The Costs of Affirmative Action: Evidence from a Male-Favoring Quota

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

March 3, 2026

Abstract

Hiring quotas can reshape who enters the workforce and how organizations perform. Such policies are often intended to redress disparities arising from past discrimination. I study a gender quota that instead preserves the representation of a historically dominant group: men in China's civil service. After merit-based hiring boosted women's share of new recruits, many county tax bureaus introduced one-to-one gender quotas mandating equal hires by sex. Using a new dataset I built through web-scraping and digitizing position-level civil service exam records and county tax revenues, 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 on mechanisms suggests these effects stem from reduced personnel ability and gendered institutional practices. These findings show that "reverse" affirmative action can be costly when institutions that serve advantaged groups persist.

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, Ursina Schaede, Shan Aman-Rana, Sandip Sukhtankar, Zoe Cullen, Alex Eble, Jaya Wen, 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, Northeastern Universities Development Consortium (NEUDC) 2025, Applied Young Economist Webinar - Labor Economics Workshop 2025, for their helpful comments and suggestions. All errors are my own.

1 Introduction

Policies to redress historical discrimination aim to increase women's participation in traditionally male-dominated fields.[1] But in some fields - such as higher education, health professions, and parts of the public sector - women now outnumber men, raising claims of reverse discrimination (Cohen, 2024; Pollitt, 2023; Dominus, 2023; Nietzel, 2024). Rising in prominence worldwide are quotas that aim to maintain gender balance by preserving men's representation.[2] This paper asks two questions: what predicts 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 overall representation and performance.

I provide new empirical evidence on a gender quota that protects men's representation in China's civil service, showing how such "reverse" affirmative action policy reshapes female representation, lowers candidate quality, and reduces institutional productivity. Specifically, 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

[1] These policies also aim to support other disadvantaged groups, such as racial and ethnic minorities. For instance, affirmative action in universities and workplaces seeks to increase representation among both women and minorities (Holzer and Neumark, 2000). Similar redistributive rules include 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 examples highlight how affirmative and redistributive hiring policies can expand opportunities for historically excluded groups.

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

[3] County tax bureaus are the lowest administrative units under the State Administration of Taxation (SAT) and responsible for collecting central taxes (both central and local taxes after the 2018 merger of state and local tax bureaus). In 2023, county bureaus hired 24,975 of 37,100 new civil servants nationwide, accounting for 67.3% of all recruits.

2011 introduction of the National Civil Service Examination, women have comprised about two-thirds of new recruits (see Figure A.2). Some bureaus 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. What drives these county tax bureaus to adopt the quotas in hiring, and what are the consequences?

To answer 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 measures of productivity, I also collect the number of annual administrative approvals from county tax bureau reports for robustness checks.

The analysis consists of two parts. The first part examines what predicts adoption of 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 implement quotas. Since not all positions are subject to the quota, I also examine

[4]Each posting reports department and bureau names, website, position title and codes, 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.

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 between 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. In 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 view, 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 has a small positive effect on revenue. Adjusted event-study estimates confirm robustness

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

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 instrument variable (IV) based on even-numbered vacancies, controlling for spillovers from neighboring counties, and applying Conley (1999) spatial adjustments to standard errors. Together, these checks reinforce the conclusion that the productivity decline is a direct consequence of quota adoption.

I find evidence that the decline in productivity I document operates through two main channels: a direct effect from reduced personnel ability and an amplifying effect from gendered institutional practices. First, 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. Second, gendered institutional practices amplify the negative impact through two mechanisms. The first is the reinforcement of corruption networks: in counties with higher corruption, revenue declines following quota adoption are substantially larger, suggesting that the quota provides cover for patronage hiring and strengthens male-dominated ties. Supporting this interpretation, I document the prevalence of alcohol-consumption culture and weak regulatory enforcement, both of which facilitate such practices. The second is biased role 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. Finally, I rule out gender composition as an alternative mechanism, pointing instead to reduced worker quality and institutionalized gender bias as the central explanations for the productivity decline.

This study contributes to the literature on the impact of "reverse" affirmative action on productivity. 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 favoring men remain largely unexplored. This paper examines such a policy: a gender quota that protects men's representation. The closest work is Schaede and Mankki (2024), which investigates the abolition of a male quota for primary school teachers in Finland and finds long-term benefits for pupils. Although both papers study quotas favoring men, the contexts

differ fundamentally. First, the quota I study arises in China's civil service, a bureaucratic setting where men are historically advantaged, creating very different motivations for its adoption. Second, unlike quotas typically mandated by national law, the policy here is voluntarily adopted by local governments. Third, rather than focusing on outcomes for the group served by quota-affected individuals, I examine how the quota shapes the productivity of the institutions that implement it. This paper documents the adoption and impact of a gender quota favoring men, offering new insights into the potential costs of "reverse" affirmative action for historically dominant groups that have become outnumbered.

More broadly, this paper contributes to debates on the trade-off between representation and efficiency in affirmative action. Recent studies show 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 highlights inefficiencies from mismatch (Oliveira et al., 2024; Antonovics and Backes, 2014) or backlash (Ip et al., 2020; Leibbrandt et al., 2018). This paper instead examines a voluntary quota favoring a historically dominant group, showing how such policies can undermine the representation of disadvantaged groups and reduce efficiency, revealing the unintended costs of affirmative action.

Another strand of related literature examines the determinants of bureaucratic performance. Some studies emphasize the role of personnel ability in shaping productivity (Fenizia, 2022; Limodio, 2021; Colonnelli et al., 2020; Besley et al., 2017; Dal Bó et al., 2017), while others highlight how promotion incentives influence civil servant performance (Bianchi et al., 2023; Bertrand et al., 2020). This paper contributes by showing how hiring policies, particularly gender quotas, shape organizational outcomes. Existing research has examined how mandated female representation affects firms' performance (Ahern and Dittmar, 2012; Matsa and Miller, 2013; Eckbo et al., 2022; Ferrari et al., 2022), but few studies analyze the impact of gender quotas on bureaucrats' performance, especially in the context of civil service recruitment 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 gender quota favoring men in the recruitment of China's county tax bureaus, demonstrating

its consequences for tax revenue and, more broadly, for 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 service 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 5.1, before analyzing its effects on gender representation, recruitment outcomes, economic performance, and robustness checks in Section 5. Mechanisms are discussed in Section 6. The final section concludes.

2 Institutional Background

2.1 Civil Service 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, much like civil service exams in other developing countries (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]

[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

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

local Hukou registration.

[8]The sharp decline in 2019 reflects the merger of local and state taxation administrations, which increased staffing levels in tax bureaus by 79% 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. Each gender-targeted posting appears as part of a pair, comprising one male-designated and one female-designated position. A county is defined as adopting the quota policy if it posts any gender-targeted positions.

Applicants and admission data. I also web-scraped the number of applicants and the written exam cut-off scores for each position since 2016. [12] 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.

[12] The data are from one education training institution (https://www.huatu.com/).

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 Yearbooks and digitized information on total tax revenue, corporate tax revenue, individual tax revenue, and value-added tax revenue.[13] Not all City Yearbook publish the tax revenue data, so I collect all available City Yearbooks and record all available data.[14] 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 Predicts 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 control-

[13]The tax revenue data follow the First Digit Law Test (Benford's Law), supporting its validity (see Figure A.4).

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

ling for province fixed effects, I compare counties ever adopted the policy to those that never adopted within the same province. 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.

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 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. Female share among prior new hires tends to be higher in wealthier, more urbanized counties with more migrants, more women, and lower corruption (see Figure A.6). 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 share 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. Officials report difficulties in recruiting male candidates, citing greater opportunities for men in the private sector and a lower inclination toward public sector jobs. They also often express the view that men may be better suited for roles involving frequent social interaction, such as business travel, networking events, or client entertainment.

Moreover, to examine institutional preferences related to gender roles, I analyze quota implementation 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.

One concern is that the share of female applicants may differ across positions requiring different majors, leading positions in female-dominated fields to be more likely to adopt the quota.[15] To address this concern, I categorize major requirements into four groups: Arts and Humanities, STEM, Business, and Law, and I then control the major fixed effects in the regression. The results remain robust, suggesting that quota implementation reflects underlying preferences of tax bureau officials regarding gender roles (see Figure 1).

To sum up, counties with higher GDP, larger populations, more local residents, lower urbanization rates, and higher levels of corruption are more likely to adopt the quota. These correlations can be only partly explained by the female share of the population, which may predict a higher proportion of female applicants and, consequently, a higher share of female hires even before the quota. This explanation aligns with the recruitment pattern analysis, which shows that counties with a higher female ratio among recent hires are more likely to adopt the quota. Furthermore, the analysis by task category suggests that governments tend to perceive tax-related positions or positions of greater responsibility as requiring more

[15]Figure A.7 compares the major requirements of tax-related and administrative positions.

men and thus adopt quotas more often for these roles, while the opposite holds for administrative positions. Taken together, these findings suggest that, beyond the female share of recent hires, institutional gender norms within a male-dominated bureaucracy also play an important role in determining quota adoption.

5 Impacts of the One-to-one Quota

5.1 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.[16] 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.[17] This method generates a counterfactual group that is comparable to treated counties

[16] Each gender-targeted posting must appear in a pair, with at least one male and one female position posted together. Figure A.9 shows the share of gender quota positions in each county over the study period.

[17] 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.

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 6.2.1. 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 between 2016 and 2023. Although approximately 56% of these counties pause the policy for one or more years, either by stopping 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 with and without quota in each year, incorporating county fixed effects and year fixed effects:

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

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

where Y_{ct} represents the outcome of interests within county c in year t, including female share of new hires, average standardized cut-off score, number of applicants, county-level tax revenue. $Treat_{ct}$ equals 1 when county c initiates the gender quota from year t onwards. $TreatInt_{ct}$ indicates the treatment intensity measuring the mean share of positions with gender quota in county c from year t onwards. δ_t accounts for the year fixed effects, γ_c denotes the county fixed effects, 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 dynamic 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):

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

$$Y_{ct} = \sum_{k=-a}^{b} \theta_k I\{t = t_c^* + k\} * \Lambda_c + \delta_t + \gamma_c + \Theta_c \cdot t + \epsilon_{ct} \tag{4}$$

Equations 3 and 4 present the event-study specifications based on quota adoption and quota intensity. 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 share, 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. Λ_c measures the mean share of positions with gender quota in county c from year t_c^* onwards. 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 effects, γ_c denotes the county fixed effects, and ϵ_{ct} represents the error

term. $\Theta_c \cdot t$ indicates the city-specific year trend.

5.2 First-Stage Results

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 female share of new hires before the gender quota adoption exhibits a parallel trend. I find a one-percentage-point increase of quota intensity (gender-targeted position share) leads to a 0.25 percentage point decrease in the female share of current-year new hires, relative to a baseline average of 67.7%. Among all new hires before the adoption of gender quota, one percentage increase of gender quota intensity corresponds to a 0.08 percentage point 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 before quota adoption. Similar patterns emerge in the basic regression estimates (without matching) 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.

5.3 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 therefore 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, which is 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 0.002 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.

Because tax-related positions are more likely than others to adopt the quota, it is important to assess how the policy affects candidates across different 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 30.47% 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.[18] 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

[18]Specifically, I use the "ngender" package. One caveat, however, is that it has higher accuracy for predicting male names than female names.

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 are not linearly correlated with applicant volume; the relationship 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.

5.4 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 quota share of 55%, the one-to-one gender quota is estimated to reduce total tax revenue by approximately 4.7%. With average county-level tax revenue exceeding 6 billion RMB each year, this translates to an estimated loss of 30 million RMB, or roughly 4 million USD per county each year. I also control for the pre-treatment matching variables and year fixed effects 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. Focusing on the treatment intensity effect, the event study in Figure 6 shows no pre-trends and a significant decline in total tax revenue following the gender quota. The average effects of gender quota introduction, measured as a dummy variable that equals 1 after quota adoption, are provided in Column (2) and (4) of Table 1. The estimates are meaningful but not statistically significant at conventional levels. This is likely because 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). As a robustness check, I also show the effect of treatment intensity before matching method in Figure A.16, which

remains similar.

5.5 Robustness Checks

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

Macroeconomic Robustness. The first concern in this paper is distinguishing whether the observed decline in tax revenue is a consequence of quota adoption or instead reflects broader economic shifts occurring at the same time, which may influence taxpayer behavior independently of tax officials' qualifications. To rule out this concern, I first 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, if anything, is associated with a slight increase, mitigating concerns that revenue losses reflect broader economic downturns.

Taxpayer Behavior. An immediately related concern is that GDP may not adequately capture shifts in taxpayer behavior, raising the possibility that the decline in tax revenue reflects only changes in taxpayers' actions rather than the conduct of tax officials. To evaluate this possibility, I use an alternative measure for personnel productivity - 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 the outcomes of routine interactions with taxpayers and potentially the ability of personnel. 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. If taxpayer behavior were the main margin of adjustment, one would expect to observe changes in punishments. Figure 7 presents the effects of the gender quota on both administrative approvals and penalties. The annual number of penalties remains unchanged, while administrative approvals decline significantly following the adoption of the gender quota, consistent with the argument that the decline is driven by quota adoption rather than concurrent changes in taxpayer behaviors. 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.[19]

[19]The data for number of administrative cases covers only 2019 to 2023, restricting the pre-treatment

Hiring Effects. 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 shows that hiring has a slightly positive effect on tax revenue, ruling out hiring as a confounding factor (see Figure A.20). 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.

