonesia," *American Economic Review*, 2021, *111* (12), 3827–3871.

Bergeron, Augustin, Gabriel Tourek, and Jonathan L Weigel, "The state capacity ceiling on tax rates: Evidence from randomized tax abatements in the drc," *Econometrica*, 2024, *92* (4), 1163–1193.

_ , Pedro Bessone, John Kabeya Kabeya, Gabriel Z Tourek, and Jonathan L Weigel, "Optimal assignment of bureaucrats: Evidence from randomly assigned tax collectors in the DRC," Technical Report, National Bureau of Economic Research 2022.

Bertrand, Marianne, Robin Burgess, Arunish Chawla, and Guo Xu, "The glittering prizes: Career incentives and bureaucrat performance," *The Review of Economic Studies*, 2020, *87* (2), 626–655.

Besley, Timothy and Torsten Persson, "Why do developing countries tax so little?," *Journal of economic perspectives*, 2014, *28* (4), 99–120.

_ , Olle Folke, Torsten Persson, and Johanna Rickne, "Gender quotas and the crisis of the mediocre man: Theory and evidence from Sweden," *American economic review*, 2017, *107* (8), 2204–2242.

Bhavnani, Rikhil R and Alexander Lee, "Does affirmative action worsen bureaucratic performance? Evidence from the Indian administrative service," *American Journal of Political Science*, 2021, *65* (1), 5–20.

Bianchi, Nicola, Giulia Bovini, Jin Li, Matteo Paradisi, and Michael Powell, "Career spillovers in internal labour markets," *The Review of Economic Studies*, 2023, *90* (4), 1800–1831.

Black, Sandra E, Jeffrey T Denning, and Jesse Rothstein, "Winners and losers? the effect of gaining and losing access to selective colleges on education and labor market outcomes," *American Economic Journal: Applied Economics*, 2023, *15* (1), 26–67.

Bleemer, Zachary, "Affirmative action, mismatch, and economic mobility after California's Proposition 209," *The Quarterly Journal of Economics*, 2022, *137* (1), 115–160.

Bó, Ernesto Dal, Frederico Finan, Olle Folke, Torsten Persson, and Johanna Rickne, "Who becomes a politician?," *The Quarterly Journal of Economics*, 2017, *132* (4), 1877–1914.

Brollo, Fernanda and Ugo Troiano, "What happens when a woman wins an election? Evidence from close races in Brazil," *Journal of Development Economics*, 2016, *122*, 28–45.

Callaway, Brantly and Pedro HC Sant'Anna, "Difference-in-differences with multiple time periods," *Journal of econometrics*, 2021, *225* (2), 200–230.

Chaisemartin, Clément De and Xavier d'Haultfoeuille, "Two-way fixed effects estimators with heterogeneous treatment effects," *American economic review*, 2020, *110* (9), 2964–2996.

Chan, Jimmy and Erik Eyster, "Does banning affirmative action lower college student quality?," *American Economic Review*, 2003, *93* (3), 858–872.

Chattopadhyay, Raghabendra and Esther Duflo, "Women as policy makers: Evidence from a randomized policy experiment in India," *Econometrica*, 2004, *72* (5), 1409–1443.

Chen, Ting, James Kai sing Kung, and Chicheng Ma, "Long live Keju! The persistent effects of China's civil examination system," *The economic journal*, 2020, *130* (631), 2030–2064.

Cohen, "Does Admissions Fairness Make My Campus Look Too Female?," *The Washington Post*, March 2024.

Colonnelli, Emanuele, Mounu Prem, and Edoardo Teso, "Patronage and selection in public sector organizations," *American Economic Review*, 2020, *110* (10), 3071–3099.

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

Decarolis, Francesco, Raymond Fisman, Paolo Pinotti, Silvia Vannutelli, and Yongxiang Wang, "Gender and bureaucratic corruption: Evidence from two countries," *The Journal of Law, Economics, and Organization*, 2023, *39* (2), 557–585.

Dollar, David, Raymond Fisman, and Roberta Gatti, "Are women really the "fairer" sex? Corruption and women in government," *Journal of Economic Behavior & Organization*, 2001, *46* (4), 423–429.

Dominus, "There Was Definitely a Thumb on the Scale to Get Boys," *The New York Times Magazine*, September 2023.

Eckbo, B Espen, Knut Nygaard, and Karin S Thorburn, "Valuation effects of Norway's board gender-quota law revisited," *Management Science*, 2022, *68* (6), 4112–4134.

Ellison, Glenn and Parag A Pathak, "The efficiency of race-neutral alternatives to race-based affirmative action: Evidence from Chicago's exam schools," *American Economic Review*, 2021, *111* (3), 943–975.

Feng, Bo, Qiwei He, Xin Jin, and Xu Xu, "Authoritarian State Building and Talent Attraction: Evidence from China's Civil Servant Fever," *Available at SSRN 4779617*, 2024.

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

Francis-Tan, Andrew and Maria Tannuri-Pianto, "Black Movement: Using discontinuities in admissions to study the effects of college quality and affirmative action," *Journal of Development Economics*, 2018, *135*, 97–116.

Holzer, Harry and David Neumark, "Assessing affirmative action," *Journal of Economic literature*, 2000, *38* (3), 483–568.

Iacus, Stefano M, Gary King, and Giuseppe Porro, "Multivariate matching methods that are monotonic imbalance bounding," *Journal of the American Statistical Association*, 2011, *106* (493), 345–361.

_ , _ , and _ , "Causal inference without balance checking: Coarsened exact matching," *Political analysis*, 2012, *20* (1), 1–24.

Ip, Edwin, Andreas Leibbrandt, and Joseph Vecci, "How do gender quotas affect workplace relationships? Complementary evidence from a representative survey and labor market experiments," *Management Science*, 2020, *66* (2), 805–822.

Janvry, Alain De, Guojun He, Elisabeth Sadoulet, Shaoda Wang, and Qiong Zhang, "Subjective performance evaluation, influence activities, and bureaucratic work behavior: Evidence from china," *American Economic Review*, 2023, *113* (3), 766–799.

Khan, Adnan Q, Asim Ijaz Khwaja, and Benjamin A Olken, "Making moves matter: Experimental evidence on incentivizing bureaucrats through performance-based postings," *American Economic Review*, 2019, *109* (1), 237–270.

Krook, Mona Lena, *Quotas for women in politics: Gender and candidate selection reform worldwide*, Oxford University Press, 2010.

Leibbrandt, Andreas, Liang Choon Wang, and Cordelia Foo, "Gender quotas, competitions, and peer review: Experimental evidence on the backlash against women," *Management Science*, 2018, *64* (8), 3501–3516.

Li, Hongbin, Lingsheng Meng, Yanyan Xiong, and Sinclair L Cook, "Job Preferences and Outcomes for China's College Graduates," *The China Quarterly*, 2024, *258*, 529–547.

Limodio, Nicola, "Bureaucrat allocation in the public sector: Evidence from the world bank," *The Economic Journal*, 2021, *131* (639), 3012–3040.

Mitchell, Brian, *International historical statistics: Europe 1750-1993*, Springer, 1998.

Murray, Rainbow, "Quotas for men: Reframing gender quotas as a means of improving representation for all," *American Political Science Review*, 2014, *108* (3), 520–532.

Nietzel, "Women Continue to Outpace Men in College Enrollment and Graduation," *Forbes*, August 2024.

Okunogbe, Oyebola and Gabriel Tourek, "How can lower-income countries collect more taxes? The role of technology, tax agents, and politics," *Journal of Economic Perspectives*, 2024, *38* (1), 81–106.

Oliveira, Rodrigo, Alei Santos, and Edson Severnini, "Bridging the gap: Mismatch effects and catch-up dynamics under a Brazilian college affirmative action program," *Economics of Education Review*, 2024, *98*, 102501.

Pande, Rohini, "Can mandated political representation increase policy influence for disadvantaged minorities? Theory and evidence from India," *American economic review*, 2003, *93* (4), 1132–1151.

Pollitt, "Affirmative Action for Men?," *The Nation*, October 2023.

Roth, Jonathan and Pedro HC Sant'Anna, "Efficient estimation for staggered rollout designs," *Journal of Political Economy Microeconomics*, 2023, *1* (4), 669–709.

Sang-Hun, "Korean Women Flock to Government," *The New York Times Magzine*, March 2010.

Schaede, Ursina and Ville Mankki, "Long-Term Gains from an Unusual Gender Quota," 2024.

Schieder, Chelsea Szendi, "A "necessary evil"? Keeping women out of medical schools won't fix what ails the Japanese medical profession," *Asia-Pacific Journal*, 2019, *17* (7), e3.

Shi, RK, Kerenssa Kay, and Ravi Somani, "5 Facts about gender equality in the public sector," 2019.

Sousa, José De and Muriel Niederle, "Trickle-down effects of affirmative action: A case study in France," Technical Report, National Bureau of Economic Research 2022.

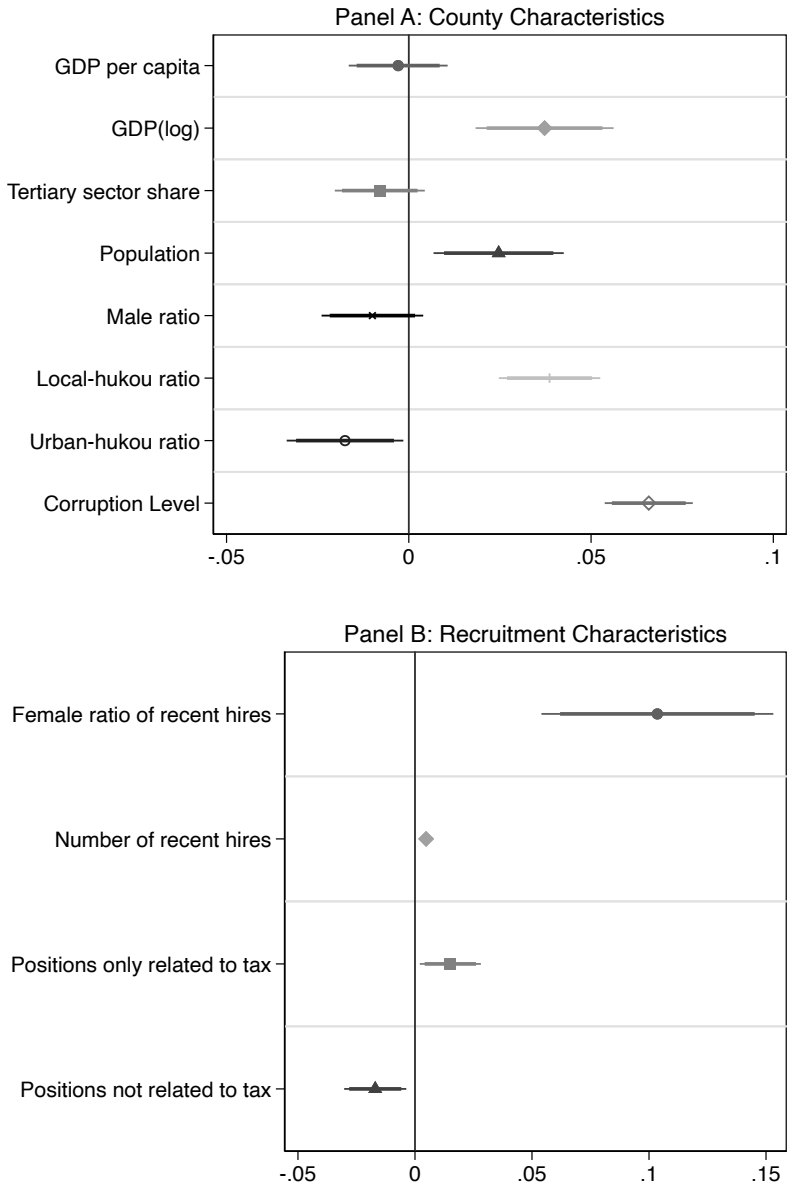
Sun, Liyang and Sarah Abraham, "Estimating dynamic treatment effects in event studies with heterogeneous treatment effects," *Journal of econometrics*, 2021, *225* (2), 175–199.

Swamy, Anand, Stephen Knack, Young Lee, and Omar Azfar, "Gender and corruption," *Journal of development economics*, 2001, *64* (1), 25–55.

Voth, Joachim and Guo Xu, "Discretion and destruction: promotions, performance, and patronage in the Royal Navy," Technical Report, CEPR Discussion Paper No. DP13963 2020.

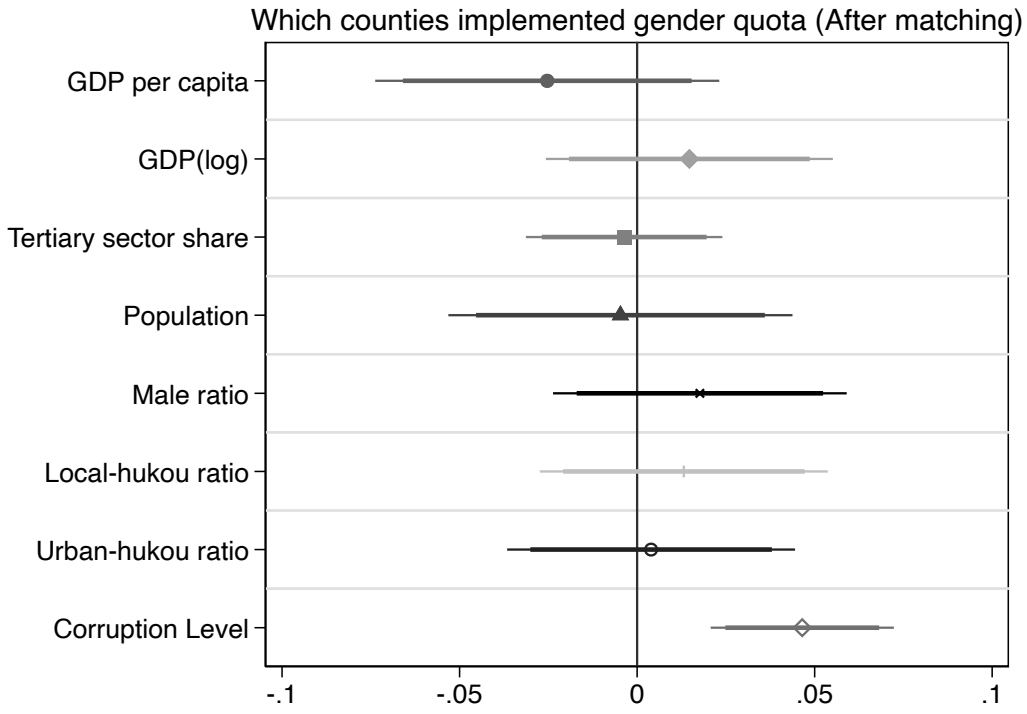
Wang, Yuhua, "China's Corruption Investigations Dataset," 2020.

Figure 1: Who Adopted the Quota?