Different Sample and Standard Errors. The findings also remain robust when clustering standard errors using Conley (1999) spatial adjustment (see Table A.4) and when restricting the sample to match that used for controlling pre-treatment matching variables (see Table A.5).

Alternative DID Estimators. 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).

IV-DiD Strategy. To complement the matching strategy, I implement an instrumental variable approach (see Appendix C Equation 5). 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

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).

the true negative effect of the quota. This pattern likely reflects positive selection: as Figure A.19 suggests, cities that adopt the quota tend to grow faster afterward, so unobserved factors linked to both stronger growth and quota adoption may bias the effect downward. Another possibility is measurement error in the quota intensity variable, which could also attenuate the OLS estimates.

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) of Table 3 are still robust.

Spillover Effects. 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 remain unchanged, suggesting that spillover effects do not drive the findings. The event study after controlling for neighbor counties' quota adoption also shows no pre-trend and the effect is robust (see Figure A.21).

6 Channels

To understand the mechanisms through which the one-to-one gender quota reduces institutional productivity, I examine two channels: reduced personnel ability and gendered institutional practices. First, I find that revenue losses stem from the admission of less qualified men

(see Section 6.1), as shown by the positive correlation between exam cut-off scores and tax revenue. Second, the amplifying effect of gendered institutional practices operates through two mechanisms: the male-dominated system may reinforce corruption networks and may also favor men for key roles and leadership positions. Consistent with this, revenue losses are larger in more corrupt counties, and men are more often assigned to important jobs. Additional evidence on the prevalence of an alcohol-consumption culture, weak regulatory enforcement, and gender gaps in promotion further supports these interpretations. I am able to rule out alternative explanations including gender composition effects, showing that changes in overall male–female ratios alone do not explain the results.

6.1 Reduced Personnel Ability

Quota adoption reduces the average exam score of candidates considered for role, which may reflect declines in new hires' quality overall. Thus, the reductions in tax revenue may derive - at least in part - from lower ability among new hires. In the pre-treatment sample, cut-off scores are positively related to tax revenue conditional on county and year fixed effects and city-by-year trends (Figure 8, Panel A). Table A.7 moves from visualization to estimation, showing that a one standard deviation increase in the cut-off score raises tax revenue by 1.5%. To better capture changes in personnel ability, Panel B considers an alternative productivity measure, the annual number of administrative approvals, which reflects both qualifications and effort exerted. Recalling the decline in approvals shown in Figure 7, the evidence suggests that the revenue decline is related to reduced personnel ability, understood as the combination of qualifications and effort. Personnel ability is not the whole story. In the next section, I examine institutional practices and then test both mechanisms simultaneously.

6.2 Gendered Institutional Practices

Besides the direct impact of reduced personnel ability, gender norms within institutions also play a role. As a male-dominated organization, the county tax department may seek to attract more men for two reasons. First, preserving existing male networks can help maintain

institutional control and facilitate patronage or corruption. Second, gender bias may lead to the belief that men are better suited for certain tasks and therefore should be allocated to more important positions.

6.2.1 Male-Dominated Networks and Corruption

Prior research finds that corruption is more prevalent in male-dominated systems (Decarolis et al., 2023; Dollar et al., 2001; Swamy et al., 2001; Brollo and Troiano, 2016). Gendered institutional practices that afford men greater opportunities to engage in corruption may explain this stylized fact. In other words, the way men operate in a role versus how women do - shaped by social norms or rules of behavior - can generate differences in corrupt behavior. In China, a clear behavioral difference lies in socializing: men in business and government more often fraternize over meals and alcohol, while women engage less in this mode of networking. In the tax context, interactions with taxpayers create opportunities for corruption - for example, overlooking evasion in exchange for bribes - which strictly reduces tax revenue. Such corruption would strictly decrease tax revenue. In this section, I examine whether gendered socialization norms help explain the documented reductions in tax revenue.

Corruption as a Driver. First, I observe in Section 4 that counties with higher levels of corruption are more likely to adopt the policy. Though correlational, this pattern is consistent with tax bureaus seeking to "protect" existing corruption networks by adopting a policy favoring male workers who can integrate into them. Furthermore, the effect is largest for "other taxes" (excluding individual income, corporate income, and VAT; Figure A.23) - a less standardized category with more space for corruption - suggesting that corruption may be part of the story in the average effects.[20]

Heterogeneity by Corruption. I then examine whether the effects of quota adoption vary with a county's pre-treatment level of corruption, measured by the total number of corruption cases recorded before 2016.[21] If men and women are equally likely to engage in corruption or integrate with the corruption-friendly office culture, we should observe no

[20]Other taxes include environmental protection tax, property tax, urban land use tax, land appreciation tax, resource tax, and tobacco leaf tax.

[21]The data is from China's Corruption Investigations Dataset (Wang, 2020). China's Corruption Investigations Dataset includes information on almost 20,000 officials who were investigated during Xi Jinping's anti-corruption campaign.

differential trends. In fact, counties that were more corrupt ex ante experience larger declines in tax revenue following quota adoption than those that were less corrupt (Figure 9). These patterns are consistent with male-dominated corruption networks expanding as more men replace women in tax bureaus.

Socialization and Corruption. To support the idea that corruption operates through socializing with taxpayers over meals and alcohol, I examine whether the effect of the gender quota varies with local drinking culture. We would expect no differential effects if men and women are equally likely to socialize over alcohol. I collect the city list known for strong drinking culture, ranked by weighted city-level alcohol industry GDP.[22] Table A.10 shows that counties located in these cities experience a larger negative effect of the gender quota on productivity. The result aligns more closely with the idea that gendered institutional norms create greater opportunities for men, who are more likely to engage in social interactions that foster corruption.

Corruption and Personnel Ability. If tax bureaus prefer men and consequently hire more men, one concern is that highly corrupt counties may recruit less qualified men, contributing to larger declines in tax revenue. To address this concern, I examine whether changes in candidate quality - measured by cut-off scores - vary with counties' pre-existing levels of corruption and disentangle the effects of personnel ability from those of corruption.

First, I find that the decline in candidate quality, reflected in lower cut-off scores, is indeed more pronounced in highly corrupt counties than in less corrupt ones, though the difference is not significant (see Table A.8 and Figure A.22).

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

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

associated with tax revenue. This suggests that 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. Since I find 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.

6.2.2 Gender Beliefs and Role Allocation

Another channel operating through gendered institutional practices may involve the belief that men are better suited for tax collection and thus receive greater resources or responsibilities. Such preferential allocation could amplify the negative effects of hiring less qualified men and reduce the productivity of incumbent female workers by weakening their work incentives. In this section, I provide suggestive evidence that men are more likely to be assigned to important positions and promoted.

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 web-scraped 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 (Anhui, Guangdong, Guizhou, and Zhejiang), 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, since promotion speed can mechanically differ across position levels, as shown in 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 ratio 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.

Gender Composition and Performance. A potential mechanism is that the one-to-one gender quota alters the gender composition of tax bureaus, which in turn may affect worker productivity. I examine whether the female share of new entrants – both current-year new entrants and cumulative new entrants – affect the tax revenue. Table A.11 shows the female share of cumulative new entrants is positively correlated with the tax revenue, though the results for other kinds of female share are not significant. This result suggests that gender composition plays only a very limited role in explaining the productivity decline.

The analysis above suggests that the decline in tax revenue is driven by both reduced personnel ability and male-dominated institutional norms. Tax bureaus appear to adopt the gender quota policy with the intent of preserving existing male networks, leading to the recruitment of less qualified men in counties where male preference is stronger. This intent generates two consequences: a decline in personnel quality and the strengthening of male-dominated networks that foster corruption and biased resource allocation. The simultaneous and interactive effects of these two forces drive the observed decline in institutional productivity.

7 Conclusion

I provide the first evidence on a one-to-one gender quota favoring men in China's civil service, a setting in which women's representation has expanded but the institution remains male-dominated. I examine both the motivations for adoption and the policy's consequences, offering a new perspective on "reverse" affirmative action. The quota, which enforces numerical parity among new hires, reduces candidate quality, discourages women's participation, and lowers bureaucratic productivity. Mechanism analysis shows that it weakens personnel ability and reinforces gendered institutional practices, thereby diminishing institutional performance.

When women begin to outnumber men in certain fields, such as the declining share of male college students, debates often arise over whether "reverse" affirmative action is needed. A central policy question is whether such measures help or harm institutions. While it is important to be cautious in claiming that all "reverse" affirmative action is costly, this paper shows that "reverse" affirmative action is costly when adopted in contexts where men remain dominant and when motivated not by gender balance but by biased gender perceptions.

More importantly, the results of this paper extend beyond gender, applying more broadly to cases where historically disadvantaged groups outnumber advantaged groups, including racial minorities, immigrants, and other marginalized populations. It also opens avenues for studying "reverse" affirmative action across different institutional and cultural settings. In higher education, for example, Japan's medical schools have been found to manipulate admissions to favor male applicants, limiting opportunities for women despite their academic advantage (Schieder, 2019). In the United States, elite universities historically reserved spots for men to preserve the "college boy" ideal (Dominus, 2023). Similar dynamics may arise in private firms or professional organizations where historically advantaged groups face the prospect of losing numerical dominance yet continue to control decision-making. Comparative evidence across countries and sectors would help clarify when and why such policies emerge, and whether they systematically trade off representation and efficiency. Examining these broader contexts would provide a fuller understanding of how "reverse" gender quotas shape equity, meritocracy, and institutional performance.

Given data limitations, the paper cannot fully examine applicant behavior, but I provide suggestive evidence of a post-quota decline in applications (Figure 5). Future research could further examine how such policies shape applicant decisions, particularly how men and women respond differently to gender restrictions in hiring. Relatedly, studying how quotas affect women's long-term career trajectories, promotion opportunities, and access to leadership positions would shed light on the broader dynamics of gender inequality in bureaucracy.

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.

Ahern, Kenneth R and Amy K Dittmar, "The changing of the boards: The impact on firm valuation of mandated female board representation," *The quarterly journal of economics*, 2012, *127* (1), 137–197.

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.

Ferrari, Giulia, Valeria Ferraro, Paola Profeta, and Chiara Pronzato, "Do board gender quotas matter? Selection, performance, and stock market effects," *Management Science*, 2022, *68* (8), 5618–5643.

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.

Matsa, David A and Amalia R Miller, "A female style in corporate leadership? Evidence from quotas," *American Economic Journal: Applied Economics*, 2013, *5* (3), 136–169.

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.

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.

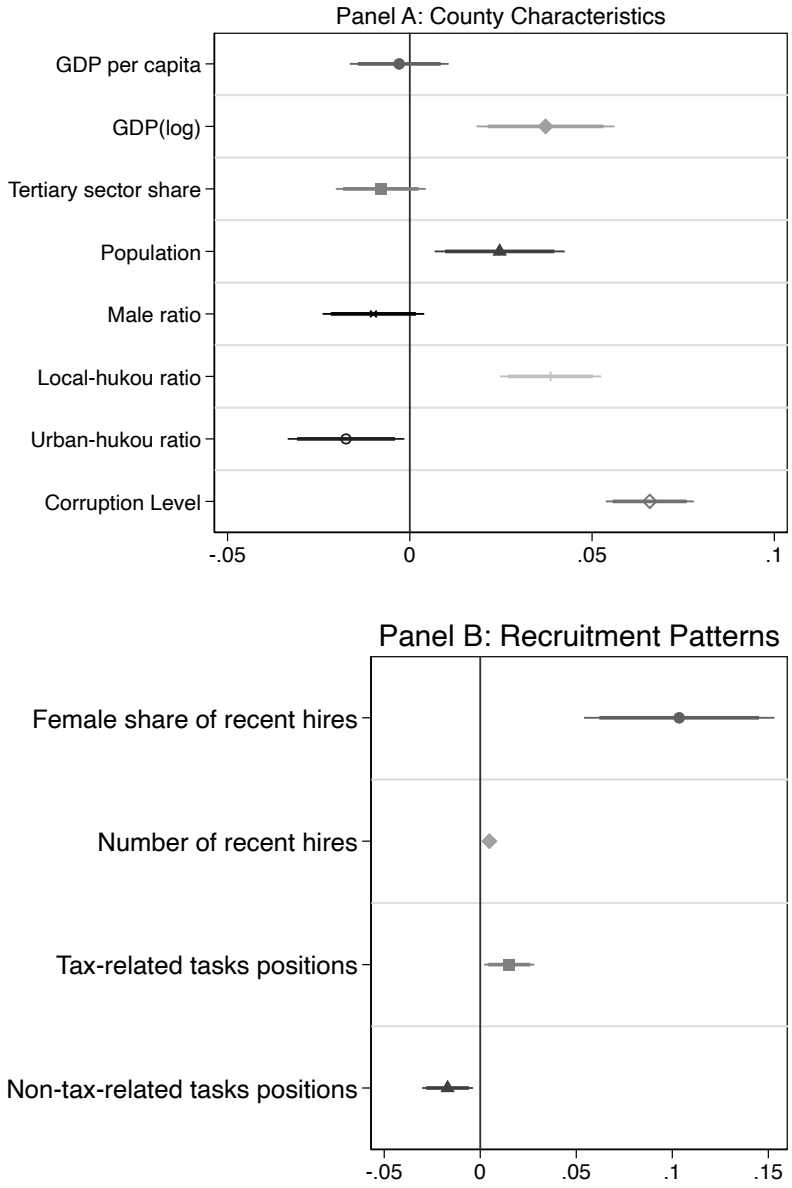
Sun, Liyang and Sarah Abraham, "Estimating dynamic treatment effects in event studies with heterogeneous treatment effects," *Journal of econometrics*, 2021, *225* (2), 175–199.

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.

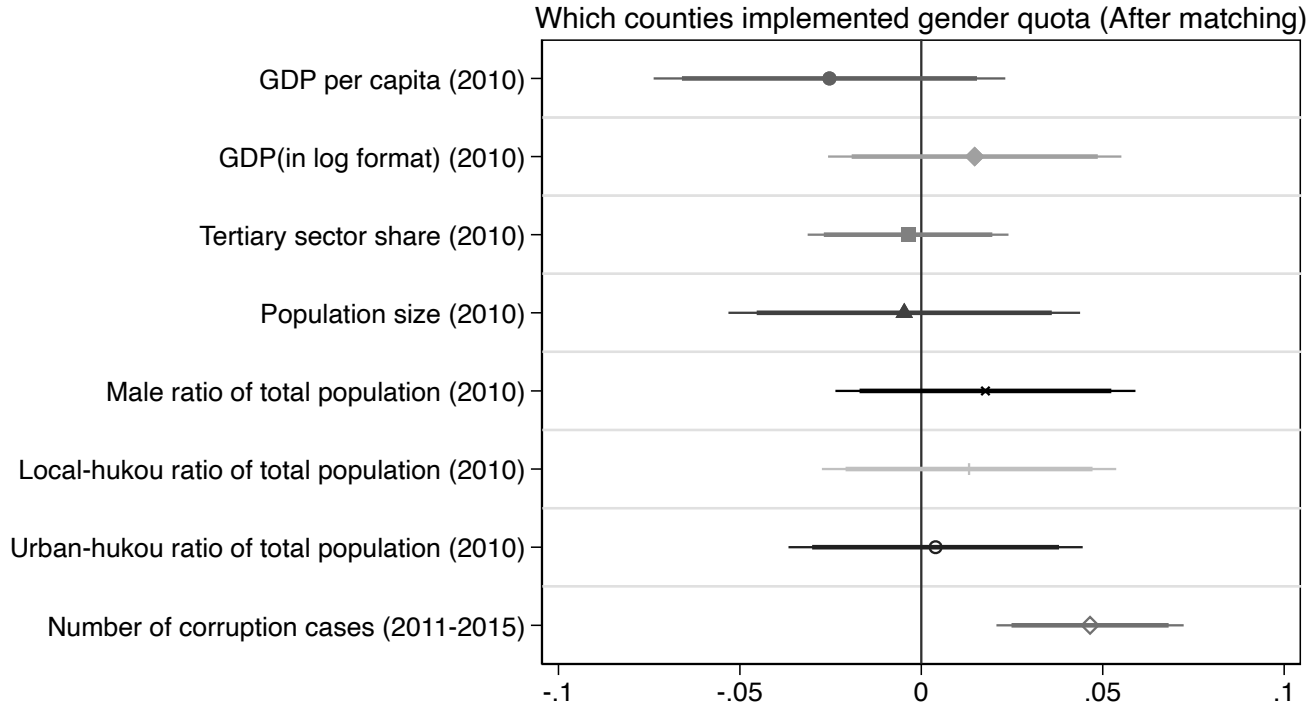
Wang, Yuhua, "China's Corruption Investigations Dataset," 2020.

Figure 1: What Predicts Quota Qdoption?



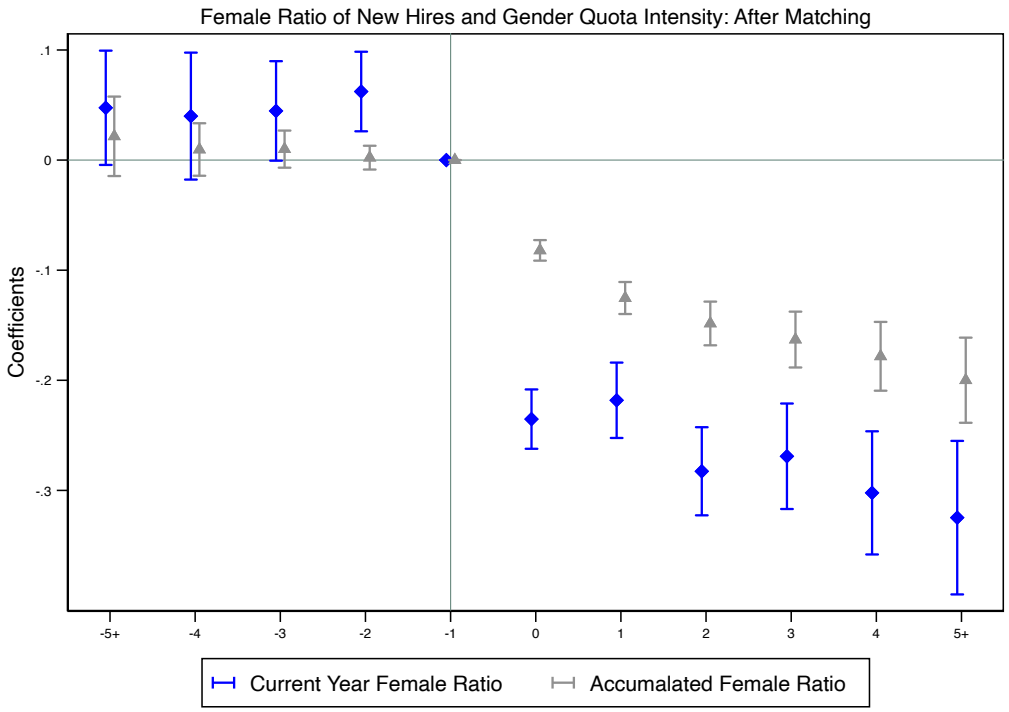
Notes - This figure presents the determinants of gender quota adoption, incorporating both county-level and recruitment patterns. *Panel A* displays the correlations between quota adoption and pre-treatment county characteristics, using county-level data from the 2010 Census. Corruption levels are sourced 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 patterns. 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



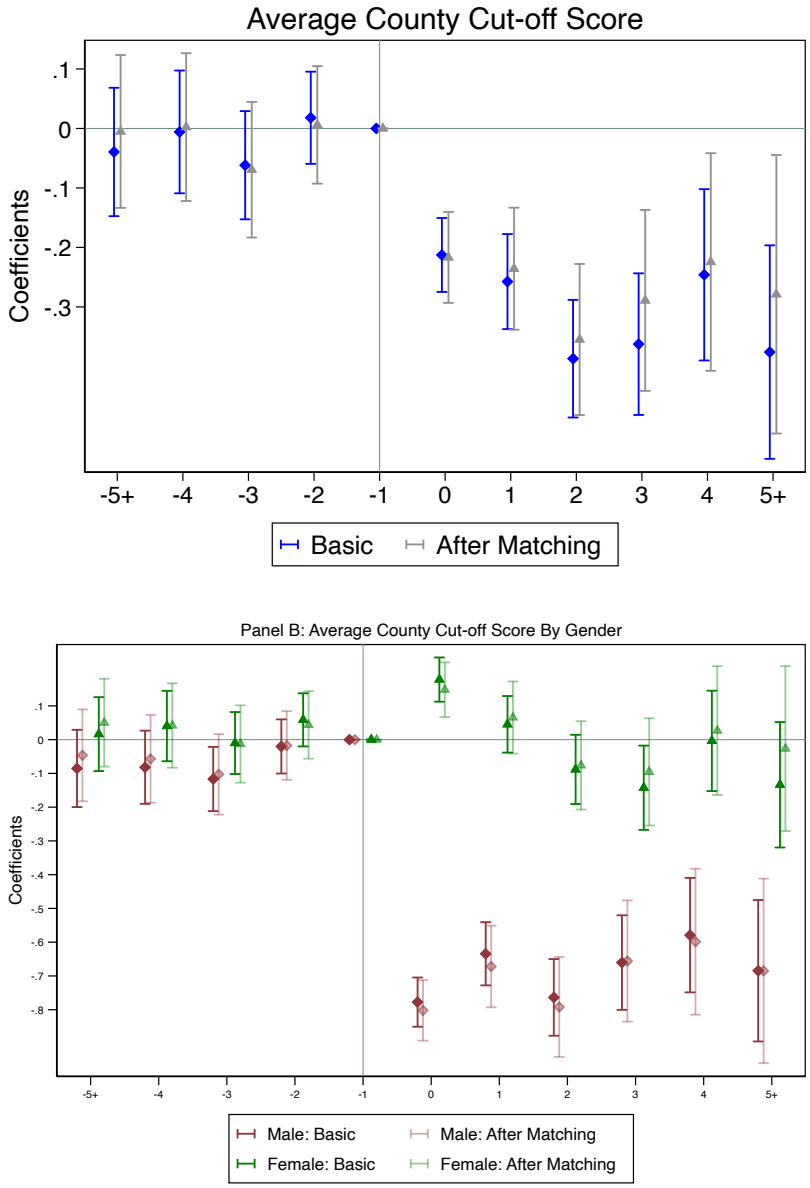
Notes - This figure presents the correlation between pre-treatment county characteristics and quota adoption after matching. The matching successfully balances both the variables used in the matching process-tertiary sector share, population, male ratio, and urban-hukou ratio-and other covariates not 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



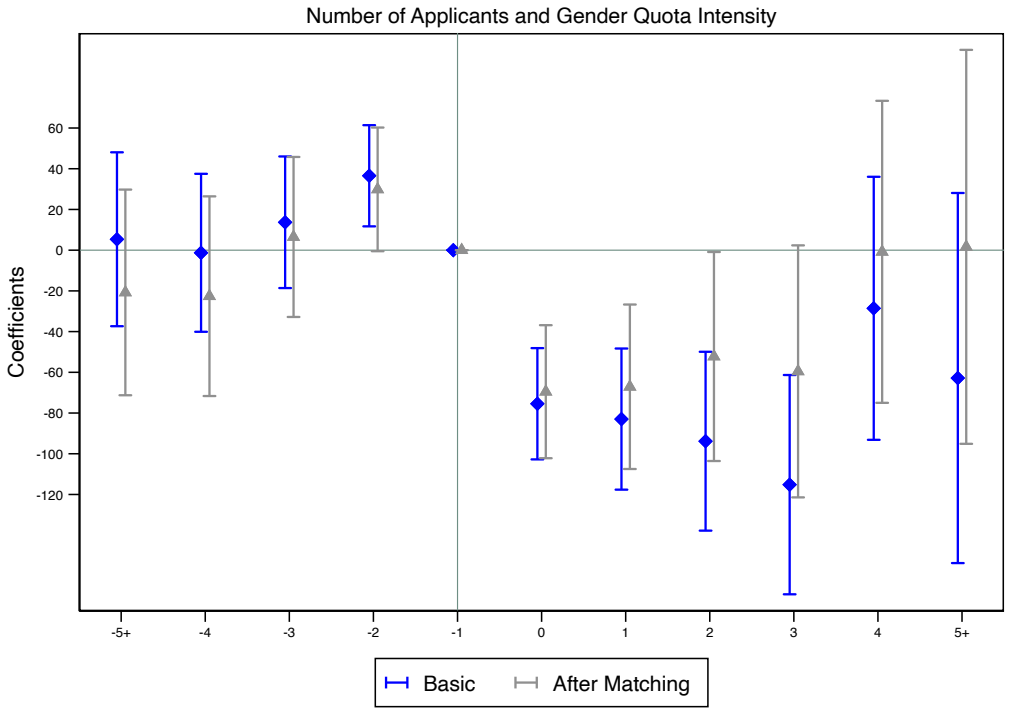
Notes - This figure presents the first-stage results of gender quota intensity, measured as the average share of gender-targeted positions across all treated years. Gender composition is derived from admission data starting in 2016. The *current-year female share* refers to the share of female new hires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