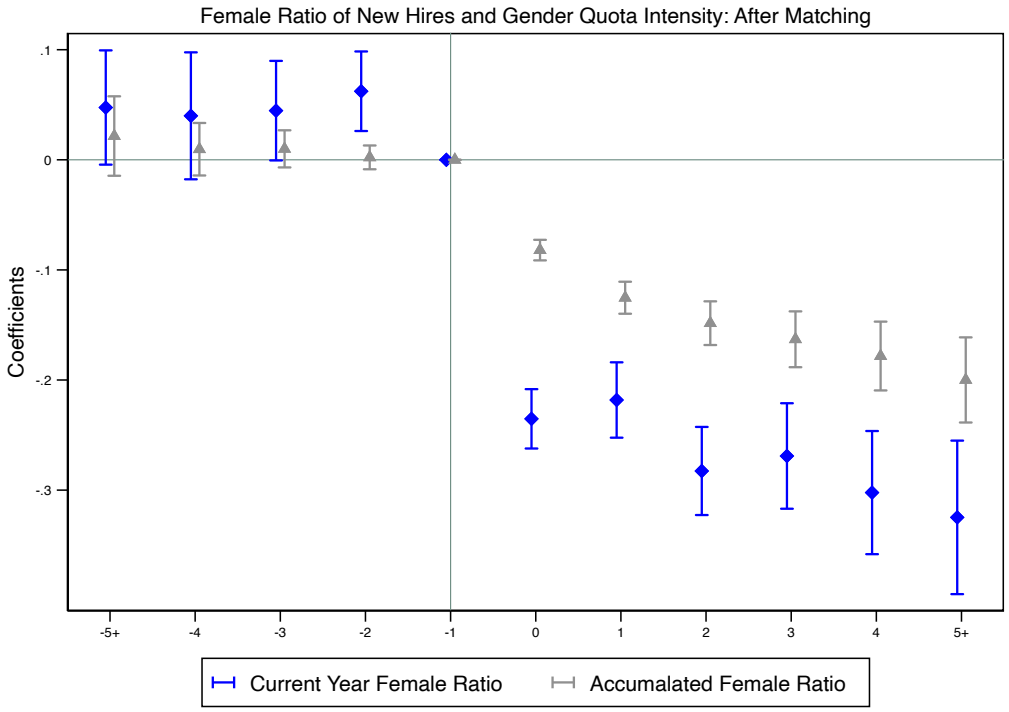
Notes - This figure presents the determinants of gender quota adoption, incorporating both county-level and recruitment characteristics. *Panel A* displays the correlations between quota adoption and pre-treatment county characteristics, using county-level data from the 2010 Census. Corruption levels are sourced from China's Corruption Investigations Dataset (Wang, 2020). Each regression in Panel A includes year and province fixed effects. *Panel B* illustrates the relationship between quota adoption and recruitment characteristics. The female share of recent hires and the total number of recent hires are calculated using admission data from 2016 onward. Positions are classified into tax-related and non-tax-related roles: tax-related positions include tax collection, tax services, and tax inspection, while non-tax-related positions include administration, human resources, internal finance, and IT. The first two regressions in Panel B are conducted at the county level and control for year fixed effects. The third and fourth regressions are at the position level, controlling for position and year fixed effects, with additional controls for major requirements.

Figure 2: Balance Test of County Characteristics: After Matching



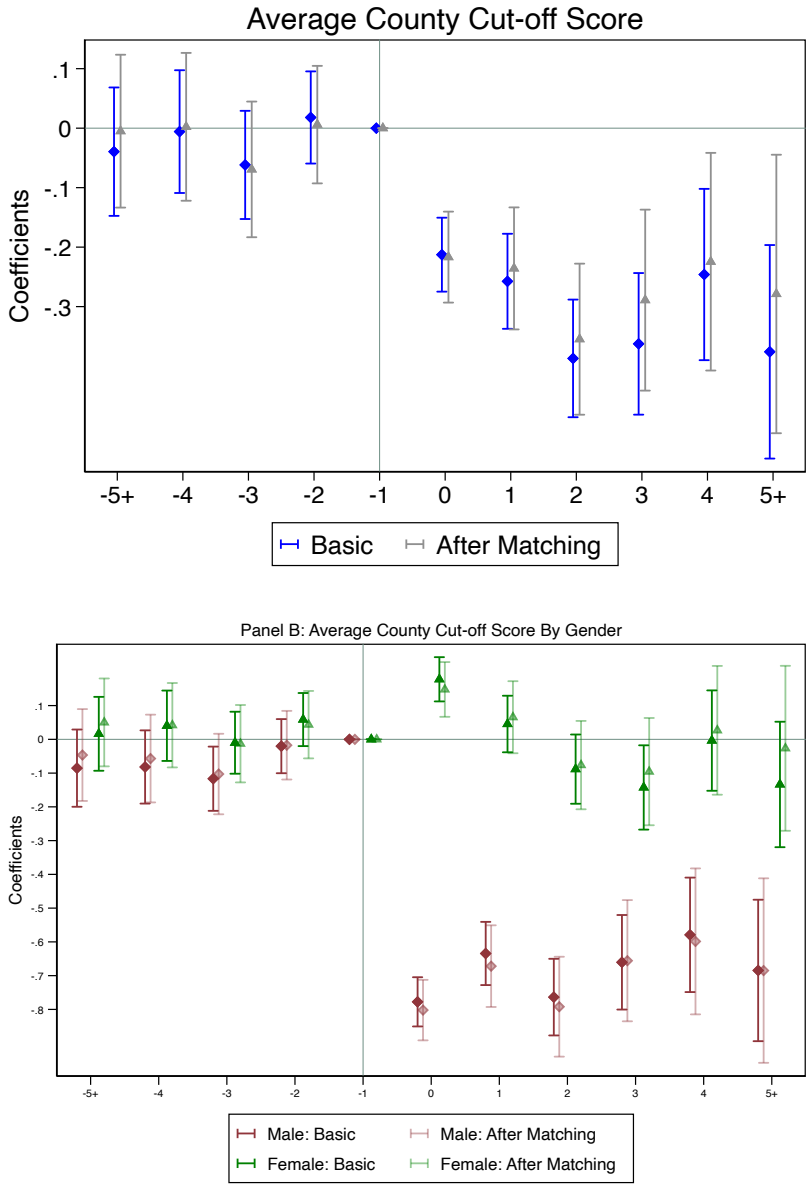
Notes - This figure presents the correlation between pre-treatment county characteristics and quota adoption after matching. The matching successfully balances both the variables used in the matching process—tertiary sector share, population, male ratio, and urban-hukou ratio—and other covariates not directly used for matching. Additionally, I conduct a robustness check for corruption level, which remains slightly unbalanced.

Figure 3: Female Share of New Hires after the Quota: After Matching



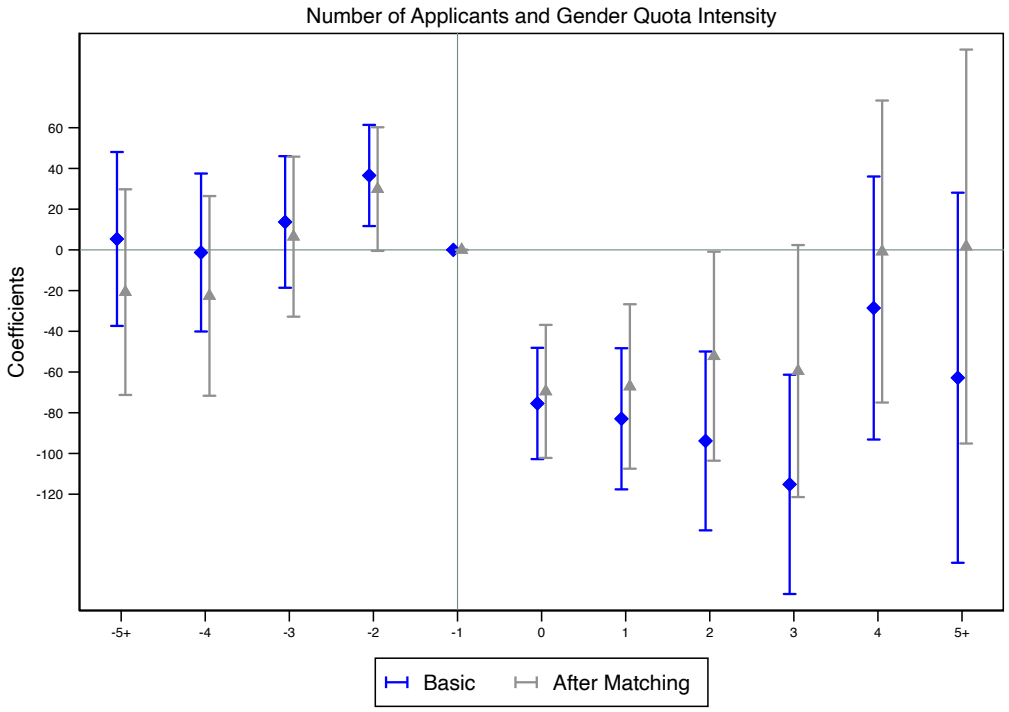
Notes - This figure presents the first-stage results of gender quota intensity, measured as the average share of gender-targeted positions across all treated years. Gender composition is derived from admission data starting in 2016. The *current-year female ratio* refers to the share of female new hires among all new hires in the given year, and the *accumulated female ratio* captures the share of female hires among all new hires from 2016 up to the current year. All ratios are calculated at the county level.

Figure 4: The Effect of Gender Quota Intensity on Average Standardized Cut-off Scores



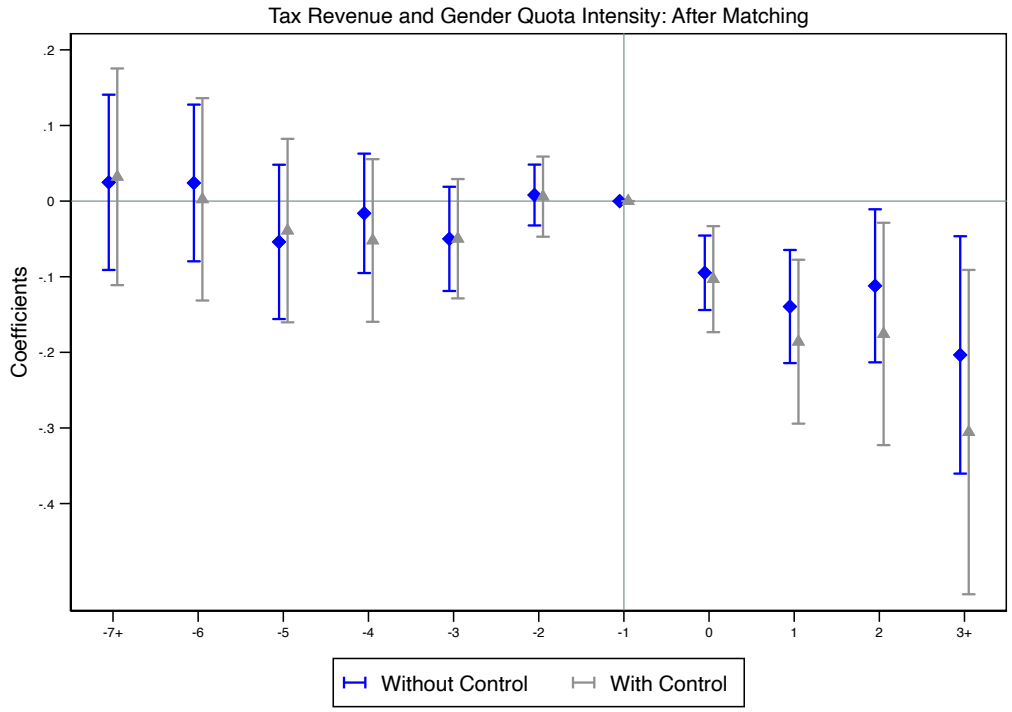
Notes - This figure shows the effect of gender quota intensity on the average county-level cut-off score. Quota intensity is measured by the average share of gender-targeted positions across all treated years. *Panel A* presents the overall effect on cut-off scores, while *Panel B* distinguishes the effects by gender, showing separate impacts for male-targeted and female-targeted positions. Cut-off scores are standardized, and the figure displays event study estimates from both the baseline regressions and the matched sample regressions.

Figure 5: The Effect of Gender Quota on Number of Applicants



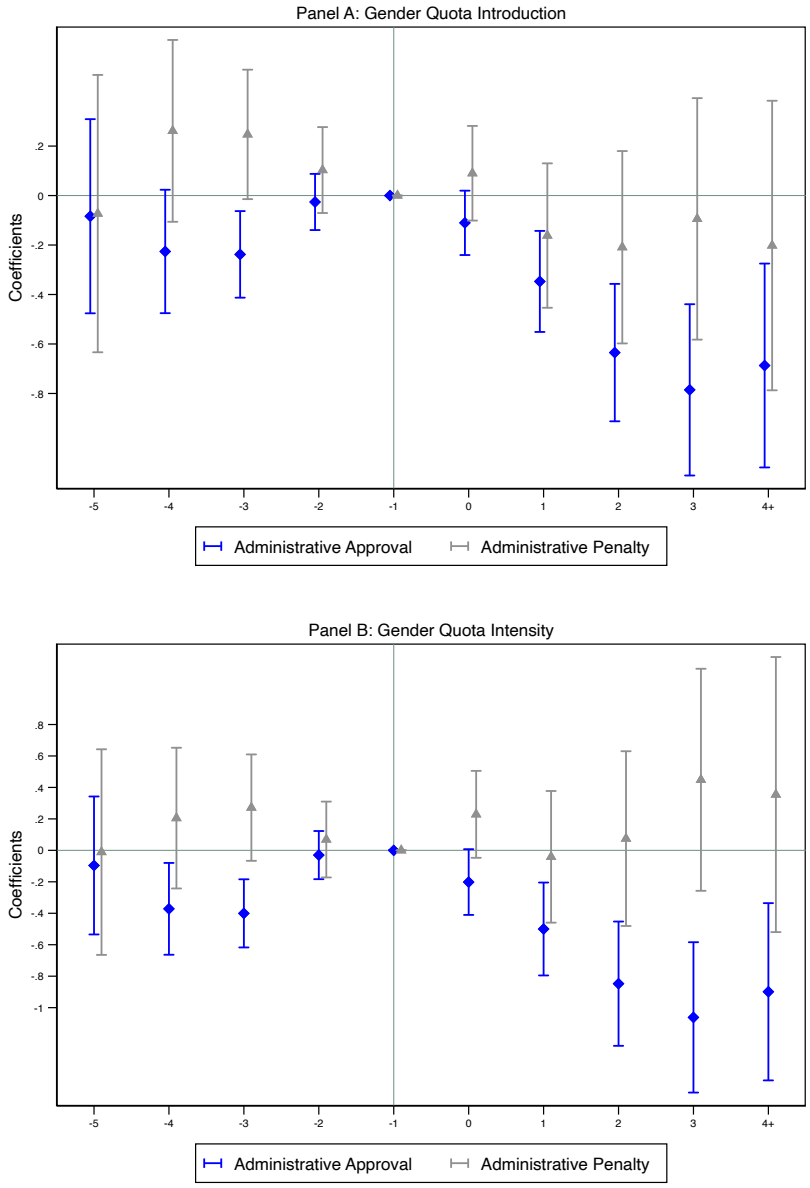
Notes - This figure shows the effect of gender quota intensity on the number of applicants. Quota intensity is measured by the average share of gender-targeted positions across all treated years. The figure displays event study estimates from both the baseline regressions and the matched sample regressions.

Figure 6: The Effect of Gender Quota on Total Revenue



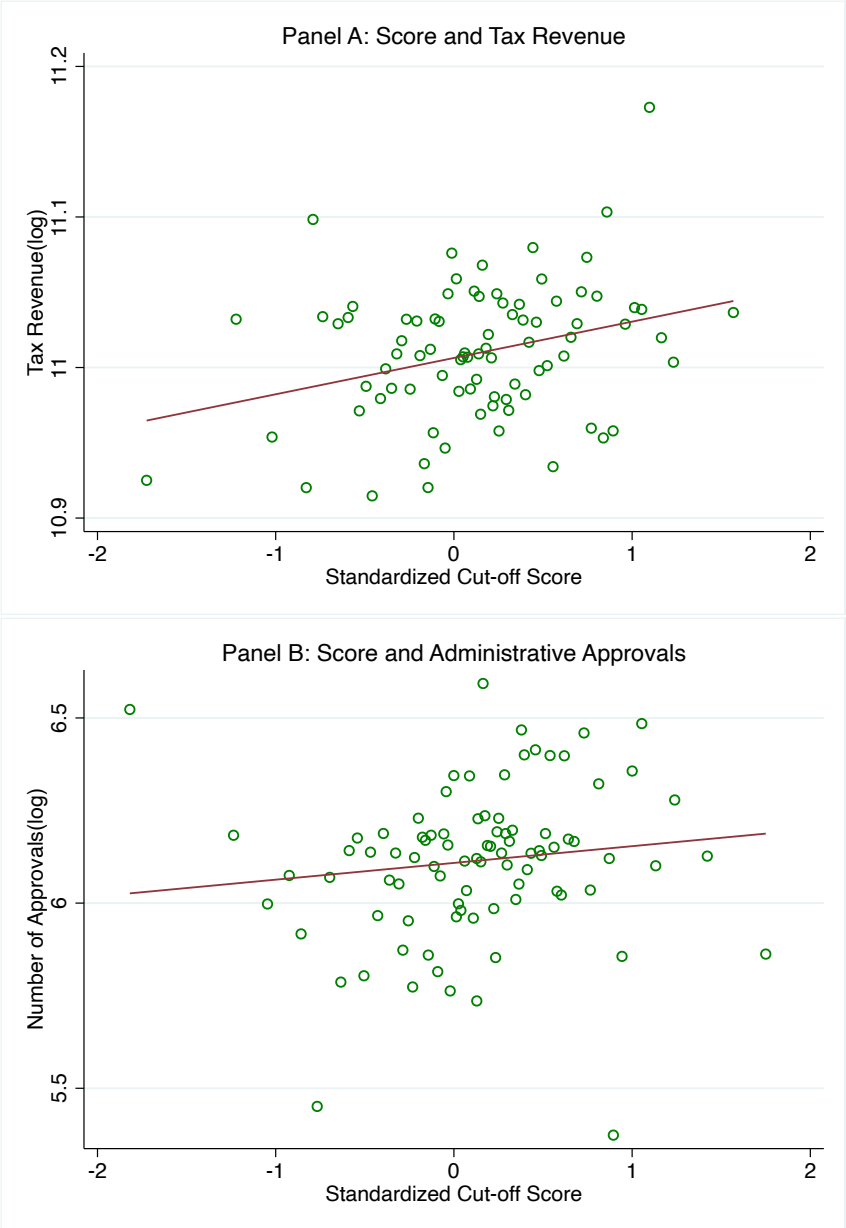
Notes - This figure shows the effect of gender quota intensity on tax revenue. Quota intensity is measured as the average share of gender-targeted positions across all treated years. Tax revenue is measured in 10,000 RMB and presented in logarithmic form. The figure displays event study estimates from regressions both with and without pre-treatment controls, using the matched sample. The controls include tertiary sector share, population size, male ratio, and urban-hukou ratio.

Figure 7: The Effect on Number of Administrative Cases



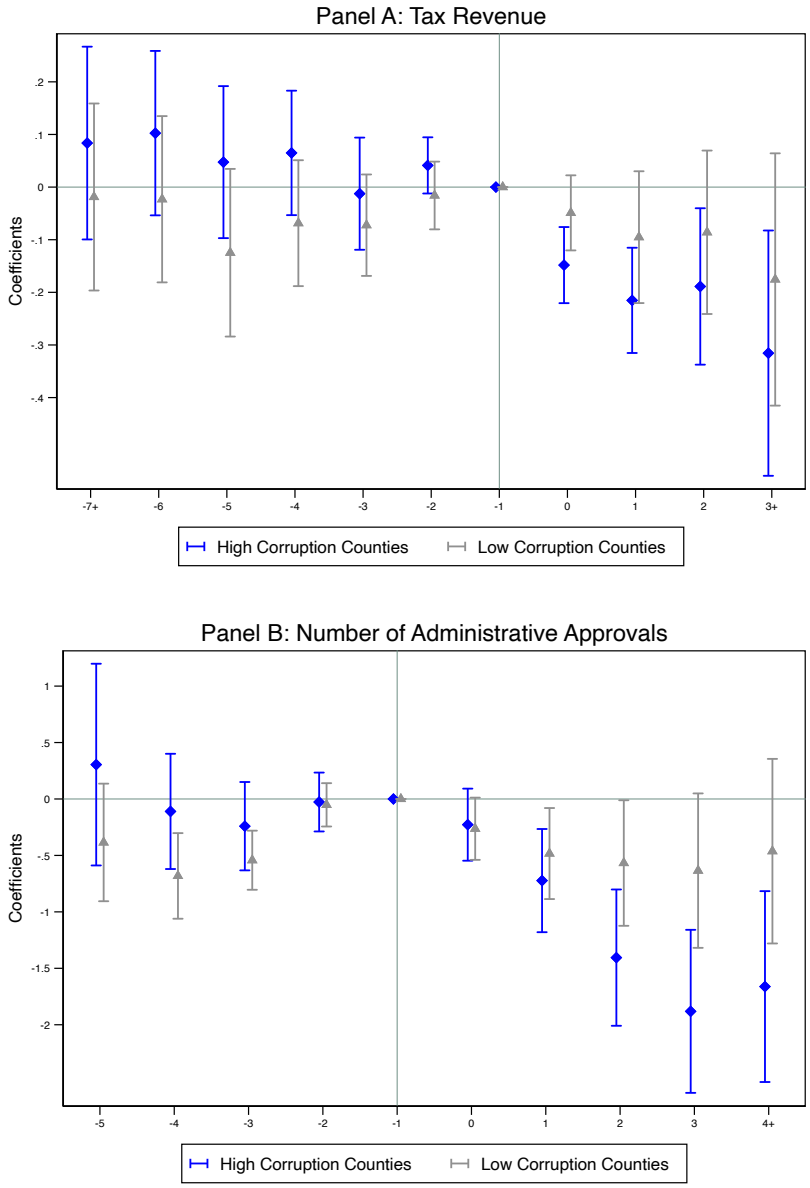
Notes - This figure shows the effect of gender quota intensity on the number of administrative cases. *Panel A* presents the effect of the introduction of gender quotas, and *Panel B* examines the effect of quota intensity, measured as the average share of gender-targeted positions across all treated years. The number of administrative cases is shown in logarithmic form. The figure reports event study estimates from regressions on both administrative approvals and administrative penalties, using the matched sample.

Figure 8: Correlation Between Ability and Productivity



Notes - This figure shows the effect of standardized cut-off score on institutional productivity. *Panel A* presents the correlation between standardized cut-off score and tax revenue, and *Panel B* shows the correlation between standardized cut-off score and number of administrative approvals. Both tax revenue and administrative approvals are expressed in logarithmic form. The correlations are estimated using the matched sample and are based on data from the pre-quota period. The specification follows that of the main regression, controlling for county fixed effects, year fixed effects, and city-by-year trends.

Figure 9: The Effect on Productivity: By Corruption Level



Notes - This figure shows the effect of the gender quota on institutional productivity, conditional on corruption levels. *Panel A* presents the effect on tax revenue, while *Panel B* shows the effect on administrative approvals. Both outcomes are expressed in logarithmic form. Corruption levels are classified based on the median number of corruption cases from 2011 to 2015, using data from Wang (2020). The median number of corruption cases is 3.

Table 1: The Effect of Gender Quota on Total Tax Revenue:After Matching

	(1) Total Tax(log)	(2) Total Tax(log)	(3) Total Tax(log)	(4) Total Tax(log)
Mean Share of Quota-position	-0.085**		-0.100**	
	(0.036)		(0.047)	
Ever Treated		-0.033		-0.041
		(0.021)		(0.027)
Observations	5069	5069	2919	2919
Mean	65447	65447	46444	46444
r2	0.967	0.967	0.973	0.973
PreTreatControl_year			X	X
Year_FE	X	X	X	X
County_FE	X	X	X	X
City_Year_Trend	X	X	X	X
Cluster_County	X	X	X	X

Notes - This table reports the effect of the gender quota on tax revenue using the Coarsened Exact Matching (CEM) method. The matching variables include tertiary sector share, population, male ratio, and urban-hukou ratio. Columns (1) and (2) present results based on the main specification, using county-year as the unit of observation. Columns (3) and (4) add interactions between pre-treatment county characteristics and year fixed effects. The pre-treatment controls include tertiary sector share, population, male ratio, and urban-hukou ratio. The dependent variable in all regressions is the log of county-level tax revenue (in 10,000 RMB). The key explanatory variables are the average share of gender-targeted positions across treated years and an indicator for whether the quota was adopted. All regressions include county fixed effects, year fixed effects, and city-by-year trends. Standard errors, reported in parentheses, are clustered at the county level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 2: The Effect of Gender Quota on Position Cutoff Score: By different Gender and Position Type

	(1) Male	(2) Female	(3) Male-Tax-related	(4) Female-Tax-related	(5) Male-Non-Tax-related	(6) Female-Non-Tax-related
Ever treated	-0.453***	0.057***	-0.521***	-0.025	-0.368**	0.020
	(0.020)	(0.018)	(0.026)	(0.023)	(0.143)	(0.092)
Observations	23375	23541	13853	13949	1101	1101
Mean	-.139	.029	-.111	.052	.202	.26
r2	0.761	0.799	0.757	0.794	0.803	0.813
Year_FE	X	X	X	X	X	X
Unique_Posit_FE	X	X	X	X	X	X
Cluster_Position	X	X	X	X	X	X

Notes - This table reports the effect of gender quotas on standardized cut-off scores at the position level. Columns (1) and (2) present results for all positions, using position-year as the unit of observation. Columns (3) and (4) focus on tax-related positions, while Columns (5) and (6) report results for non-tax-related positions, such as administrative roles. Odd-numbered columns show the change in cut-off scores when a position shifts from gender-neutral to male-targeted, while even-numbered columns show the change when it shifts from gender-neutral to female-targeted. The dependent variable in all regressions is the standardized cut-off score for a given position-year. The main explanatory variable is an indicator for whether the position has ever been treated under the gender quota. All regressions include year fixed effects and unique position fixed effects. A unique position is defined as one within a county that shares identical criteria—education, work experience, language requirements, Hukou status, and any other special conditions. Standard errors, shown in parentheses, are corrected for clustering at the county level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

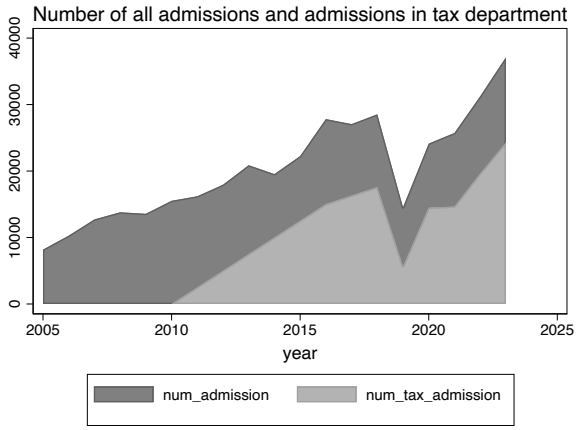
Table 3: The Effect of Quota on Tax Revenue : IV Strategy

	(1) Basic	(2) Basic-First	(3) Basic-Second	(4) Hire Even	(5) Hire Even-First	(6) Hire Even-Second
Mean Share of Quota-position	-0.055*		-0.138**	-0.101**		-0.174**
	(0.028)		(0.056)	(0.041)		(0.075)
Share of years with even number vacancies		0.630***			0.692***	
		(0.032)			(0.041)	
Observations	8188	8188	8188	3251	3251	3251
Mean	77964	.104	77964	77964	.153	77964
r2	0.961	0.819	0.199	0.962	0.879	0.236
F-statistics		386.93			288.11	
Year_FE	X	X	X	X	X	X
County_FE	X	X	X	X	X	X
City_Year_Trend	X	X	X	X	X	X
Cluster_County	X	X	X	X	X	X

Notes - This table reports the effect of gender quotas on tax revenue using an alternative empirical strategy—2SLS regression. Column (1) presents the OLS results without matching, following the main specification, with the county-year as the unit of observation. Column (2) reports the first-stage regression, showing that the share of years with an even number of vacancies is positively associated with gender quota intensity. Column (3) presents the corresponding second-stage 2SLS results. Columns (4) to (6) restrict the sample to counties that hired only in years with an even number of vacancies. Column (4) reports OLS results, while Columns (5) and (6) present the first and second stages of the 2SLS, respectively. The outcome variable in all regressions is county-level tax revenue in a given year, and the main explanatory variable is the average share of gender-targeted positions across treated years. The instrumental variable is the share of years with an even number of vacancies relative to all treated years. All regressions control for the fixed effects included in Column (1) of Table 1. Standard errors, shown in parentheses, are corrected for clustering at the county level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

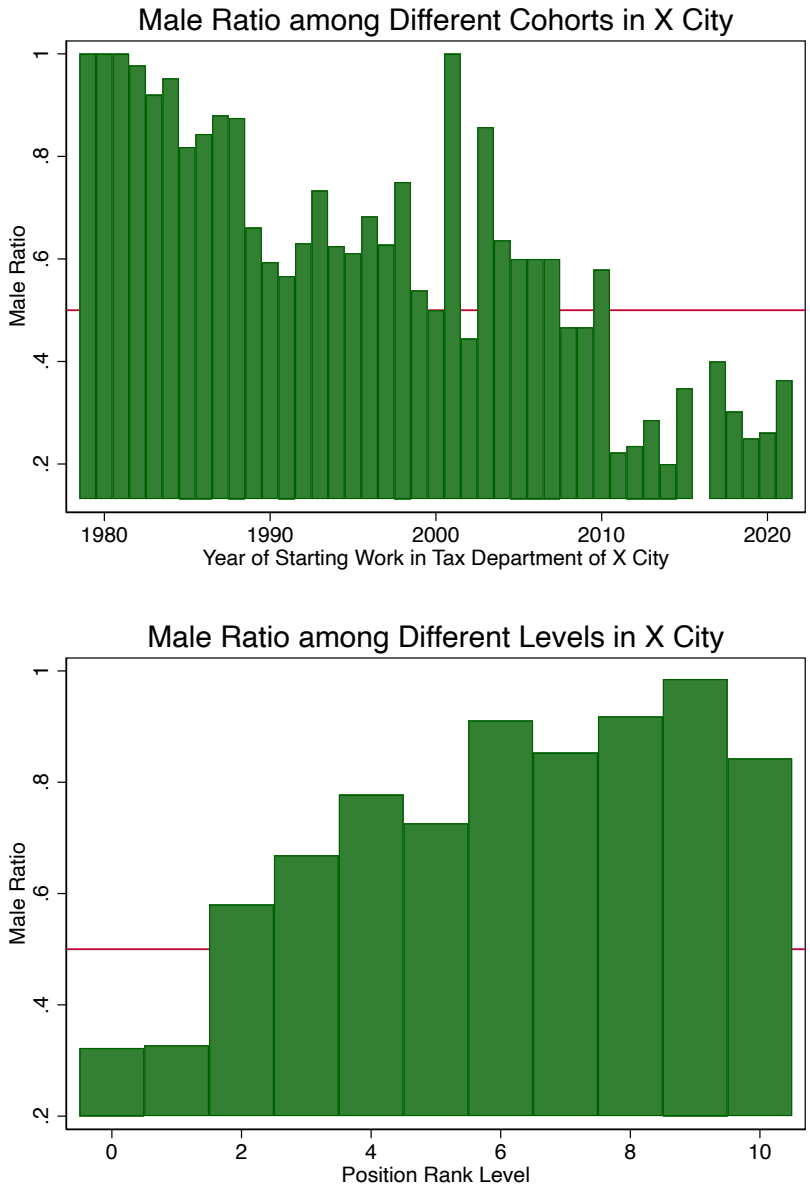
A Appendix: Supplemental Figures and Tables

Figure A.1: Share of Tax Department's Recruitment



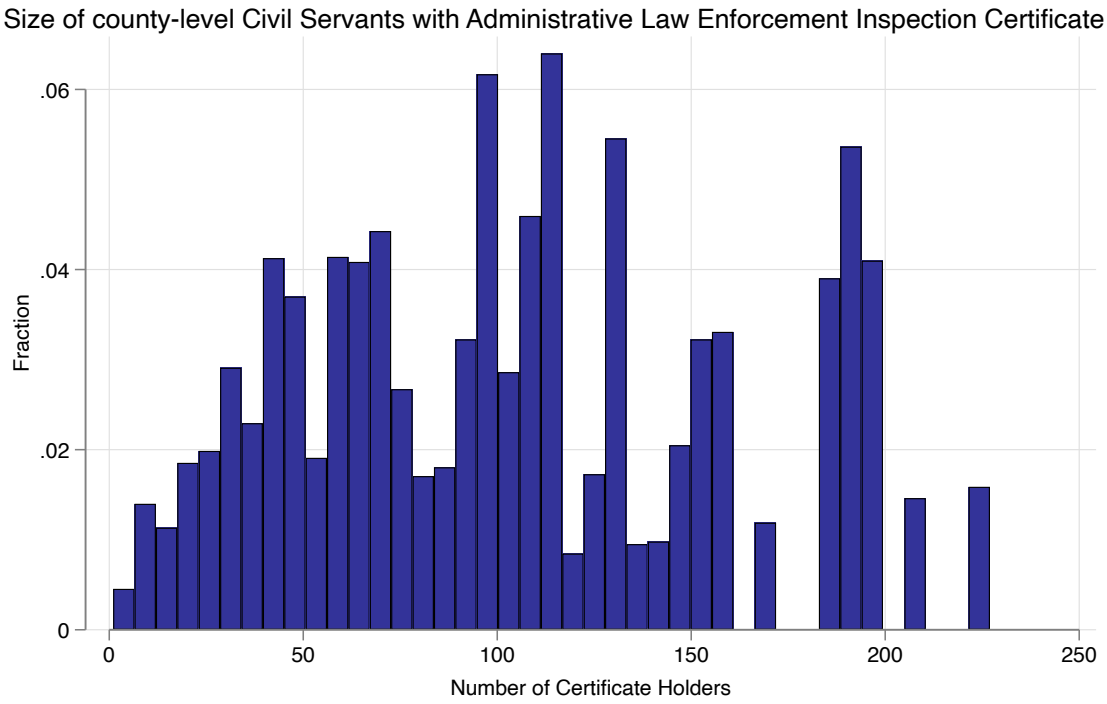
Notes - This figure shows the proportion of tax department vacancies relative to all vacancies in the National Civil Service Examination.