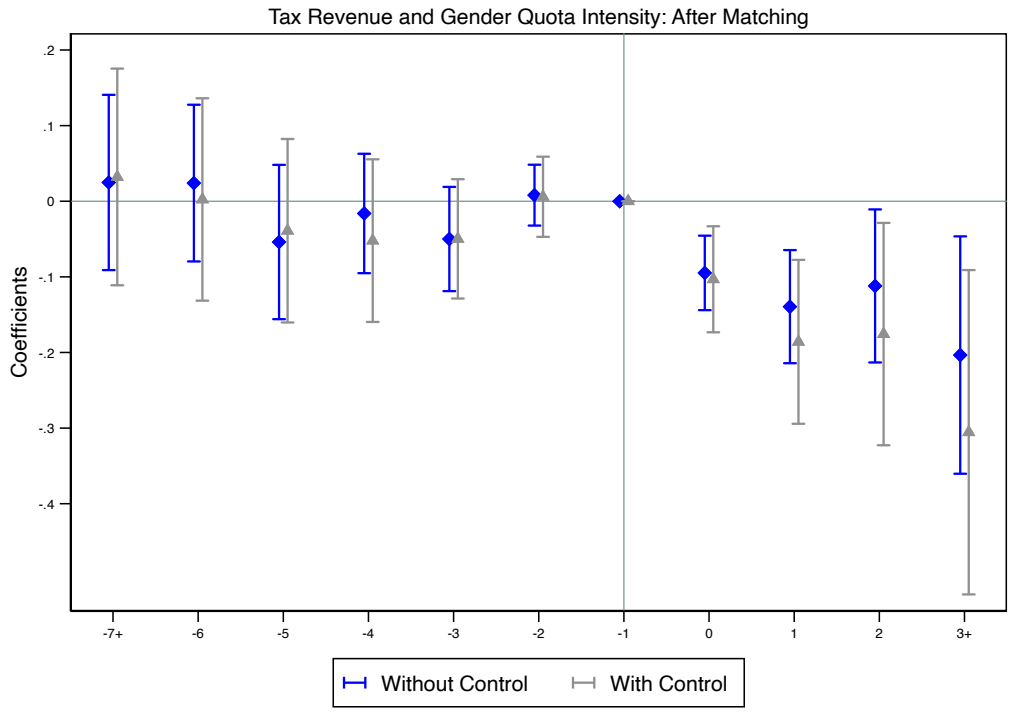
Notes - This figure shows the effect of gender quota intensity on the average county-level cut-off score. Quota intensity is measured by the average share of gender-targeted positions across all treated years. *Panel A* presents the overall effect on cut-off scores, while *Panel B* distinguishes the 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



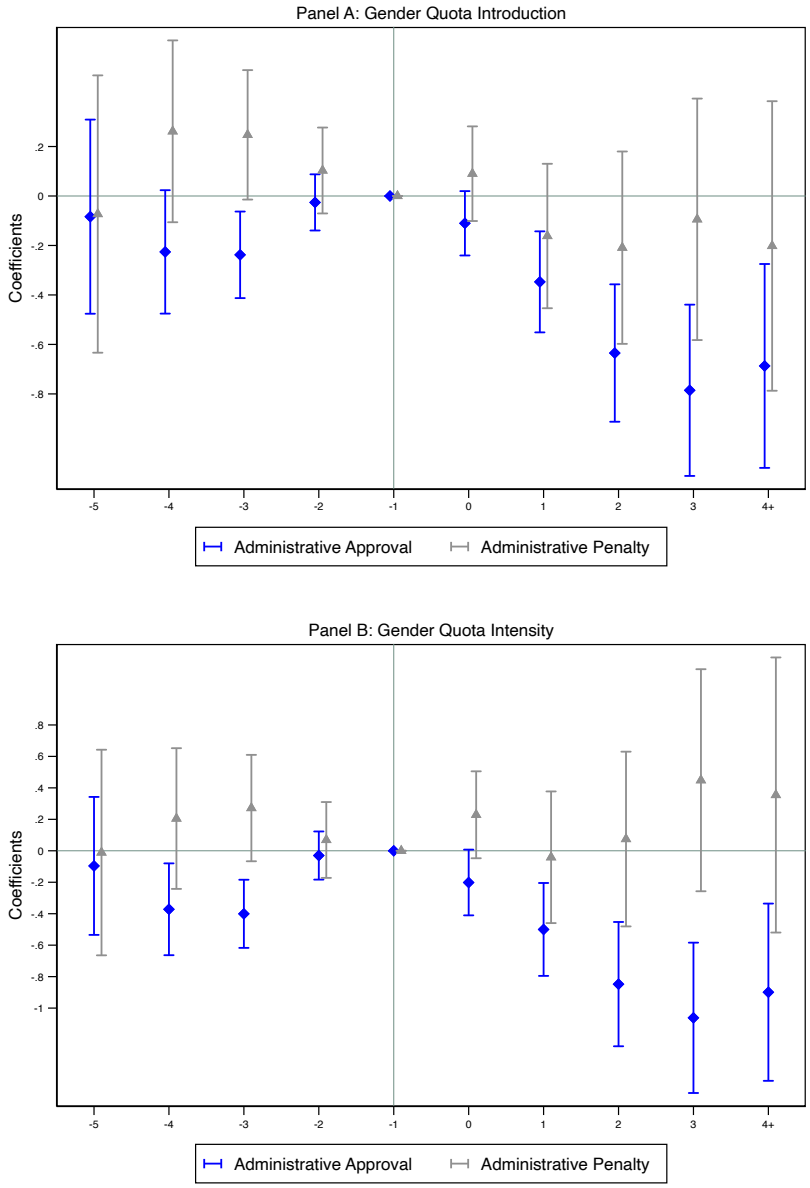
Notes - This figure shows the effect of gender quota intensity on the number of applicants. Quota intensity is measured by the average share of gender-targeted positions across all treated years. The figure displays event study estimates from both the baseline regressions and the matched sample regressions.

Figure 6: The Effect of Gender Quota on Total Revenue



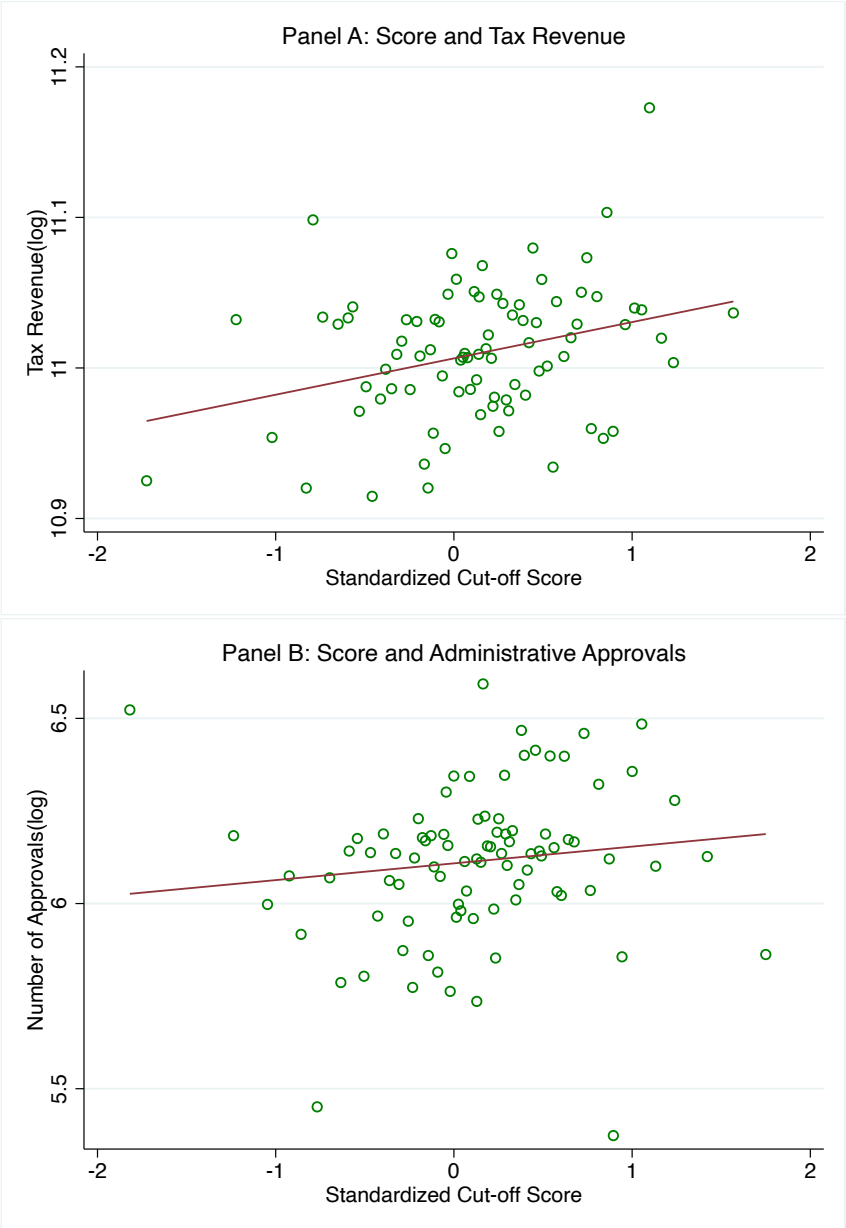
Notes - This figure shows the effect of gender quota intensity on tax revenue. Quota intensity is measured as the average share of gender-targeted positions across all treated years. Tax revenue is measured in 10,000 RMB and presented in logarithmic form. The figure displays event study estimates from 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



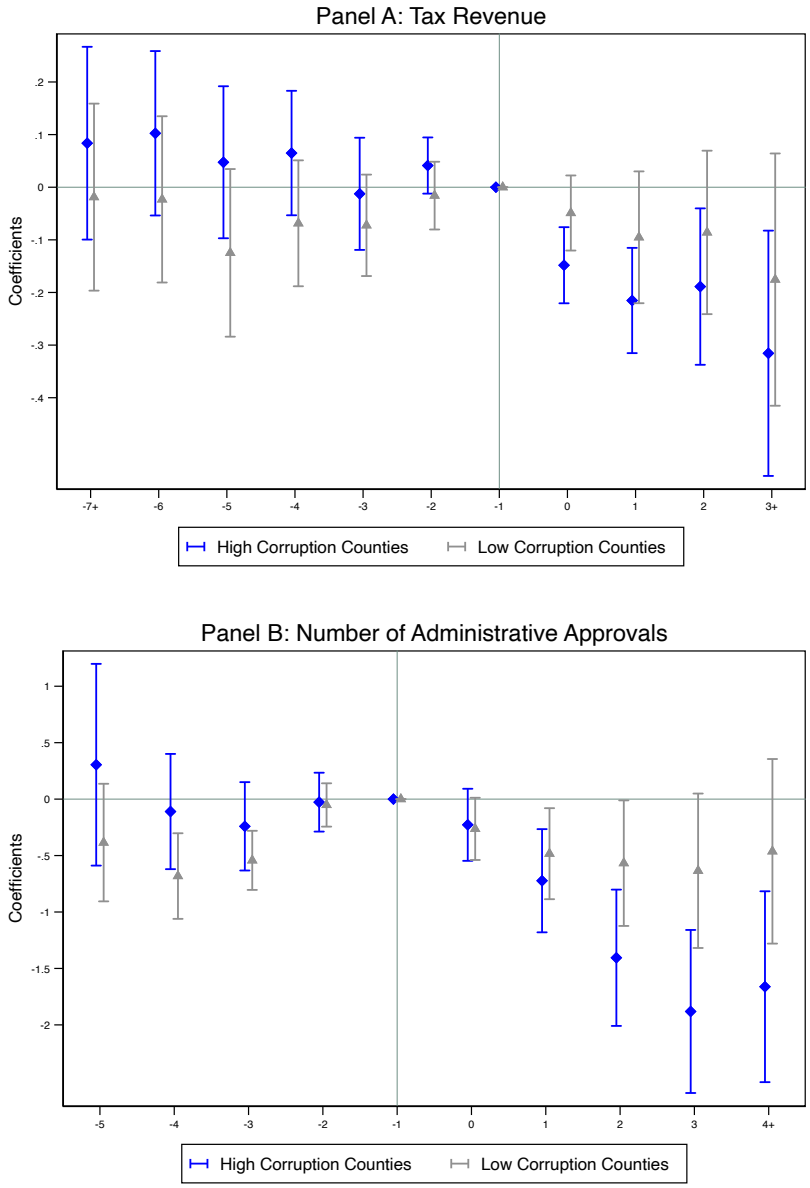
Notes - This figure shows the effect of gender quota intensity on the number of administrative cases. *Panel A* presents the effect of the introduction of gender quotas, and *Panel B* examines the effect of quota intensity, measured as the average share of gender-targeted positions across all treated 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. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