Figure A.2: Male Representation in Tax Department



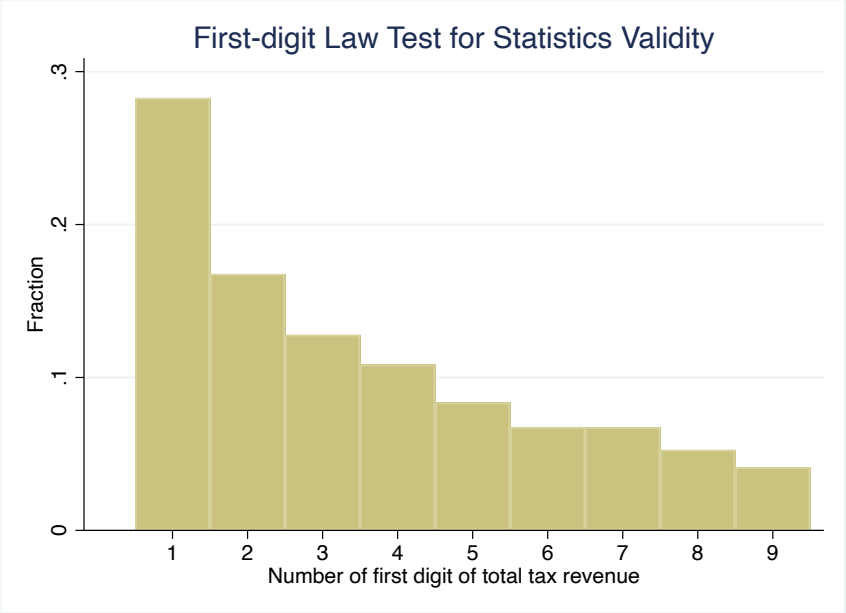
Notes - This figure presents male representation in both overall numbers and leadership roles. The left panel shows the male ratio across cohorts defined by the year of entry into the workforce. The right panel displays the male ratio across different position levels, where a higher level indicates a higher rank or leadership role within the tax department. The data are confidential and obtained from City X.

Figure A.3: Size of Administrative Law Enforcement Inspection Certificate Holders in Province Anhui



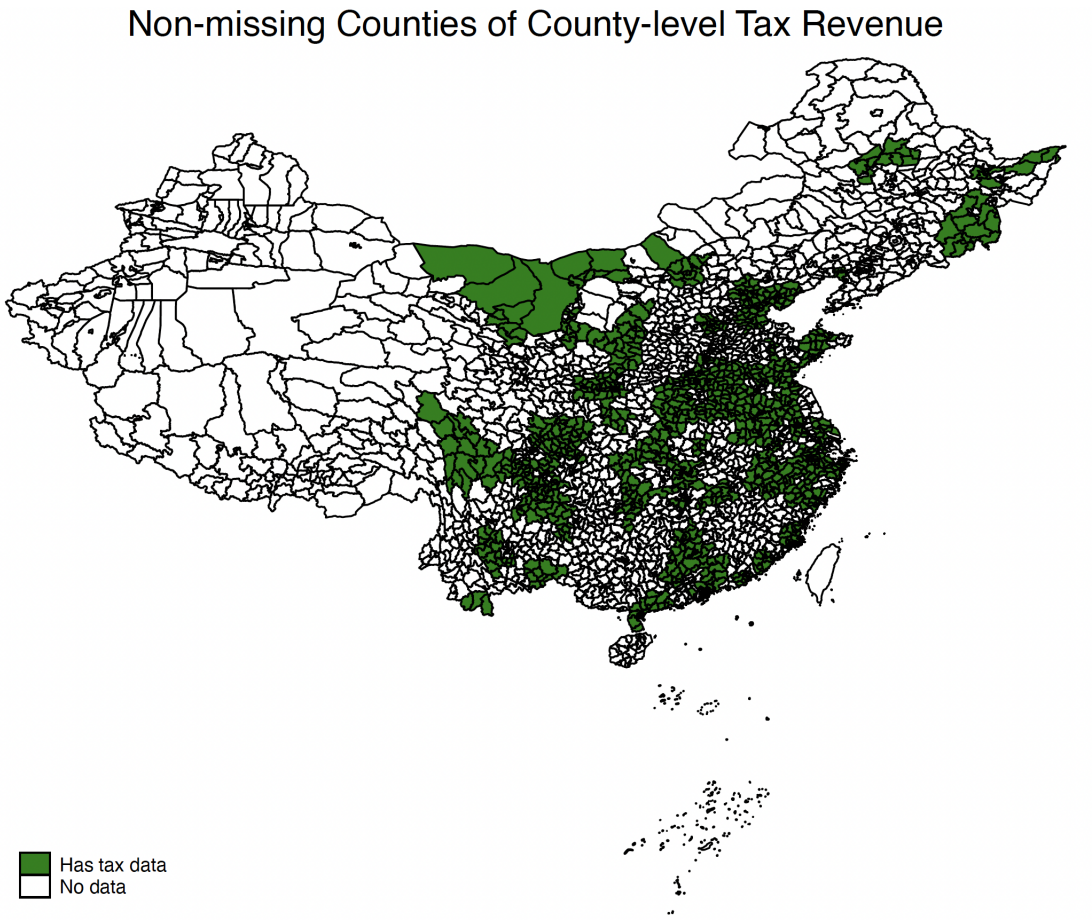
Notes - This figure shows the number of administrative certificate holders across counties in Anhui Province.

Figure A.4: Validity of Tax Revenue Data



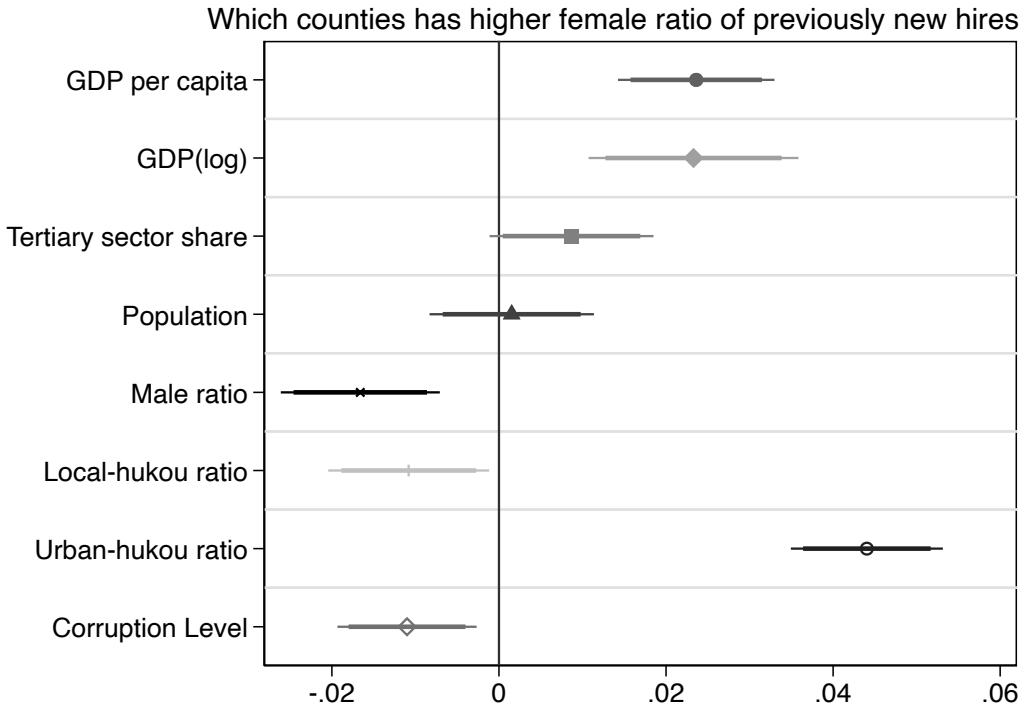
Notes - This figure shows the data validity using the First-digit Law Test.

Figure A.5: Counties with Tax Data in City Year Books



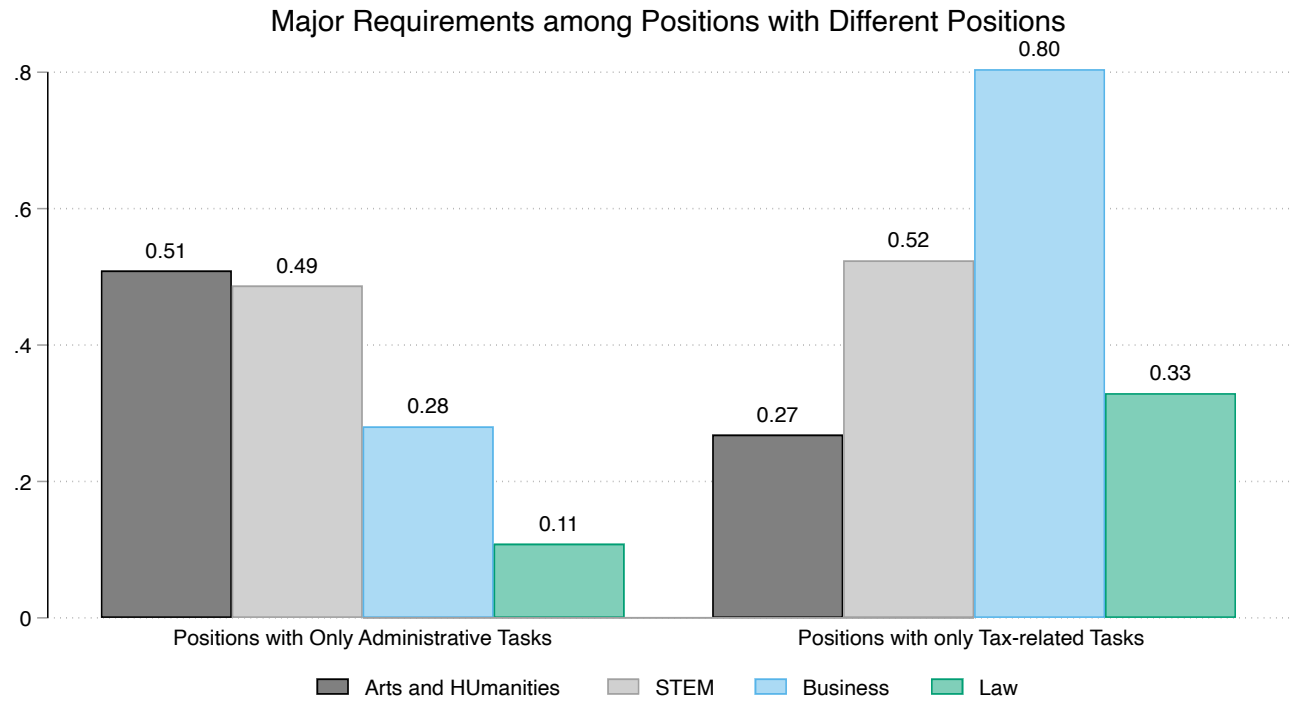
Notes - This figure displays the counties for which tax data are available from the China Statistical Yearbooks.

Figure A.6: Correlation Between Prior Female Hiring Shares and County Characteristics



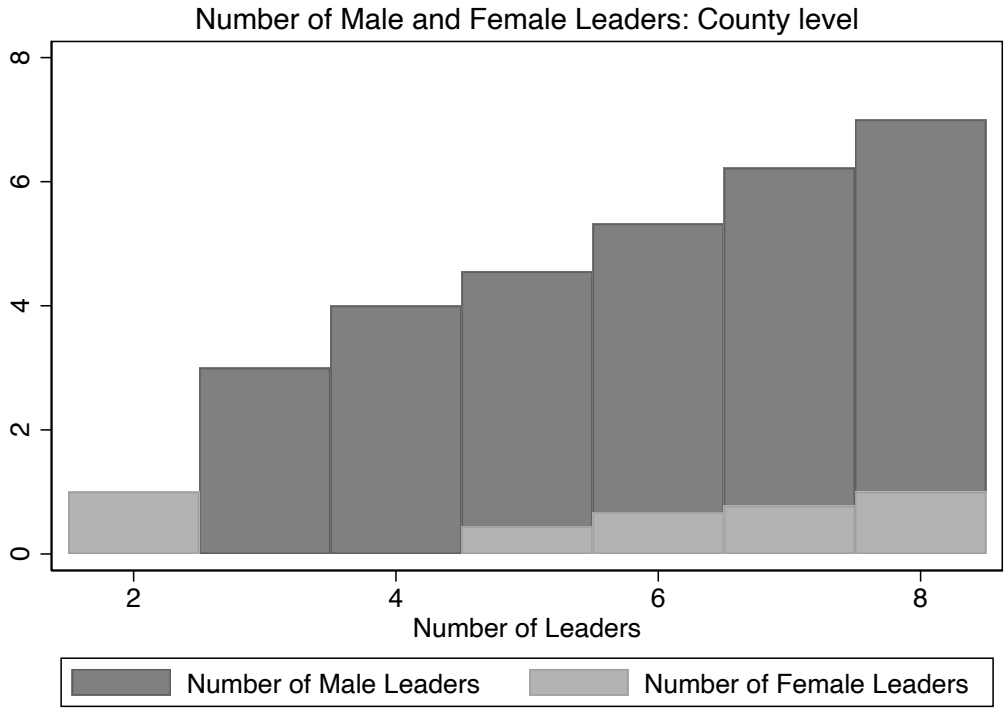
Notes - This figure shows the correlation between county characteristics and the female share of recent new hires. County characteristics data are from Census 2010 reported in county level. Corruption level is from China's Corruption Investigations Dataset(Wang, 2020). Year fixed effect and province fixed effect are controlled in each regression. The female ratio of recent new hires is restricted to pre-adoption years for treatment group.