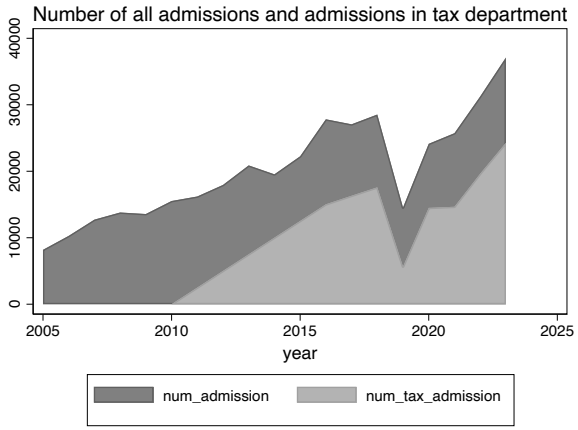
Table 3: The Effect of Quota on Tax Revenue : IV Strategy

	(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. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

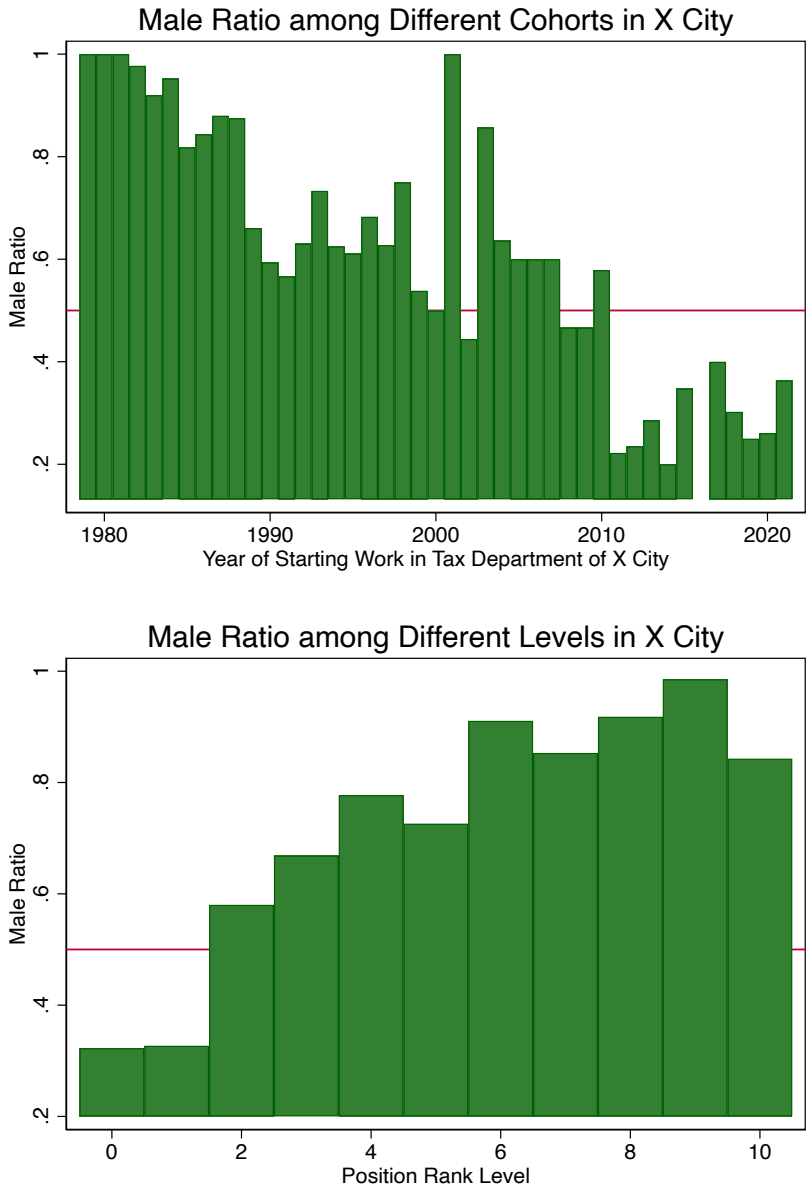
A Appendix: Supplemental Figures and Tables

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



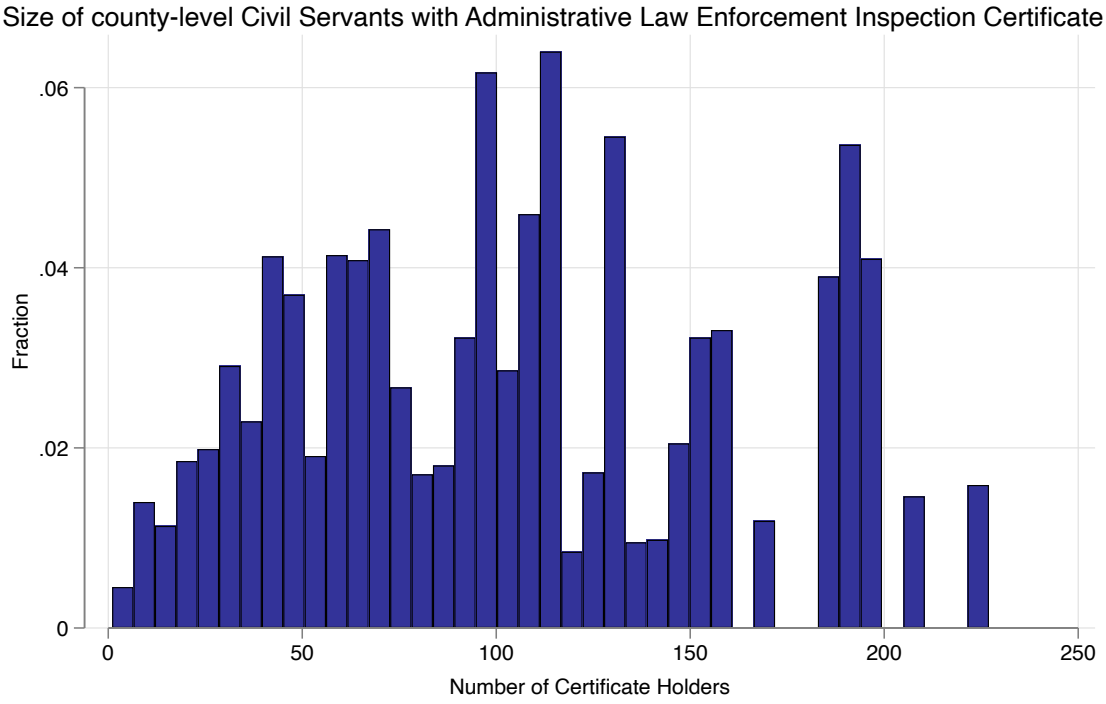
Notes - This figure shows the proportion of tax department vacancies relative to all vacancies in the National Civil Service Examination.

Figure A.2: Male Representation in Tax Department



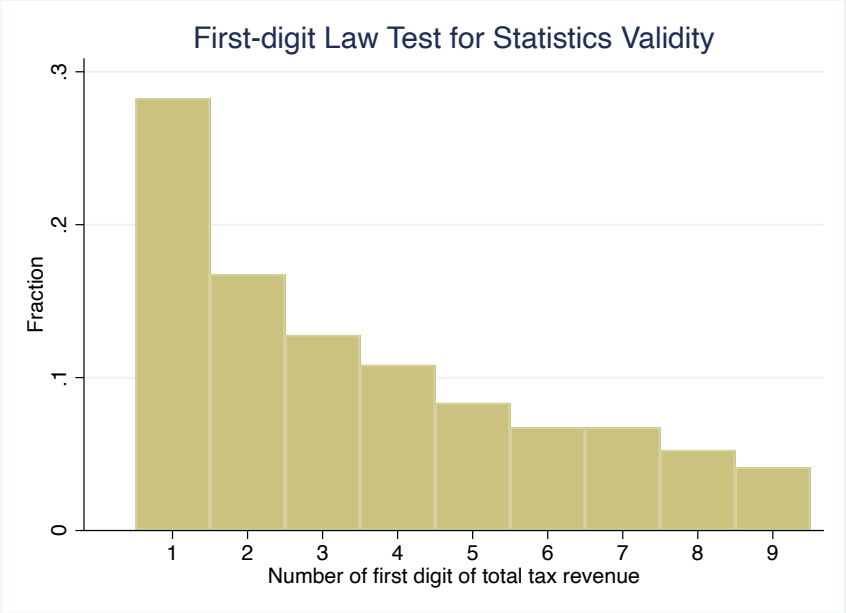
Notes - This figure presents male representation in both overall numbers and leadership roles. The left panel shows the male ratio across cohorts defined by the year of entry into the workforce. The right panel displays the male ratio across different position levels, where a higher level indicates a 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



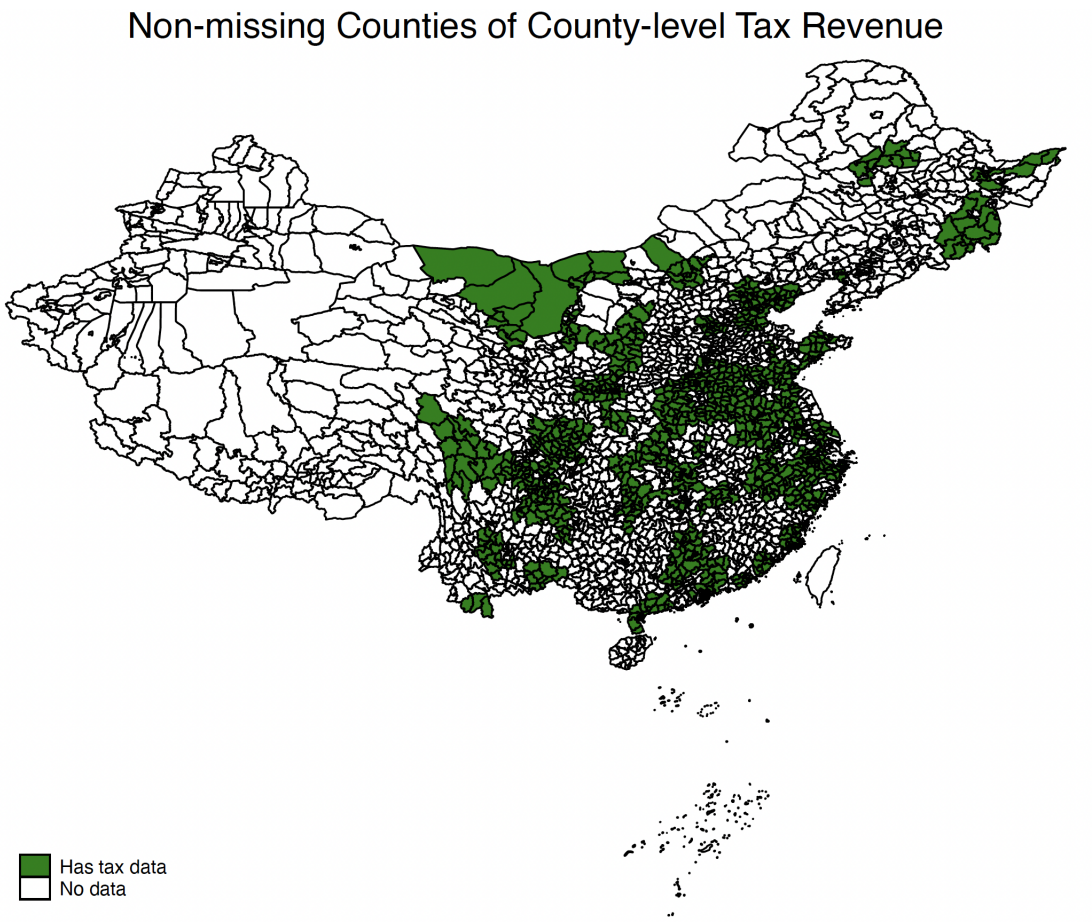
Notes - This figure shows the number of administrative certificate holders across counties in Anhui Province.

Figure A.4: Validity of Tax Revenue Data



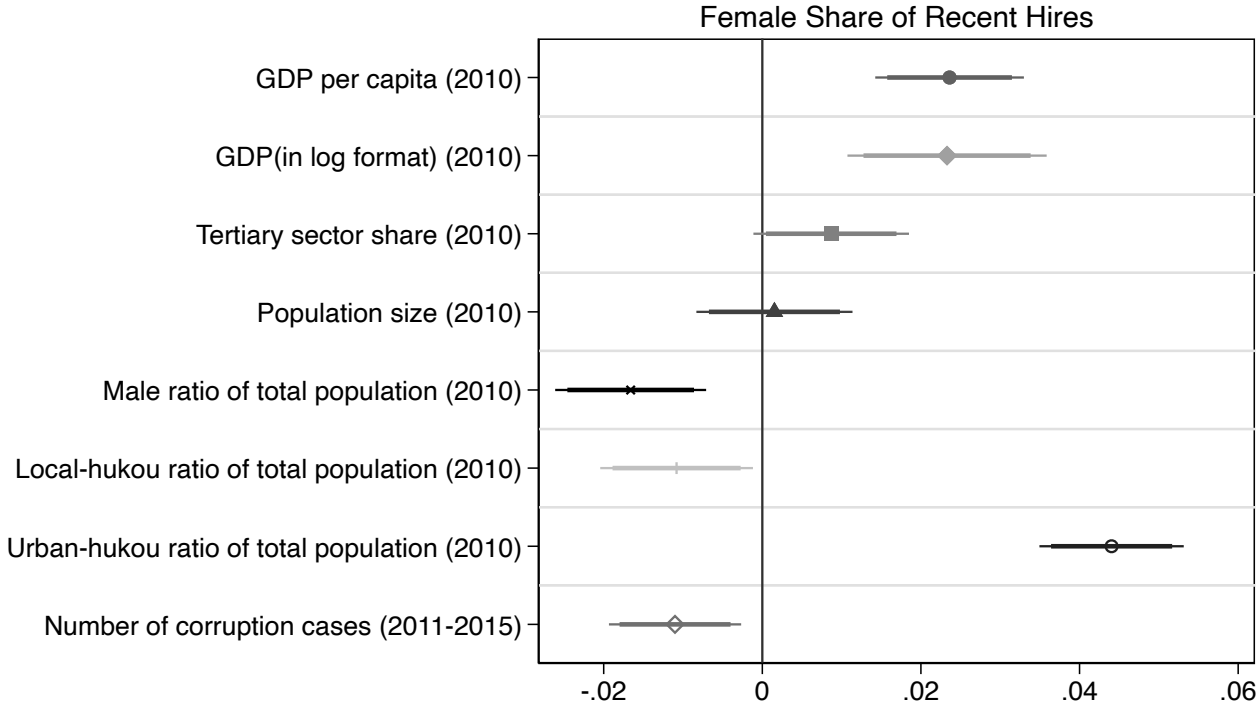
Notes - This figure shows the data validity using the First-digit Law Test.

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



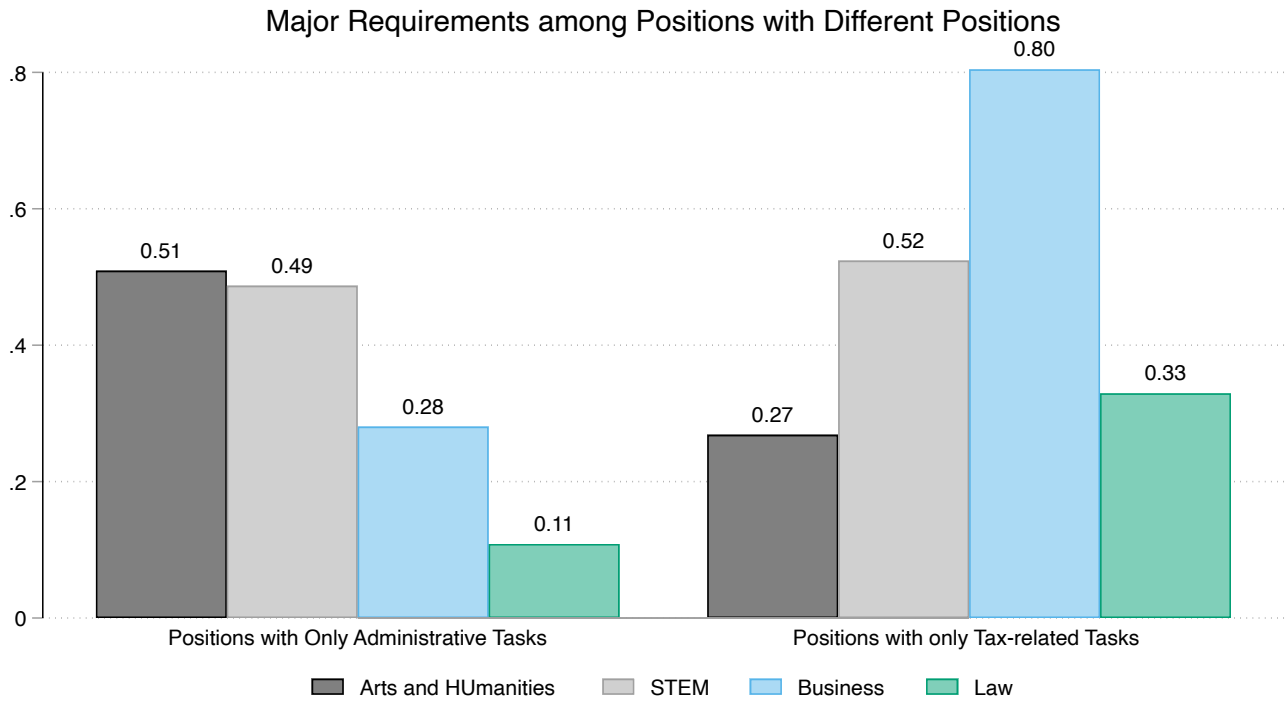
Notes - This figure displays the counties for which tax data are available from the China Statistical Yearbooks.

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



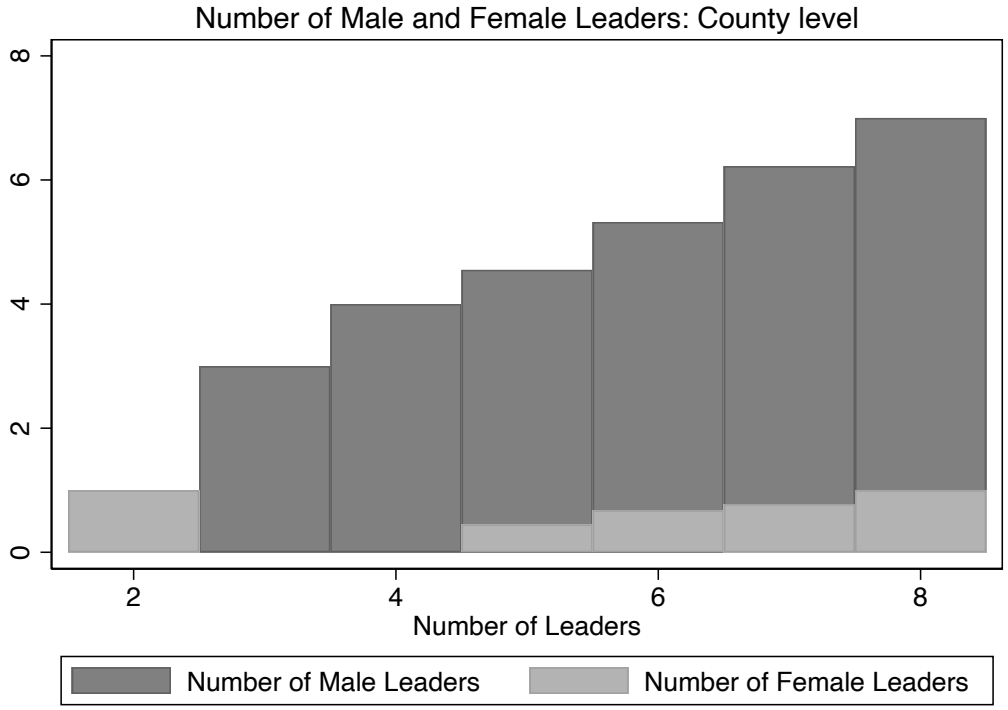
Notes - This figure shows the correlation between county characteristics and the female share of recent new hires. County characteristics data are from Census 2010 reported in county level. Corruption level is from China's Corruption Investigations Dataset(Wang, 2020). Year fixed effects and province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

Figure A.7: Position categories and major requirements



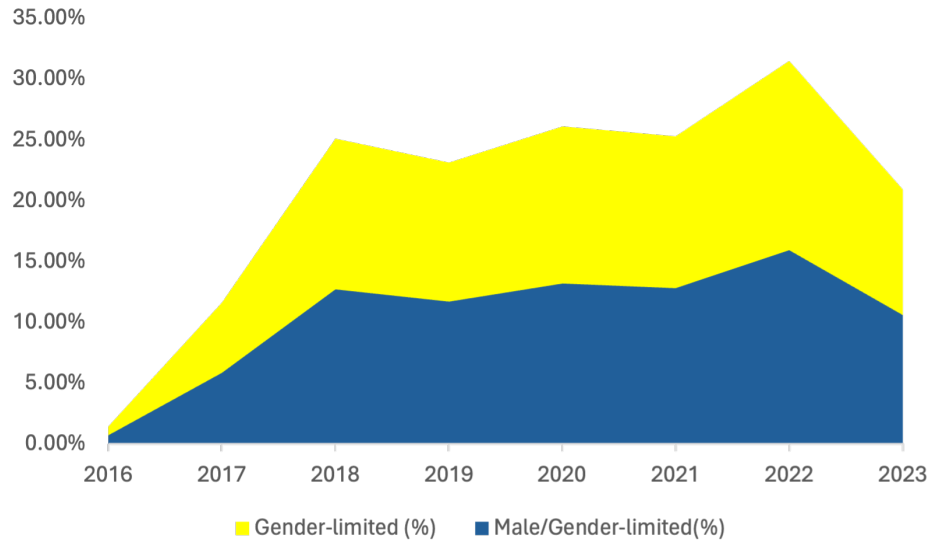
Notes - This figure displays the distribution of major requirements across different positions.

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



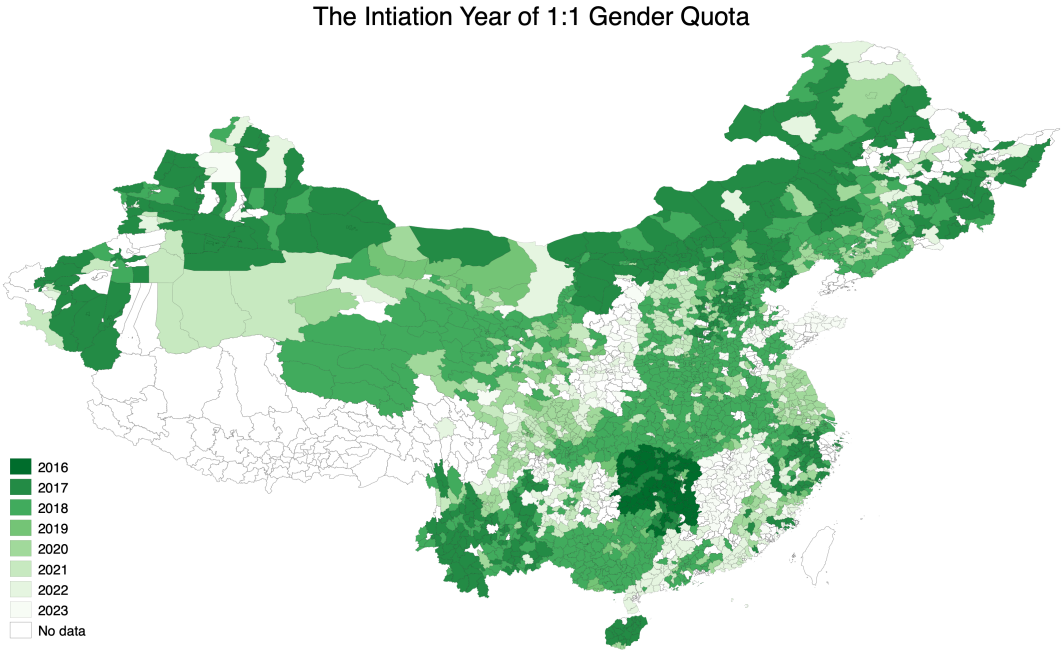
Notes - This figure shows the gender composition of leaders of county tax department. The data are web-scraped from the official websites of county-level taxation bureaus in four provinces: Anhui, Guangdong, Guizhou, and Zhejiang.

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



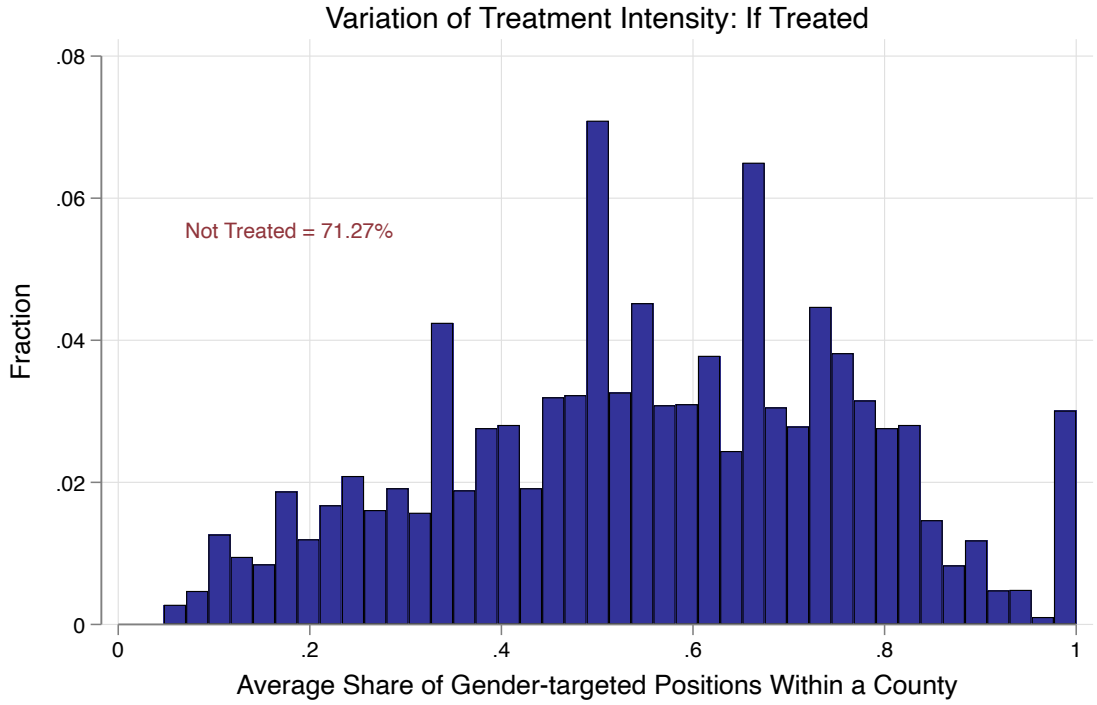
Notes - This figure shows the share of gender-targeted positions. The "Male/Gender-limited" ratio refers to the proportion of male-targeted positions among all gender-targeted positions. Male-targeted positions consistently account for 50% of gender-targeted positions.