Figure A.7: Position categories and major requirements



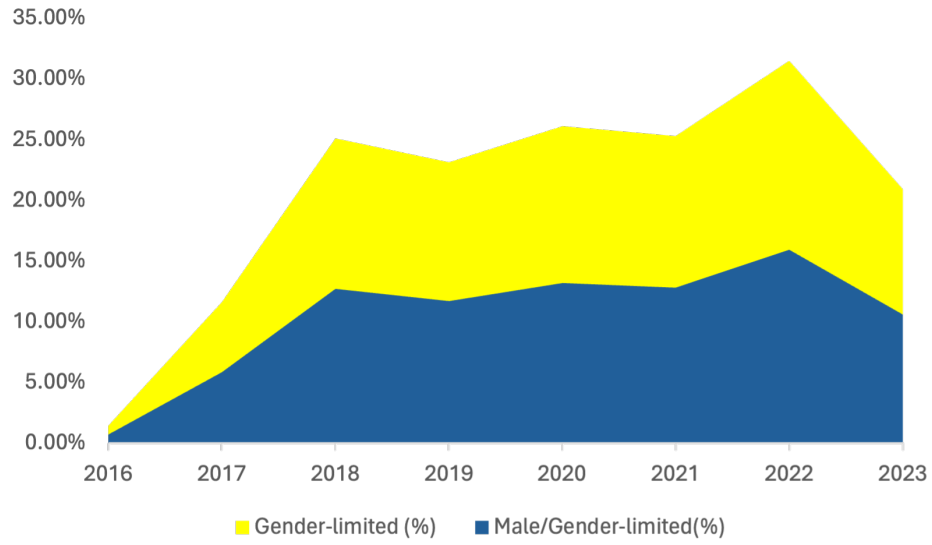
Notes - This figure displays the distribution of major requirements across different positions.

Figure A.8: Gender Composition of Leaders of County-level Taxation Bureau



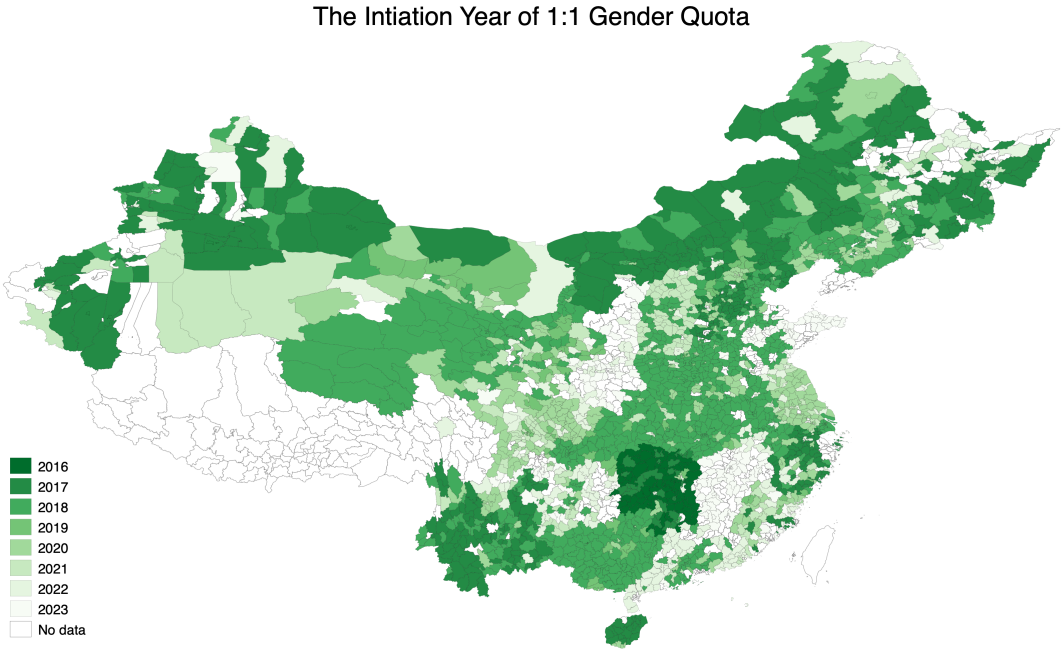
Notes - This figure shows the gender composition of leaders of county tax department. The data are web-scraped from the official websites of county-level taxation bureaus in four provinces: Anhui, Guangdong, Guizhou, and Zhejiang.

Figure A.9: Share of Gender Quota Positions in Each County



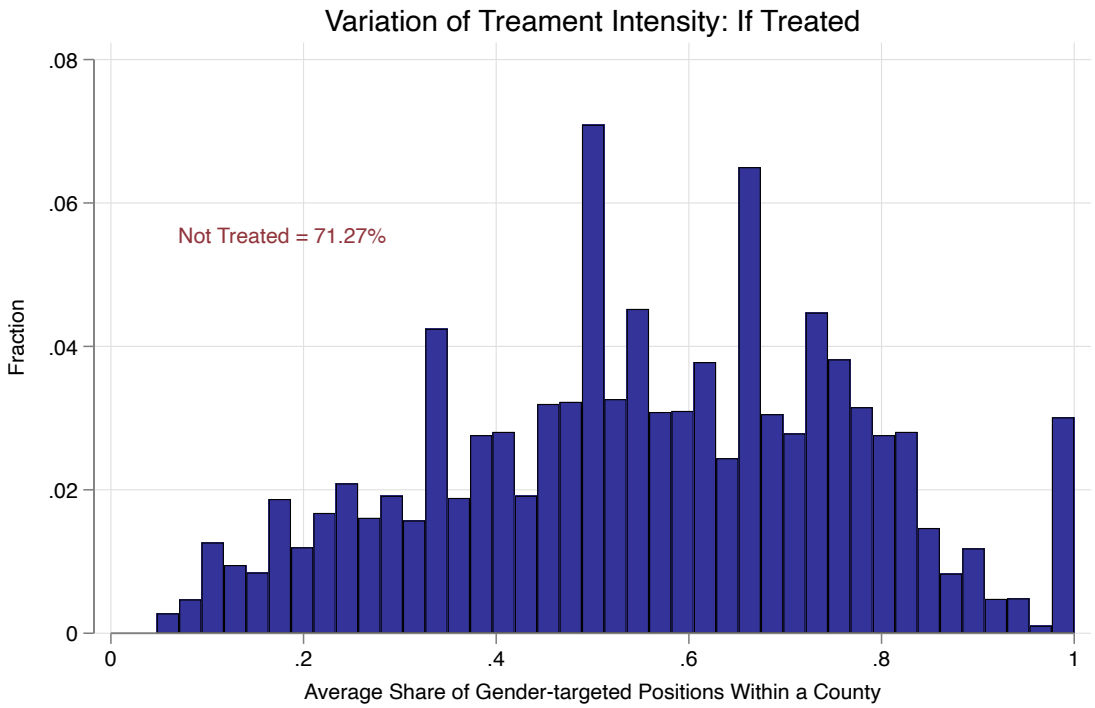
Notes - This figure shows the share of gender-targeted positions. The "Male/Gender-limited" ratio refers to the proportion of male-targeted positions among all gender-targeted positions. Male-targeted positions consistently account for 50% of gender-targeted positions.

Figure A.10: Gender Quota Enrollment Across Countries



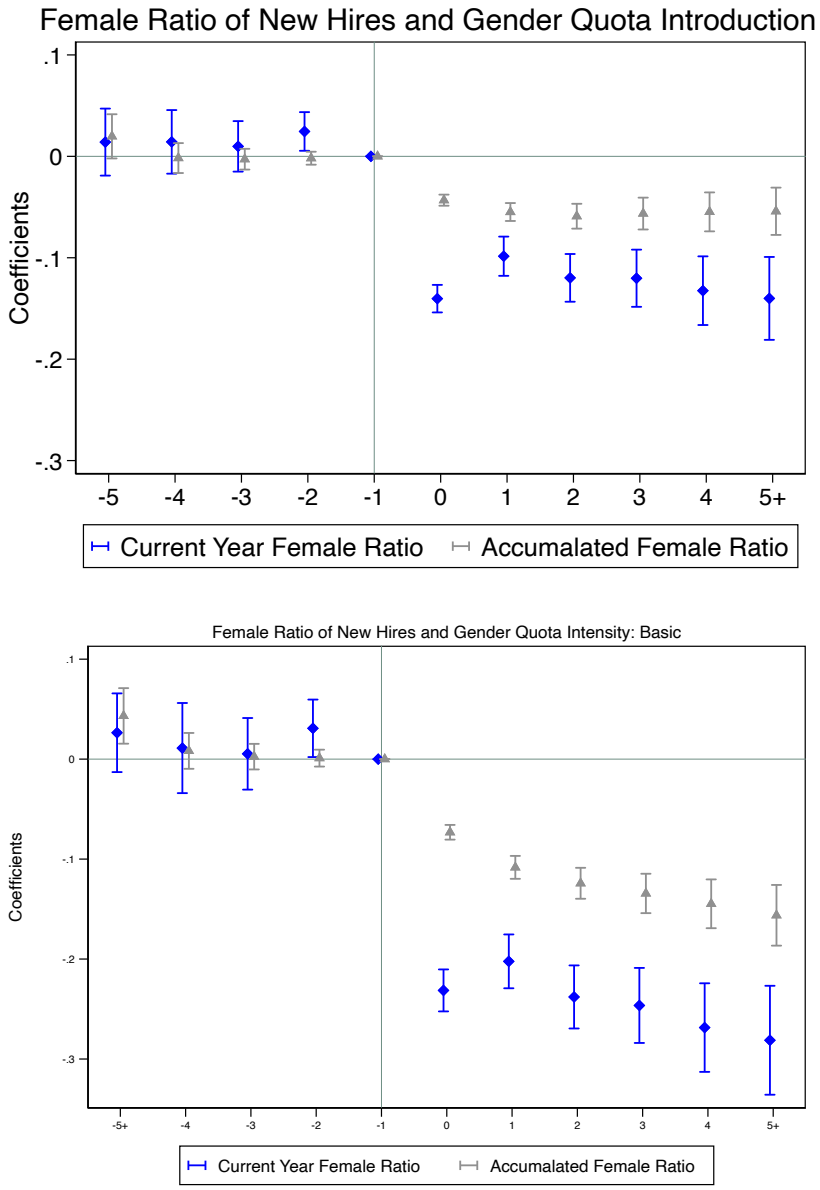
Notes - This figure shows the year in which the gender quota was first implemented across counties in China. The initiation year is defined as the first year in which at least one position in a county was designated as gender-targeted.

Figure A.11: Variation of Continuous Treatment: If positive



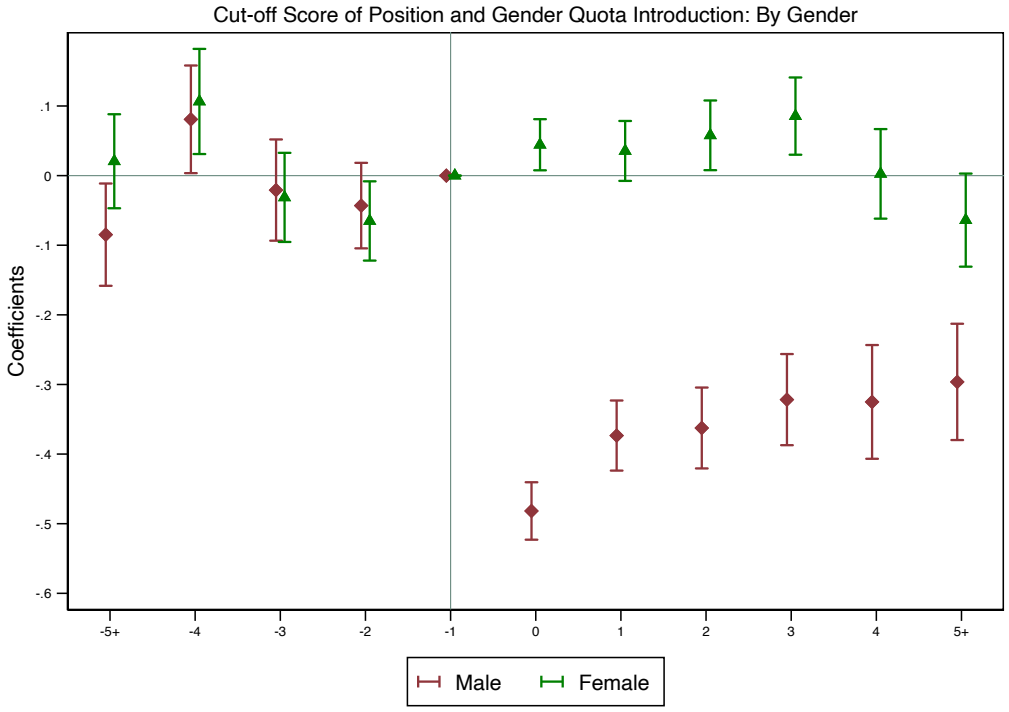
Notes - This figure shows the variation in the average share of gender-targeted positions across counties. Counties that did not implement the policy are excluded.

Figure A.12: Event Study of Female Ratio after the Quota Introduction(Basic)



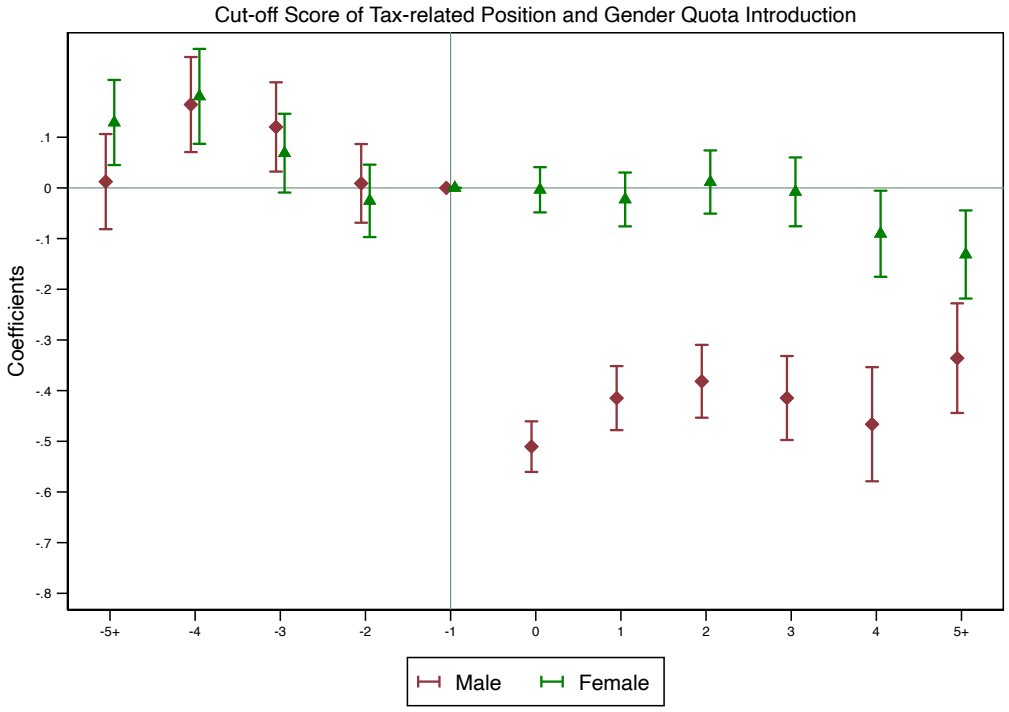
Notes - This figure presents the first-stage results of gender quota intensity without matching, measured as the average share of gender-targeted positions across all treated years. Gender composition is derived from admission data starting in 2016. The *current-year female ratio* refers to the share of female new hires among all new hires in the given year, and the *accumulated female ratio* captures the share of female hires among all new hires from 2016 up to the current year. All ratios are calculated at the county level.