Figure A.10: Gender Quota Enrollment Across Countries



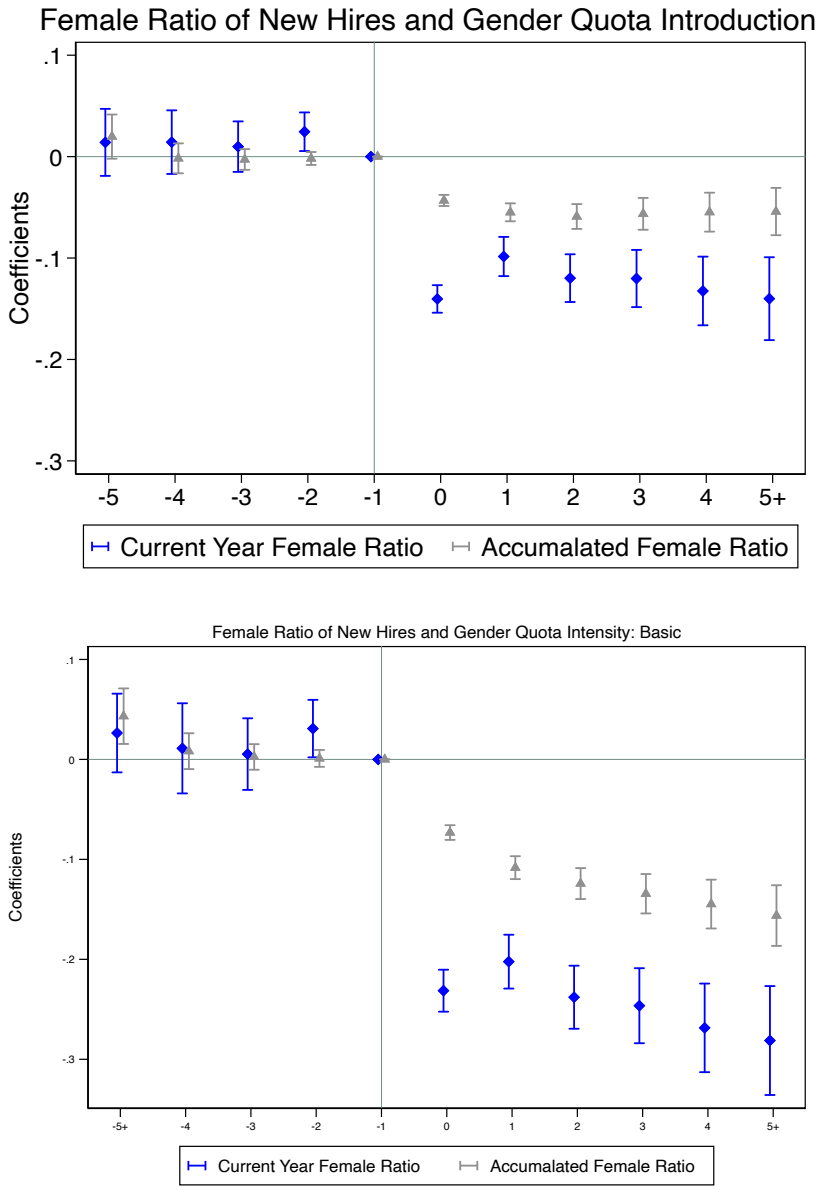
Notes - This figure shows the year in which the gender quota was first implemented across counties in China. The initiation year is defined as the first year in which at least one position in a county was designated as gender-targeted.

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



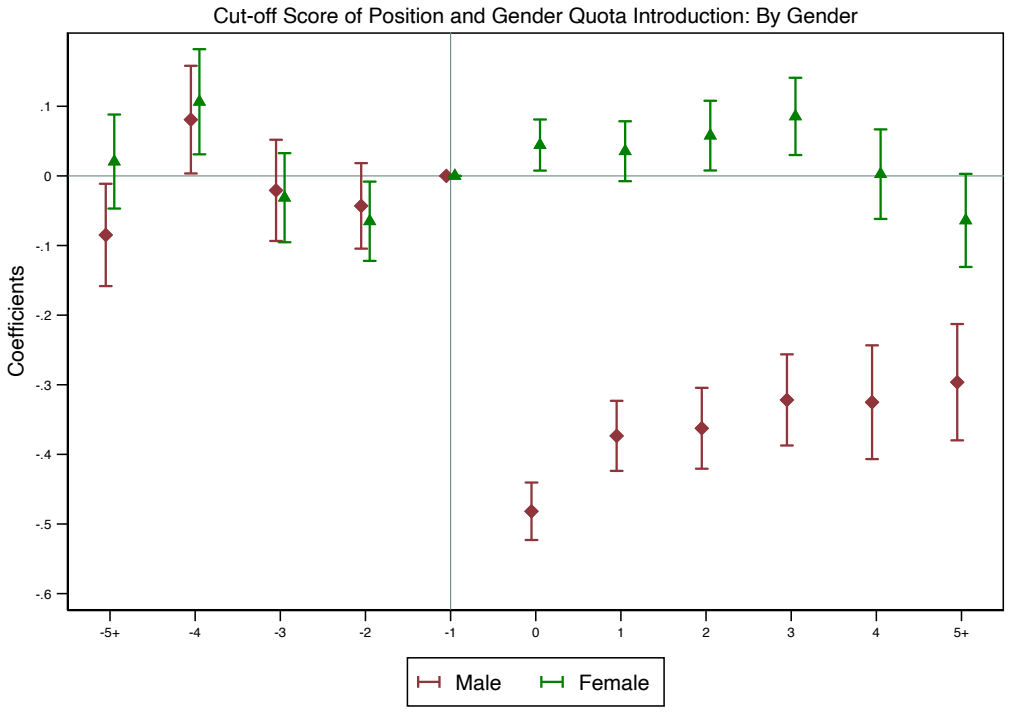
Notes - This figure shows the variation in the average share of gender-targeted positions across counties. Counties that did not implement the policy are excluded.

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



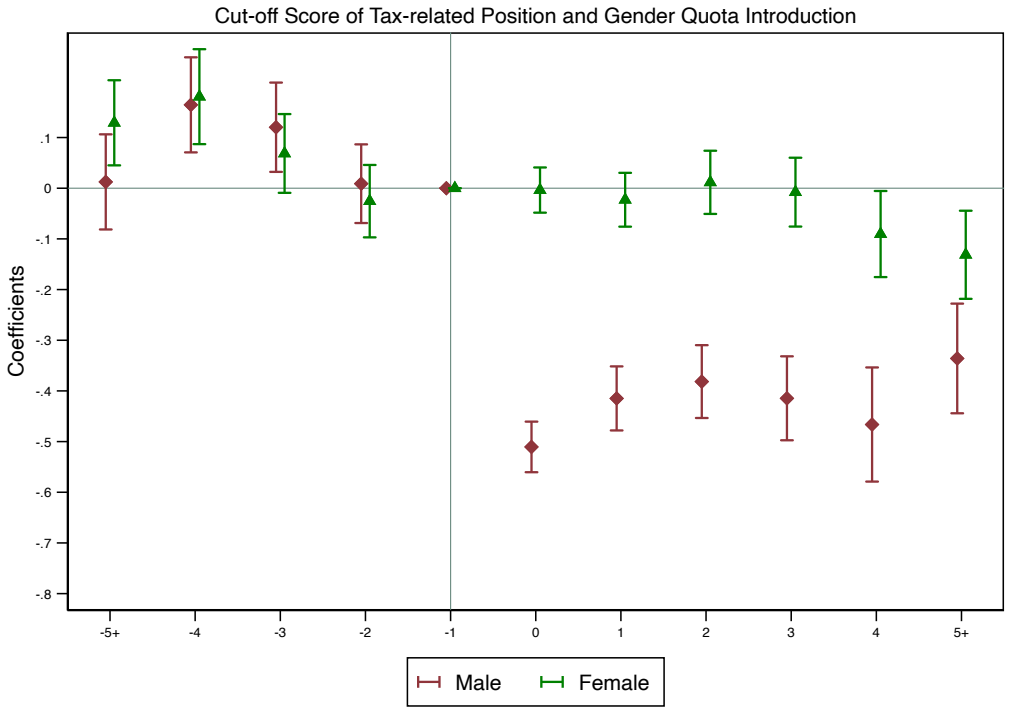
Notes - This figure presents the first-stage results of gender quota intensity without matching, measured as the average share of gender-targeted positions across all treated years. Gender composition is derived from admission data starting in 2016. The *current-year female share* refers to the share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



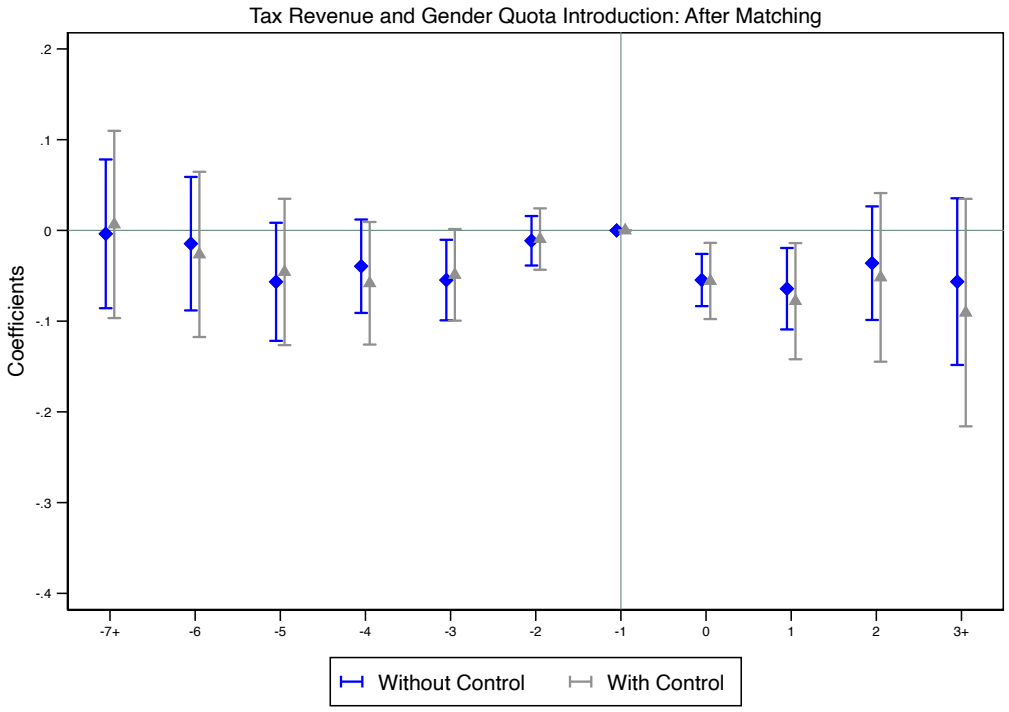
Notes - This figure shows the effect of gender quota introduction on the position-level cut-off score by gender, showing separate impacts for male-targeted and female-targeted positions. Cut-off scores are standardized. A unique position is defined as one that shares the same county and identical criteria-education,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