Figure A.13: The Effect of Gender Quota on Position-level Standardized Cut-off Scores: By Gender



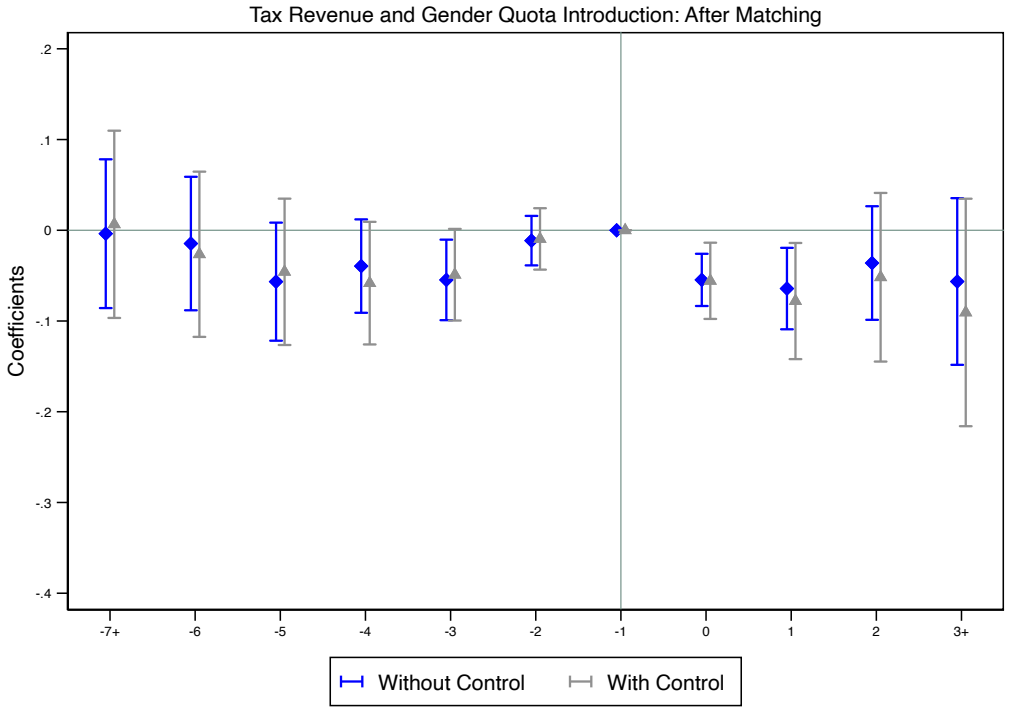
Notes - This figure shows the effect of gender quota introduction on the position-level cut-off score by gender, showing separate impacts for male-targeted and female-targeted positions. Cut-off scores are standardized. A unique position is defined as one that shares the same county and identical criteria—education, work experience, language skills, Hukou status, and any other special requirements.

Figure A.14: The Effect of Gender Quota on Tax-related Position Standardized Cut-off Scores: By Gender



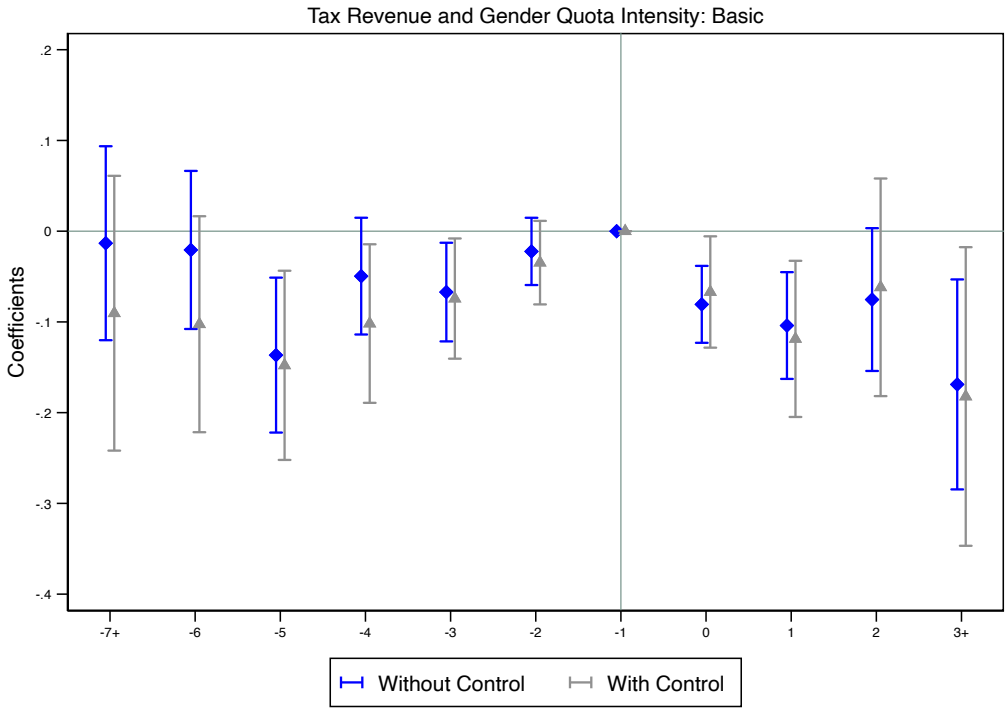
Notes - This figure shows the effect of gender quota introduction on the tax-related positions' cut-off score by gender, showing separate impacts for male-targeted and female-targeted positions. Cut-off scores are standardized. A unique position is defined as one that shares the same county and identical criteria—education, work experience, language skills, Hukou status, and any other special requirements.

Figure A.15: The Effect on Total Revenue: Ever Treated



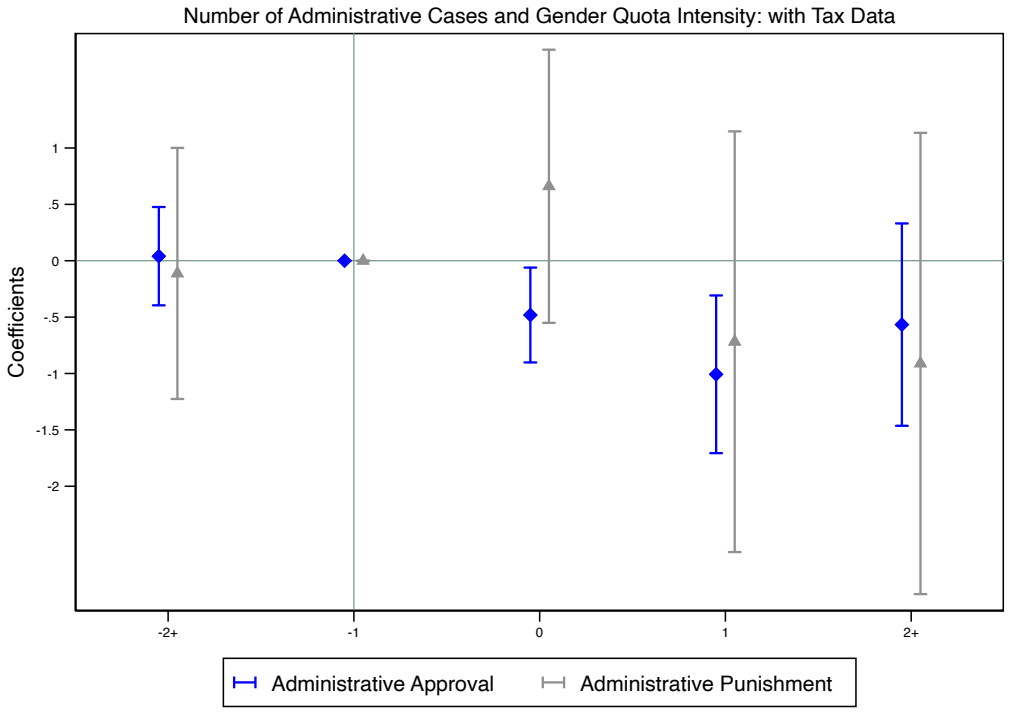
Notes - This figure shows the effect of gender quota introducation on tax revenue. Tax revenue is measured in 10,000 RMB and presented in logarithmic form. The figure displays event study estimates from regressions both with and without pre-treatment controls, using the matched sample. The controls include tertiary sector share, population size, male ratio, and urban-hukou ratio.

Figure A.16: The Effect on Total Revenue: No Matching



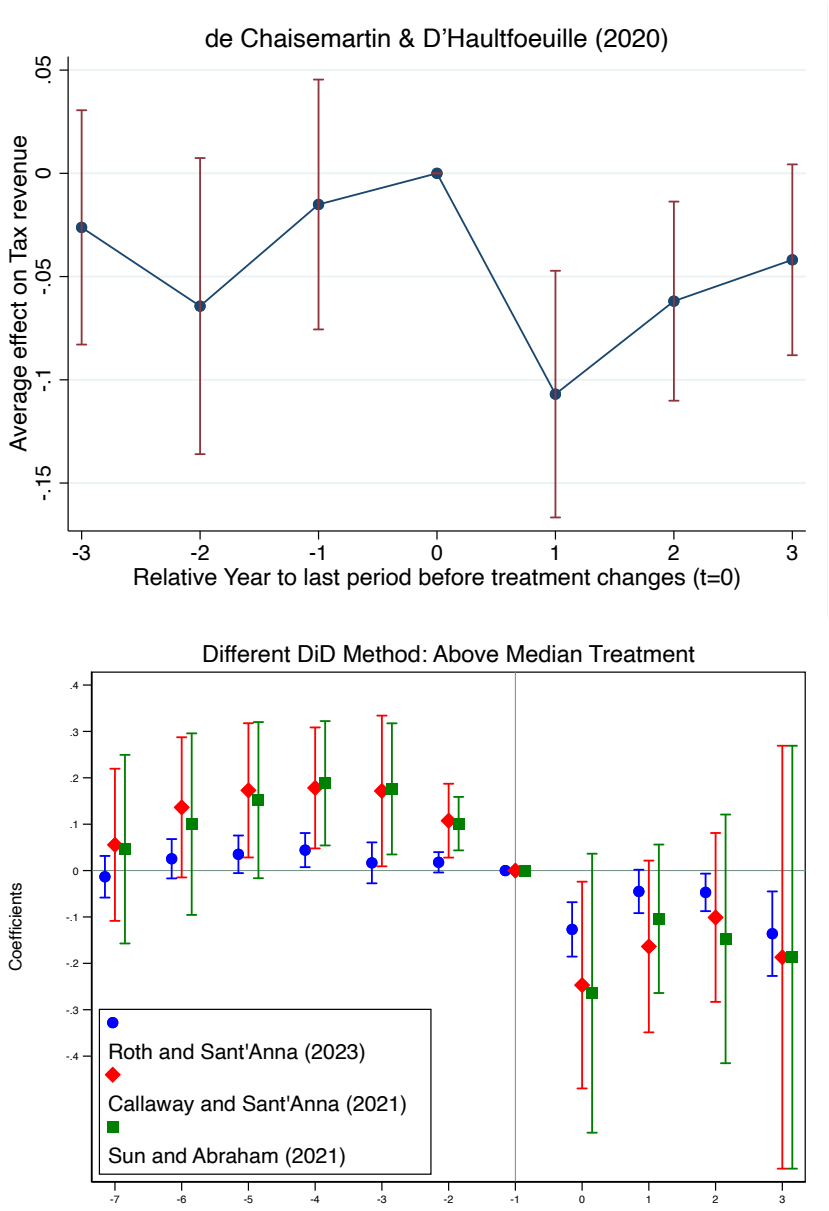
Notes - This figure shows the effect of gender quota intensity on tax revenue without matching. Quota intensity is measured as the average share of gender-targeted positions across all treated years. Tax revenue is measured in 10,000 RMB and presented in logarithmic form. The figure displays event study estimates from regressions both with and without pre-treatment controls, using the matched sample. The controls include tertiary sector share, population size, male ratio, and urban-hukou ratio.

Figure A.17: The Effect on Number of Administrative Cases: Within Main Sample



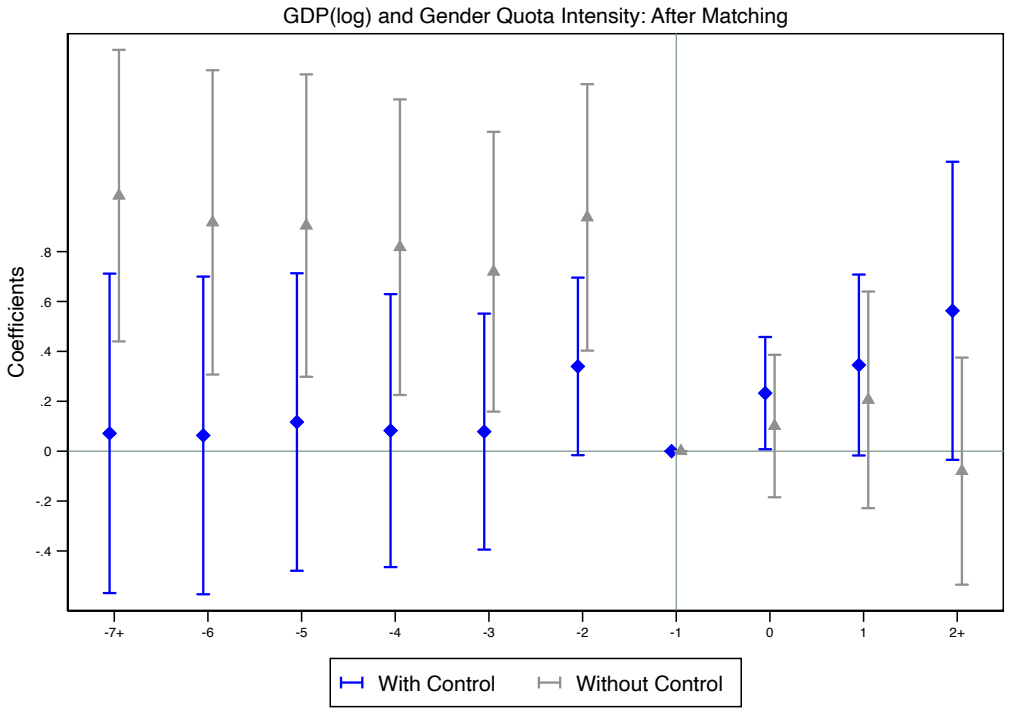
Notes - This figure shows the effect of gender quota intensity on the number of administrative cases within the sample whose tax data is available. Quota intensity is measured as the average share of gender-targeted positions across all treated years. The number of administrative cases is shown in logarithmic form. The figure reports event study estimates from regressions on both administrative approvals and administrative penalties, using the matched sample.

Figure A.18: Robustness Check using Different Staggered DiD Estimators



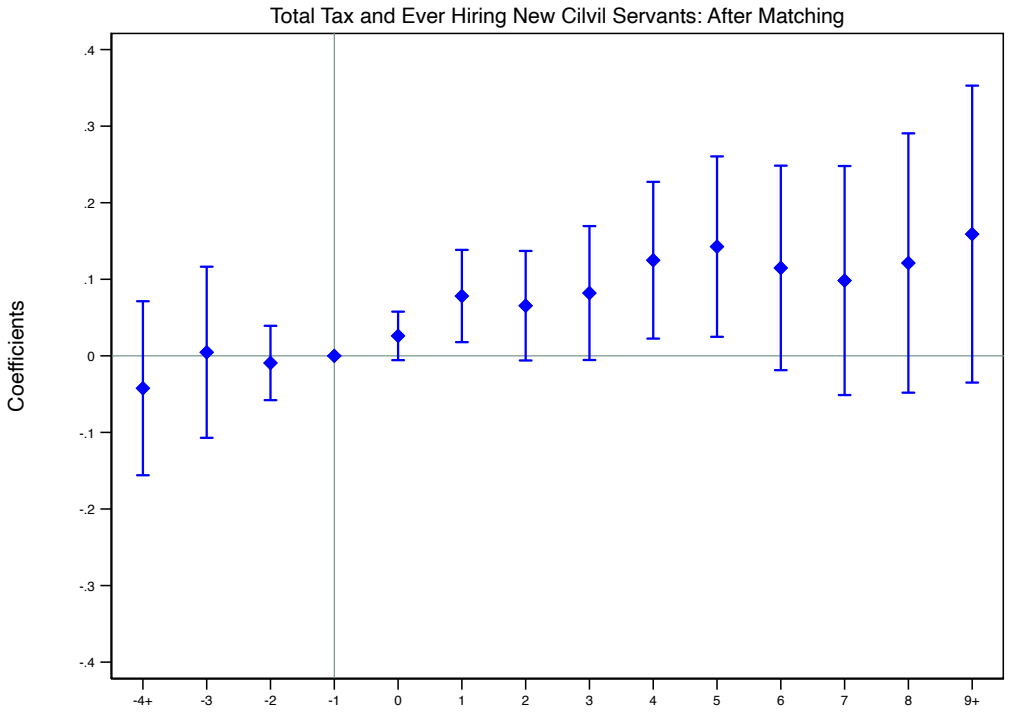
Notes - This figure overlays event-study plots based on four different estimators to examine the effect of continuous treatment intensity on tax revenue. The top panel presents results using the estimator from De Chaisemartin and d'Haultfoeuille (2020), which accommodates continuous treatment. The remaining three estimators—Roth and Sant'Anna (2023) (blue line with circle markers), Callaway and Sant'Anna (2021) (red line with diamond markers), and Sun and Abraham (2021) (green line with square markers)—require binary treatment status. To apply these methods, I define treated counties as those with a gender-targeted position share above the median (55%) and drop the remaining treated units. The outcome variable is county-level tax revenue, measured in 10,000 RMB and presented in logarithmic form. I control for city-by-year trends and use the matched sample only in De Chaisemartin and d'Haultfoeuille (2020), where applicable.

Figure A.19: The Effect on GDP(Log)



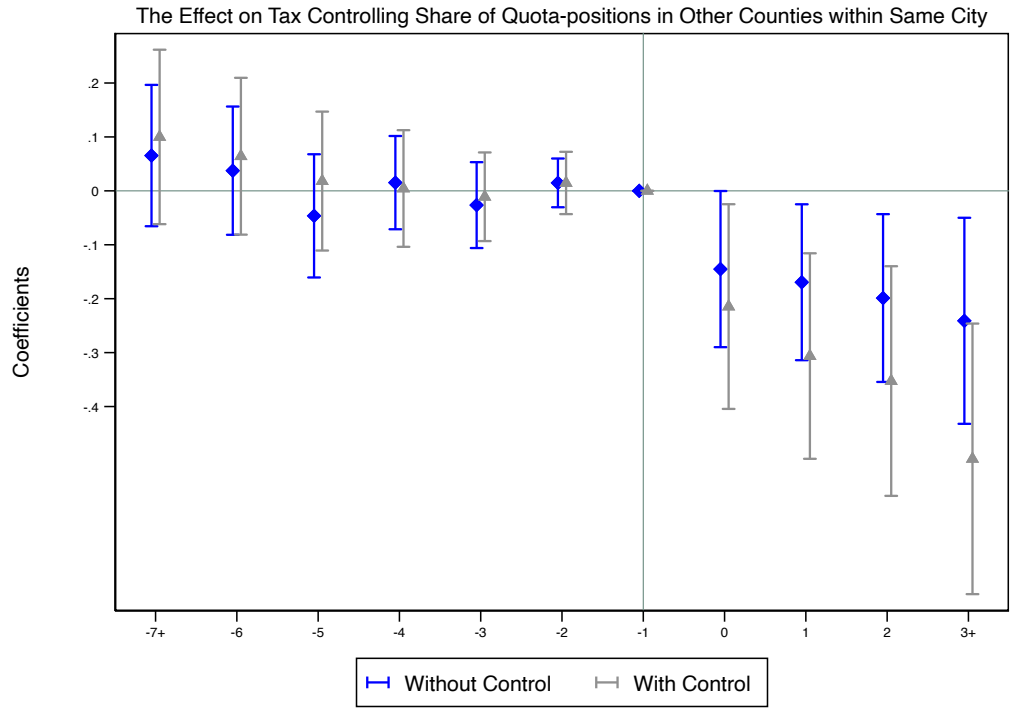
Notes - This figure shows the effect of gender quota intensity on county-level GDP. Quota intensity is measured as the average share of gender-targeted positions across all treated years. GDP is measured in 10,000 RMB and presented in logarithmic form. The figure displays event study estimates from regressions both with and without pre-treatment controls, using the matched sample. The controls include tertiary sector share, population size, male ratio, and urban-hukou ratio.

Figure A.20: The Effect of Ever Hiring on Tax Revenue



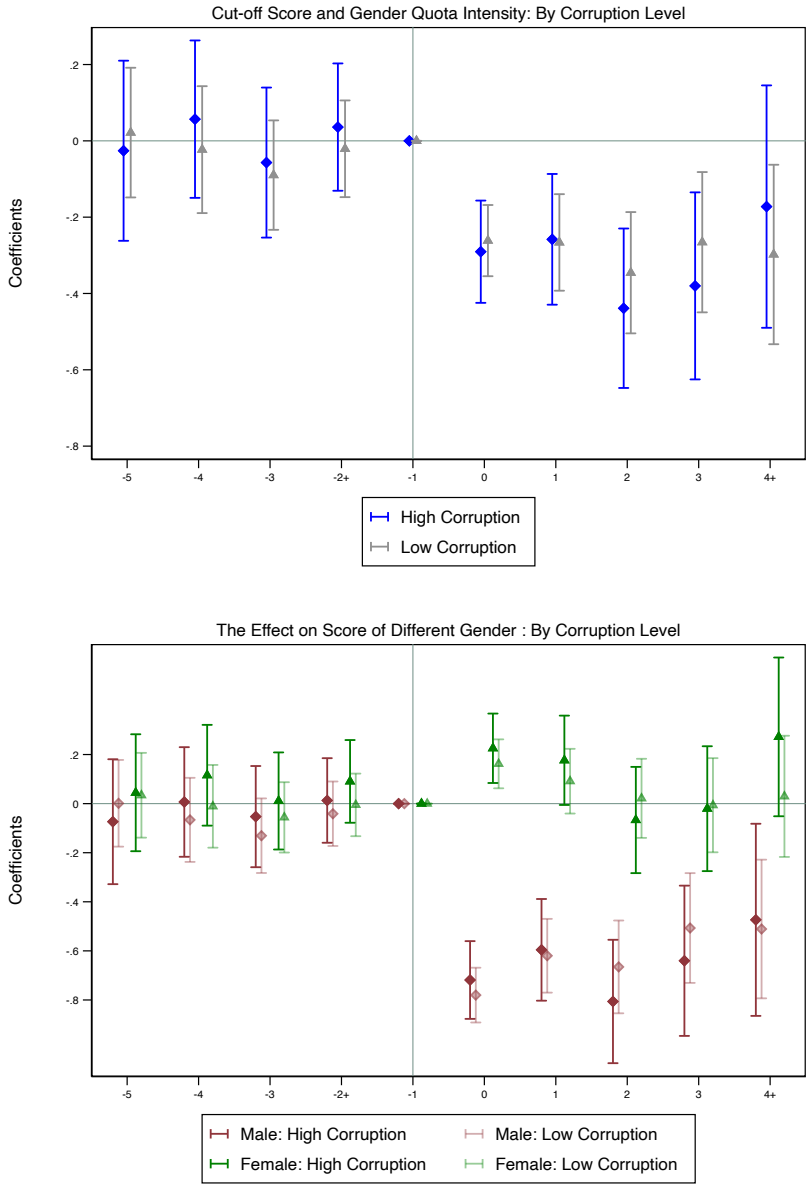
Notes - This figure shows the effect of ever hiring on tax revenue. "Ever hiring" is defined as an indicator equal to 1 if a county has hired at least once from year t onward. Tax revenue is measured in 10,000 RMB and presented in logarithmic form. All other specifications follow the main regression model.

Figure A.21: The Effect on Tax Controlling Spillover Effect



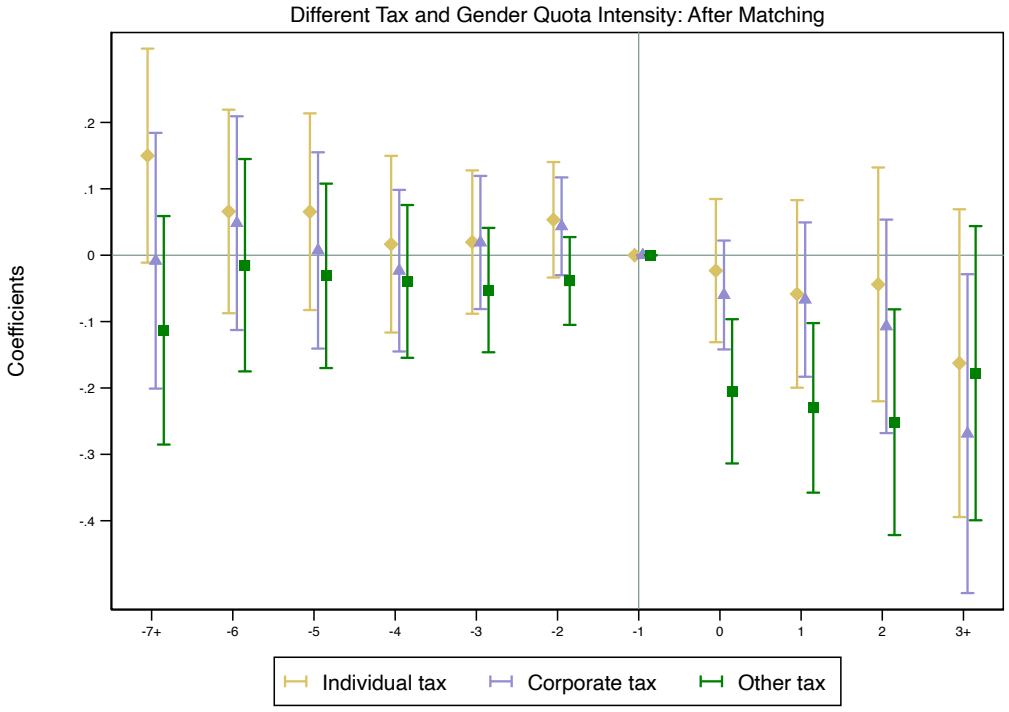
Notes - This figure shows the effect of gender quota intensity on tax revenue controlling spillover effect. Quota intensity is measured as the average share of gender-targeted positions across all treated years. Tax revenue is measured in 10,000 RMB and presented in logarithmic form. The figure displays event study estimates from regressions both with and without pre-treatment controls, using the matched sample. The controls include tertiary sector share, population size, male ratio, and urban-hukou ratio.

Figure A.22: The Effect of Quota Intensity on Score: By Corruption Level



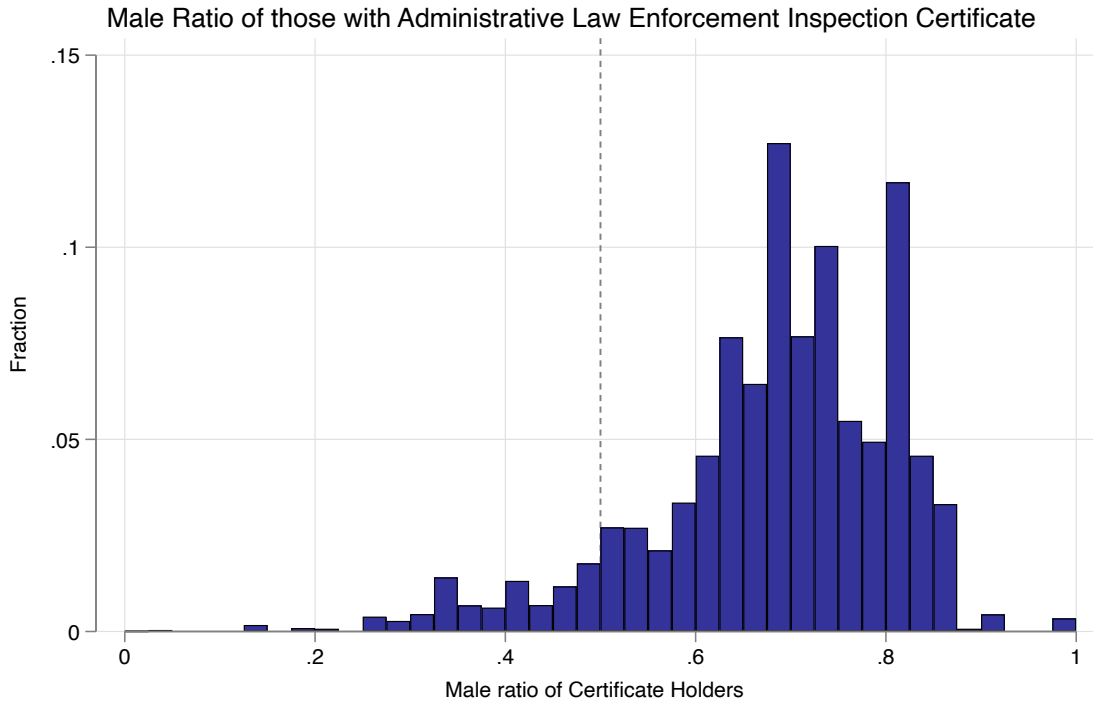
Notes - This figure shows the effect of the gender quota on cut-off scores, conditional on corruption levels. The top panel presents the effect on the county-level average cut-off score, while the bottom panel shows the effect by gender. Cut-off scores are standardized across counties. Corruption levels are classified based on the median number of corruption cases (three) from 2011 to 2015, using data from Wang (2020).

Figure A.23: The Effect on Different Tax Categories: After Matching



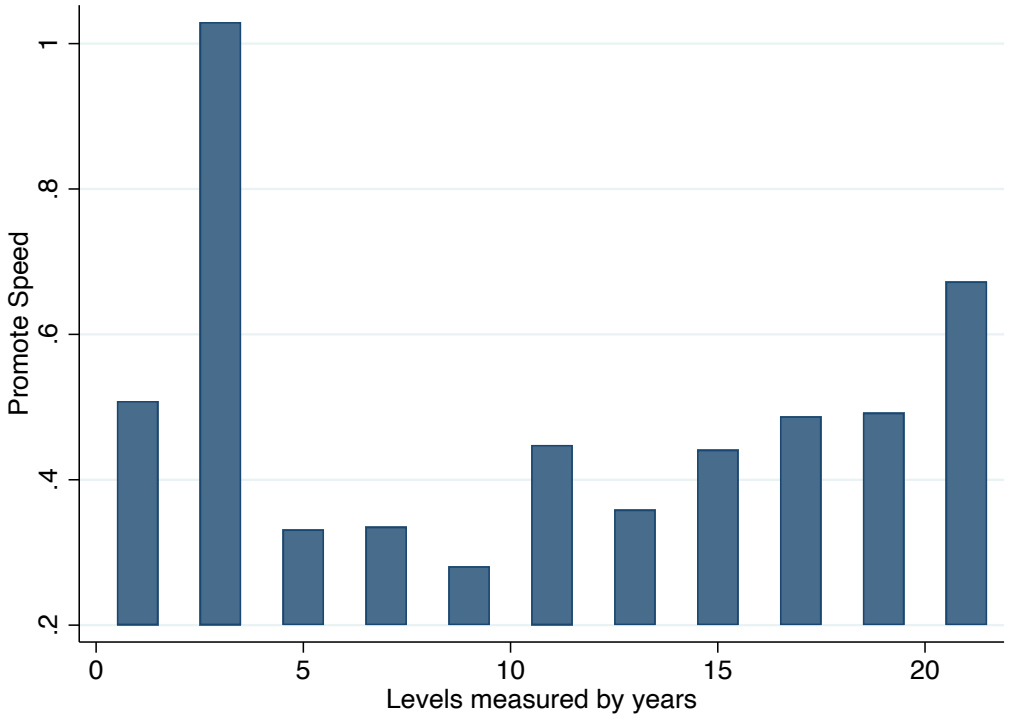
Notes - This figure shows the effect of gender quota intensity on different types of tax revenue. "Other tax" refers to tax revenue excluding individual income tax, corporate income tax, and value-added tax. Tax revenue is measured in 10,000 RMB and presented in logarithmic form. All other specifications follow the main regression model.

Figure A.24: Male ratio of Administrative Law Enforcement Inspection Certificate Holders in Province Anhui



Notes - This figure displays the male share among administrative certificate holders across counties in Anhui Province. In most counties, men comprise more than 50% of certificate holders.

Figure A.25: The promote speed across different levels



Notes - This figure shows promotion speed across different current position levels. Promotion speed is measured as the ratio of the typical number of years required to reach a given level to the actual number of years it took an employee to be promoted.

Table A.1: The Effect of Quota on County Cut-off Score

Panel A: Average County Cut-off Score				
	(1)	(2)	(3)	(4)
	Basic	Basic	After Matching	After Matching
Mean Share of Quota-position	-0.223***		-0.223***	
	(0.033)		(0.041)	
Ever Treated		-0.124***		-0.128***
		(0.022)		(0.028)
Number of vacancies	-0.022***	-0.022***	-0.025***	-0.024***
	(0.002)	(0.002)	(0.002)	(0.002)
Observations	19199	19199	13334	13334
Mean	-.007	- .007	-.011	-.011
r2	0.602	0.602	0.619	0.619
Panel B: Cut-off Score by Gender				
	(5)	(6)	(7)	(8)
	Male	Female	After Matching:Male	After Matching:Female
Mean Share of Quota-position	-0.734***	0.138***	-0.772***	0.116***
	(0.039)	(0.036)	(0.049)	(0.044)
Number of Male Vacancies	-0.029***		-0.035***	
	(0.003)		(0.004)	
Number of Female Vacancies		-0.027***		-0.031***
		(0.002)		(0.004)
Observations	18918	19057	13150	13233
Mean	-.159	.137	-.159	.123
r2	0.542	0.602	0.564	0.617
Year_FE	X	X	X	X
County_FE	X	X	X	X
City_Year_Trend	X	X	X	X
Cluster_County	X	X	X	X

Notes - This table reports the effect of gender quotas on standardized cut-off scores at the county level. *Panel A* presents the overall effect on cut-off scores, while *Panel B* distinguishes the effects by gender, showing separate impacts for male-targeted and female-targeted positions. Cut-off scores are standardized. Columns (1) (2) (5) (6) present results for all samples, using county-year as the unit of observation. Columns (3) (4) (7) (8) are matched sample. The dependent variable in all regressions is the standardized cut-off score for a given county-year. Standard errors, shown in parentheses, are corrected for clustering at the county level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A.2: The Effect of Gender Quota on Number of Applicants

Panel A: Total Number of Applicants				
	(1)	(2)	(3)	(4)
	Basic	Basic	After Matching	After Matching
Mean Share of Quota-position	-97.168***		-89.386***	
	(15.158)		(17.749)	
Ever Treated		-48.096***		-45.294***
		(9.684)		(11.125)
Number of vacancies	49.361***	49.336***	47.093***	47.063***
	(2.670)	(2.680)	(2.826)	(2.835)
Observations	22493	22493	15861	15861
Mean	317.417	317.417	291.516	291.516
r2	0.605	0.604	0.575	0.575
Panel A: Number of Applicants by Gender - After Matching	(5)	(6)	(7)	(8)
	Male	Female	Male	Female
Mean Share of Quota-position	-2.322	-43.299**		
	(12.253)	(17.006)		
Ever Treated			-0.436	-26.763***
			(6.821)	(9.406)
Number of vacancies in a county	21.814***	28.557***	21.802***	28.604***
	(1.255)	(2.158)	(1.259)	(2.161)
Observations	11264	11264	11264	11264
Mean	132.413	185.678	132.413	185.678
r2	0.579	0.611	0.579	0.611
Year_FE	X	X	X	X
County_FE	X	X	X	X
City_Year_Trend	X	X	X	X
Cluster_County	X	X	X	X

Notes - This table reports the effect of gender quotas on the number of applicants at the county level. Panel A presents the overall effect on total applicant numbers, while Panel B distinguishes the effects by gender, showing separate impacts for male-targeted and female-targeted positions. Columns (1), (2), (5), and (6) use the full sample with county-year as the unit of observation. Columns (3), (4), (7), and (8) use the matched sample. The dependent variable in all regressions is the number of applicants for a given county-year. Data on the gender ratio of applicants are not available. Instead, we assume that the gender ratio of candidates who pass the written exam reflects the gender ratio of applicants. This ratio is estimated using a gender-classification API based on candidate names. Standard errors, shown in parentheses, are corrected for clustering at the county level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A.3: The Effect of Quota on Tax Revenue within Samples Ever Hired and Never Stop Treatment

	(1) If Ever Hired	(2) Never Stop Treatment
Mean Share of Quota-position	-0.068**	-0.092**
	(0.032)	(0.036)
Observations	4258	4849
Mean	165058.23	157694.355
r2	0.970	0.968
Year_FE	X	X
County_FE	X	X
City_Year_Trend	X	X
Cluster_County	X	X

Notes - This table reports robustness checks using restricted samples. Column (1) restricts the sample to counties that have ever implemented gender-targeted hiring. Column (2) focuses on counties that never discontinued the quota once adopted. All samples are matched sample. Standard errors, shown in parentheses, are corrected for clustering at the county level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A.4: The Effect on Tax Revenue(After Matching):Conley Clustering

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	20km	50km	100km	200km	300km	400km	500km	600km
Mean Share of Quota-position	-0.085***	-0.085***	-0.085**	-0.085**	-0.085**	-0.085**	-0.085**	-0.085**
	(0.030)	(0.031)	(0.035)	(0.039)	(0.040)	(0.040)	(0.038)	(0.036)
Observations	5069	5069	5069	5069	5069	5069	5069	5069
Mean	11.089	11.089	11.089	11.089	11.089	11.089	11.089	11.089
r2	0.229	0.229	0.229	0.229	0.229	0.229	0.229	0.229
Year_FE	X	X	X	X	X	X	X	X
County_FE	X	X	X	X	X	X	X	X
City_Year_Trend	X	X	X	X	X	X	X	X
Cluster_district	X	X	X	X	X	X	X	X
Ever Treated	-0.033*	-0.033*	-0.033*	-0.033	-0.033	-0.033	-0.033	-0.033*
	(0.017)	(0.018)	(0.020)	(0.022)	(0.022)	(0.023)	(0.021)	(0.019)
Observations	5069	5069	5069	5069	5069	5069	5069	5069
Mean	11.089	11.089	11.089	11.089	11.089	11.089	11.089	11.089
r2	0.228	0.228	0.228	0.228	0.228	0.228	0.228	0.228
Year_FE	X	X	X	X	X	X	X	X
County_FE	X	X	X	X	X	X	X	X
City_Year_Trend	X	X	X	X	X	X	X	X
Cluster_district	X	X	X	X	X	X	X	X

Notes - This table reports robustness checks using different clustering radii, ranging from 20 km to 600 km. The dependent variable in all regressions is the log of county-level tax revenue (in 10,000 RMB). The key explanatory variables are the average share of gender-targeted positions across treated years and an indicator for whether the quota was adopted. All regressions include county fixed effects, year fixed effects, and city-by-year trends. Standard errors, reported in parentheses, are clustered at the county level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A.5: The Effect of Gender Quota on Total Tax Revenue: After Matching

	(1) All	(2) With Control	(3) All	(4) With Control
Mean Share of Quota-position	-0.091*	-0.100**		
	(0.047)	(0.047)		
Ever Treated			-0.038	-0.041
			(0.027)	(0.027)
Observations	2919	2919	2919	2919
Mean	10.746	10.746	10.746	10.746
r2	0.973	0.973	0.973	0.973
PreTreatControl_year		X		X
Year_FE	X	X	X	X
County_FE	X	X	X	X
City_Year_Trend	X	X	X	X
Cluster_County	X	X	X	X

Notes - This table reports the effect of the gender quota on tax revenue using the same sample when control pre-treatment county characteristics. The matching variables include tertiary sector share, population, male ratio, and urban-hukou ratio. Columns (1) and (3) present results based on the main specification, using county-year as the unit of observation. Columns (2) and (4) add interactions between pre-treatment county characteristics and year fixed effects. The pre-treatment controls include tertiary sector share, population, male ratio, and urban-hukou ratio. The dependent variable in all regressions is the log of county-level tax revenue (in 10,000 RMB). The key explanatory variables are the average share of gender-targeted positions across treated years and an indicator for whether the quota was adopted. All regressions include county fixed effects, year fixed effects, and city-by-year trends. Standard errors, reported in parentheses, are clustered at the county level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A.6: The Results controlling Spillover Effect

	(1) Tax(log)	(2) Score	(3) Male-Score	(4) Female-Score	(5) Num of Applicant
Mean Share of Quota-position	-0.156*	-0.197**	-0.641***	0.011	-109.880***
	(0.082)	(0.078)	(0.091)	(0.082)	(27.097)
Mean Share of Quota-position of Other Counties within Same City	0.080	-0.041	-0.198*	0.160	29.836
	(0.083)	(0.099)	(0.116)	(0.105)	(32.584)
Number of Vacancies within a Treated County		-0.024***			47.070***
		(0.002)			(2.830)
Number of Male Vacancies			-0.035***		
			(0.004)		
Number of Female Vacancies				-0.031***	
				(0.004)	
Observations	3656	13334	13150	13233	15861
Mean	11.023	-.011	-.159	.123	291.516
r2	0.965	0.619	0.564	0.617	0.575
Year_FE	X	X	X	X	X
County_FE	X	X	X	X	X
City_Year_Trend	X	X	X	X	X
Cluster_County	X	X	X	X	X

Notes - This table reports the effect of the gender quota on all outcomes controlling spillover effects using matched sample. The matching variables include tertiary sector share, population, male ratio, and urban-hukou ratioThe dependent variable in all regressions is the log of county-level tax revenue (in 10,000 RMB). The key explanatory variable is the average share of gender-targeted positions across treated years. All regressions include county fixed effects, year fixed effects, and city-by-year trends. Standard errors, reported in parentheses, are clustered at the county level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A.7: The Effect of Average County-level Standardized Cut-off Score on Total Tax Revenue

	(1) After Matching	(2) Basic
Average County Standardized Cut-off Score	0.015*	0.014*
	(0.008)	(0.007)
Observations	1728	2858
Mean	11.177	11.402
r2	0.984	0.983
Year_FE	X	X
County_FE	X	X
City_Year_Trend	X	X
Cluster_County	X	X

Notes - This table reports the effect of the cut-off score on tax revenue. Column(1) shows the result for matched sample, while column (2) shows the result of all sample. The dependent variable in all regressions is the log of county-level tax revenue (in 10,000 RMB). The key explanatory variable is the average share of gender-targeted positions across treated years. All regressions include county fixed effects, year fixed effects, and city-by-year trends. Standard errors, reported in parentheses, are clustered at the county level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A.8: Gender Quota and Score (Matching): By Corruption Level

	(1) Scoret High Corrupt	(2) Score Low Corrupt	(3) MaleScore High Corrupt	(4) MaleScore- Low Corrupt	(5) FemaleScore High Corrupt	(6) FemaleScore Low Corrupt
Mean Share of Quota-position	-0.338***	-0.247**	-0.748***	-0.495***	0.123	-0.005
	(0.075)	(0.108)	(0.086)	(0.132)	(0.079)	(0.114)
Observations	4605	8729	4551	8599	4572	8661
Mean	.045	.008	.204	.135	.092	.14
r2	0.543	0.658	0.509	0.601	0.538	0.658
Year_FE	X	X	X	X	X	X
County_FE	X	X	X	X	X	X
City_Year_Trend	X	X	X	X	X	X
Cluster_County	X	X	X	X	X	X

Notes - This table shows the effect of the gender quota on cut-off scores, conditional on corruption levels. Column(1) and (2) present the effect on the county-level average cut-off score, while column (3) to (6) show the effect by gender. Cut-off scores are standardized across counties. Corruption levels are classified based on the median number of corruption cases (three) from 2011 to 2015, using data from Wang (2020). Standard errors, reported in parentheses, are clustered at the county level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A.9: The Effect of Corruption and Score on Tax

	(1) Tax Revenue	(2) Tax Revenue	(3) Tax Revenue	(4) Tax Revenue
Average County Standardized Cut-off Score	0.024**		0.017	0.041**
	(0.012)		(0.022)	(0.020)
Number of Corruption Cases		0.009	0.017*	
		(0.007)	(0.010)	
Number of Corruption Cases × Average County Standardized Cut-off Score				-0.004*
				(0.003)
Observations	925	2367	667	925
Mean	11.015	10.622	10.78	11.015
r2	0.989	0.902	0.891	0.989
PreTreatControl		X	X	
Year_FE	X	X	X	X
City_FE		X	X	
County_FE	X			X
City_Year_Trend	X			X
Cluster_County	X	X	X	X

Notes - This table shows the effect of the corruption and cut-off score on tax revenue. Corruption levels are classified based on the median number of corruption cases (three) from 2011 to 2015, using data from Wang (2020). Standard errors, reported in parentheses, are clustered at the county level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A.10: Gender Quota and Total Tax Revenue: By Drinking Culture

	(1) Famous Drinking Culture	(2) Not Famous Drinking Culture	(3) Famous Drinking Culture	(4) Not Famous Drinking Culture
Mean Share of Quota-position	-0.162***	-0.062	-0.490**	0.007
	(0.060)	(0.042)	(0.201)	(0.158)
Observations	1423	3646	2340	4467
Mean	11.446	10.949	11.512	11.024
r2	0.971	0.965	0.781	0.800
Year_FE	X	X	X	X
County_FE	X	X	X	X
City_Year_Trend	X	X	X	X
Cluster_County	X	X	X	X

Notes - This table shows the effect of the corruption and cut-off score on tax revenue. The famous drinking culture cities are known for strong drinking culture, ranked by weighted city-level alcohol industry GDP. Standard errors, reported in parentheses, are clustered at the county level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A.11: The Effect of Average County-level Standardized Cut-off Score on Total Tax Revenue

	(1) Tax(log)	(2) Tax(log)	(3) Tax(log)	(4) Tax(log)	(5) Tax(log)	(6) Tax(log)
Female Ratio of Previously New Hires	0.085			0.091		
	(0.066)			(0.077)		
Female Ratio of Current-Year New Hires		0.004			-0.017	
		(0.018)			(0.023)	
Female Ratio of All New Hires			0.177*			0.159
			(0.107)			(0.112)
Observations	1422	1728	1422	955	1118	955
Mean	11.195	11.177	11.195	10.971	10.952	10.971
r2	0.987	0.984	0.987	0.985	0.982	0.985
PreTreatControl_year				X	X	X
Year_FE	X	X	X	X	X	X
County_FE	X	X	X	X	X	X
City_Year_Trend	X	X	X	X	X	X
Cluster_County	X	X	X	X	X	X

Notes - This table shows the effect of the female ratio on tax revenue. *Female Ratio of Previously New Hires* is the share of female new hires before treatment. The *Female Ratio of Current-Year New Hires* refers to the share of female new hires among all new hires in the given year, and the *Female Ratio of All New Hires* captures the share of female hires among all new hires from 2016 up to the current year. All ratios are calculated at the county level. Standard errors, reported in parentheses, are clustered at the county level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A.12: The Effect of Male Ratio of Cohort on Promotion Speed

	(1)	(2)	(3)	(4)
	Promote Speed	Promote Speed	Promote Speed	Promote Speed
Gender	-0.008	-0.465***	-0.012	-0.490***
	(0.012)	(0.117)	(0.012)	(0.118)
Male ratio in the same cohort	-0.665***	-1.312***		
	(0.145)	(0.226)		
Male ratio in the same cohort except self			-0.662***	-1.339***
			(0.144)	(0.228)
Male ratio in the same cohort × Gender		0.713***		
		(0.173)		
Male ratio in the same cohort except self × Gender				0.743***
				(0.173)
age	-0.239***	-0.199***	-0.239***	-0.197***
	(0.033)	(0.030)	(0.033)	(0.030)
age2	0.002***	0.002***	0.002***	0.002***
	(0.000)	(0.000)	(0.000)	(0.000)
Observations	928	928	928	928
Mean	.35	.35	.35	.35
r2	0.731	0.746	0.731	0.747
County_FE	X	X	X	X
Cluster_County	X	X	X	X

Notes - This table shows the effect of the male ratio within the peer cohort on promotion speed. Standard errors, reported in parentheses, are clustered at the county level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$