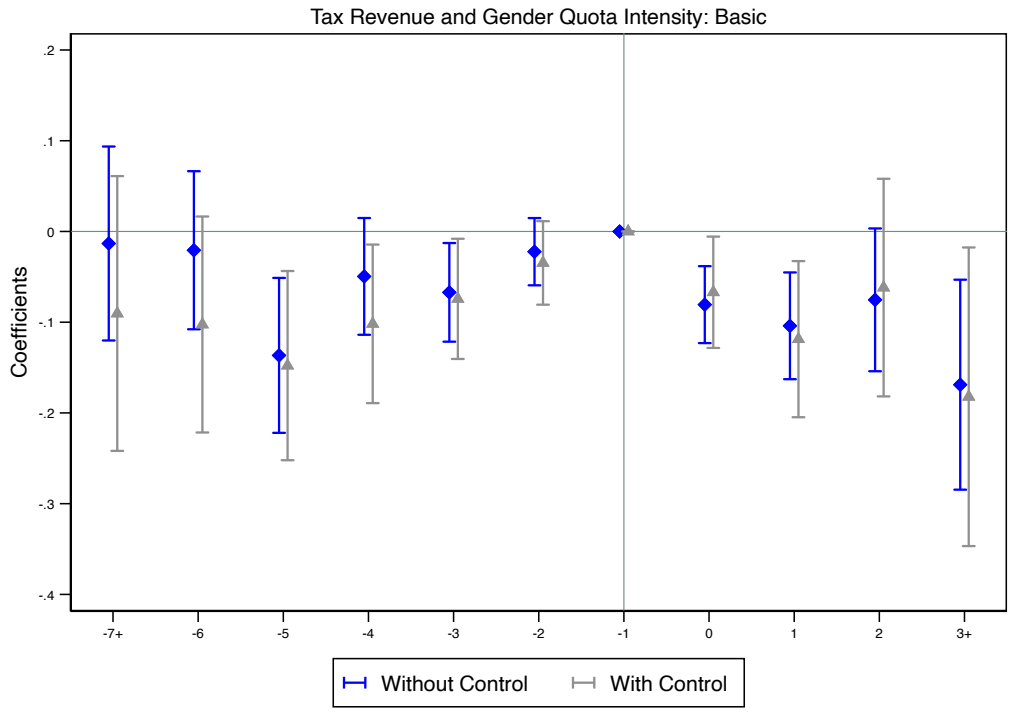
Notes - This figure shows the effect of gender quota introduction on the tax-related positions' cut-off score by gender, showing separate impacts for male-targeted and female-targeted positions. Cut-off scores are standardized. A unique position is defined as one that shares the same county and identical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

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



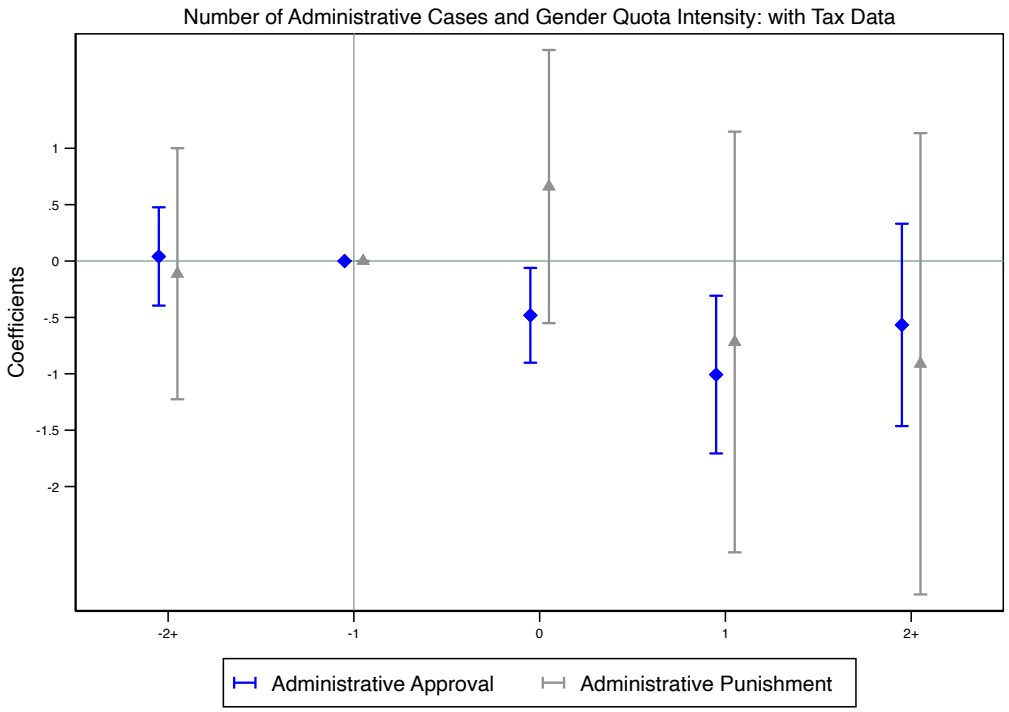
Notes - This figure shows the effect of gender quota introducation on tax revenue. Tax revenue is measured in 10,000 RMB and presented in logarithmic form. The figure displays event study estimates from regressions both with and without pre-treatment controls, using the matched sample. The controls include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

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



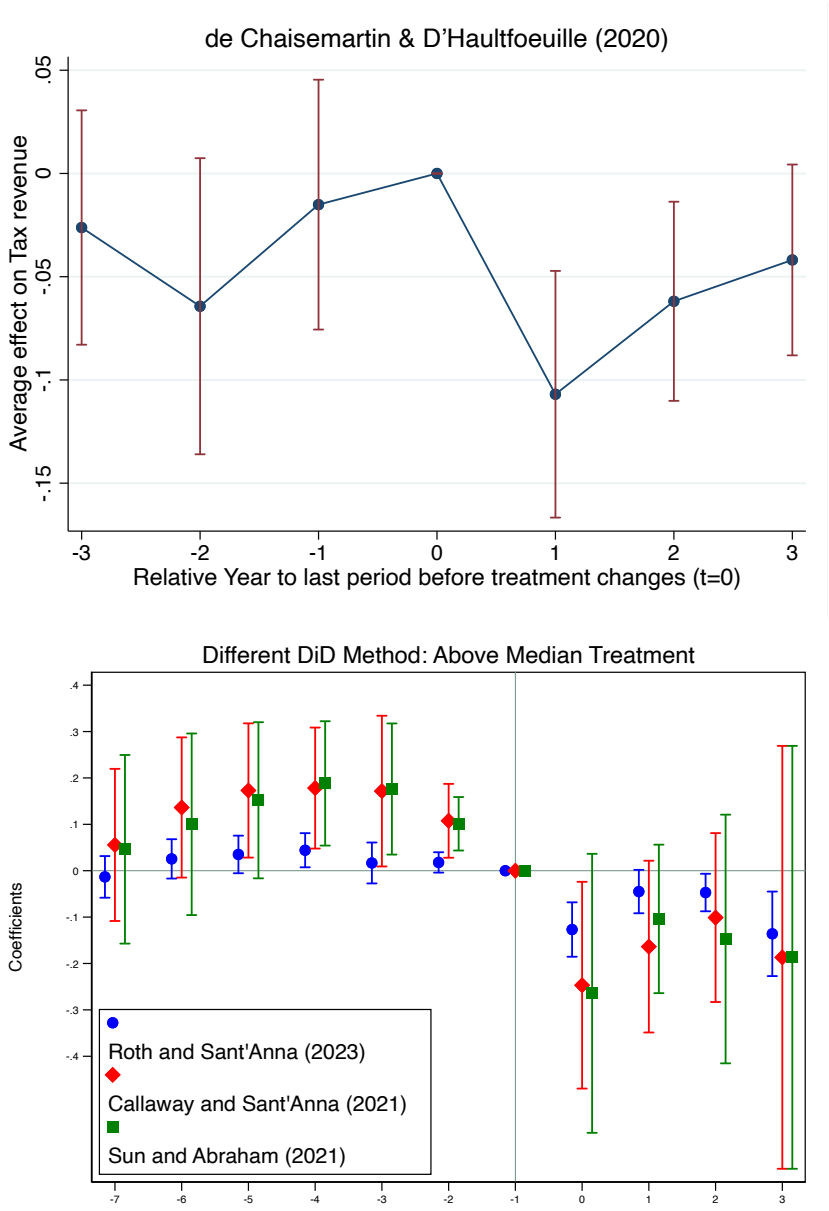
Notes - This figure shows the effect of gender quota intensity on tax revenue without matching. Quota intensity is measured as the average share of gender-targeted positions across all treated years. Tax revenue is measured in 10,000 RMB and presented in logarithmic form. The figure displays event study 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



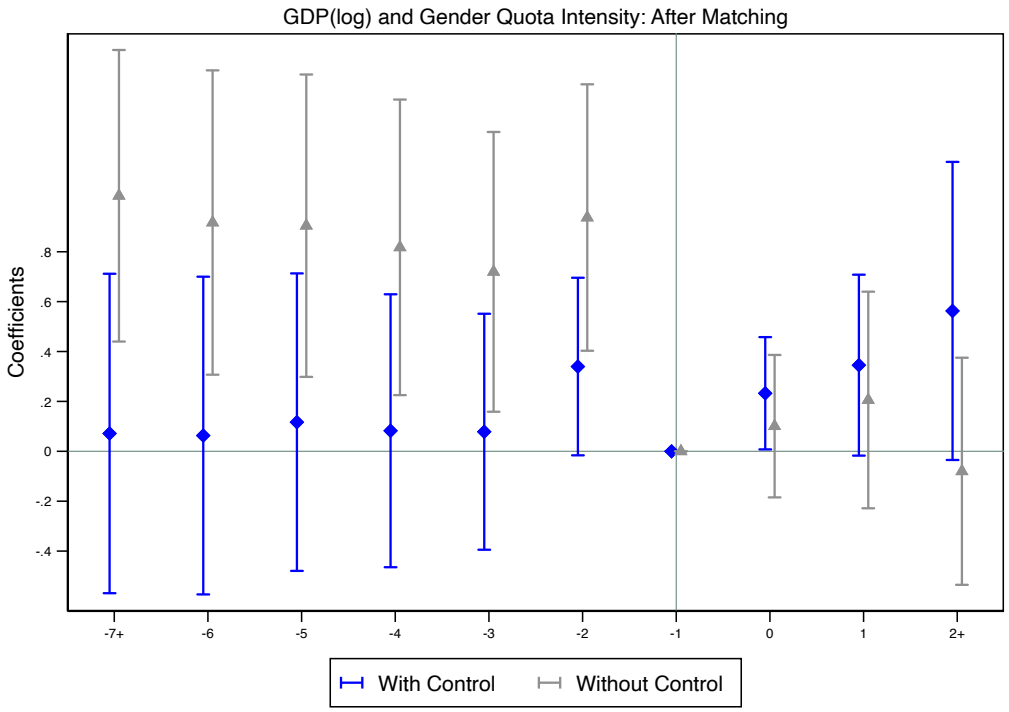
Notes - This figure shows the effect of gender quota intensity on the number of administrative cases within the sample whose tax data is available. Quota intensity is measured as the average share of gender-targeted positions across all treated years. The number of administrative cases is shown in logarithmic 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)



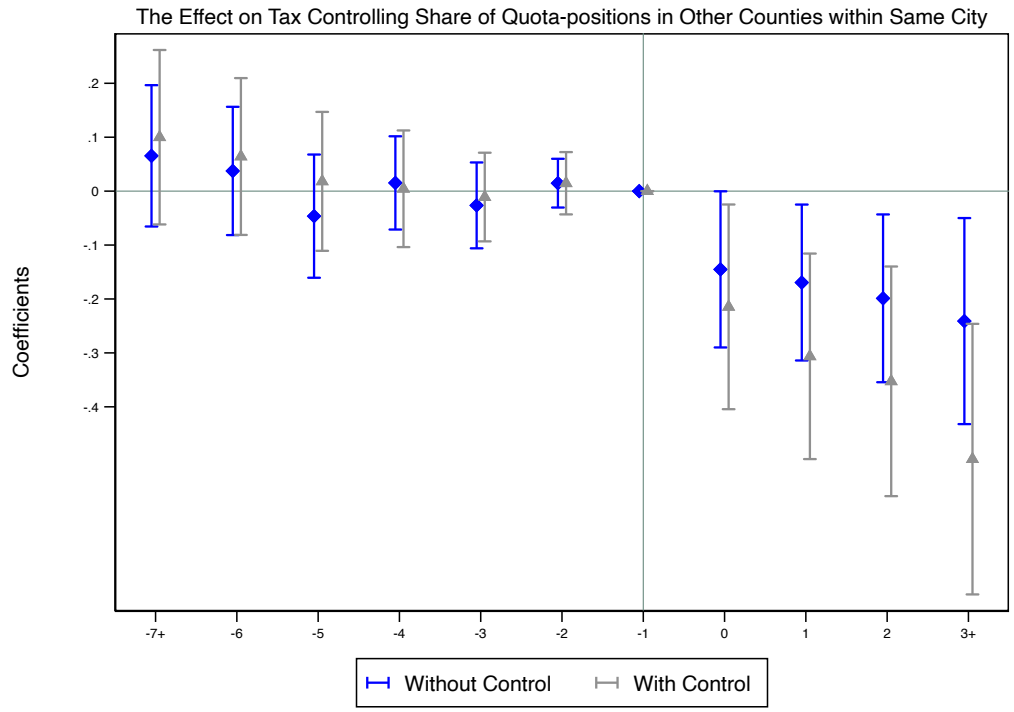
Notes - This figure shows the effect of gender quota intensity on county-level GDP. Quota intensity is measured as the average share of gender-targeted positions across all treated years. GDP is measured in 10,000 RMB and presented in logarithmic form. The figure displays event study estimates from regressions 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



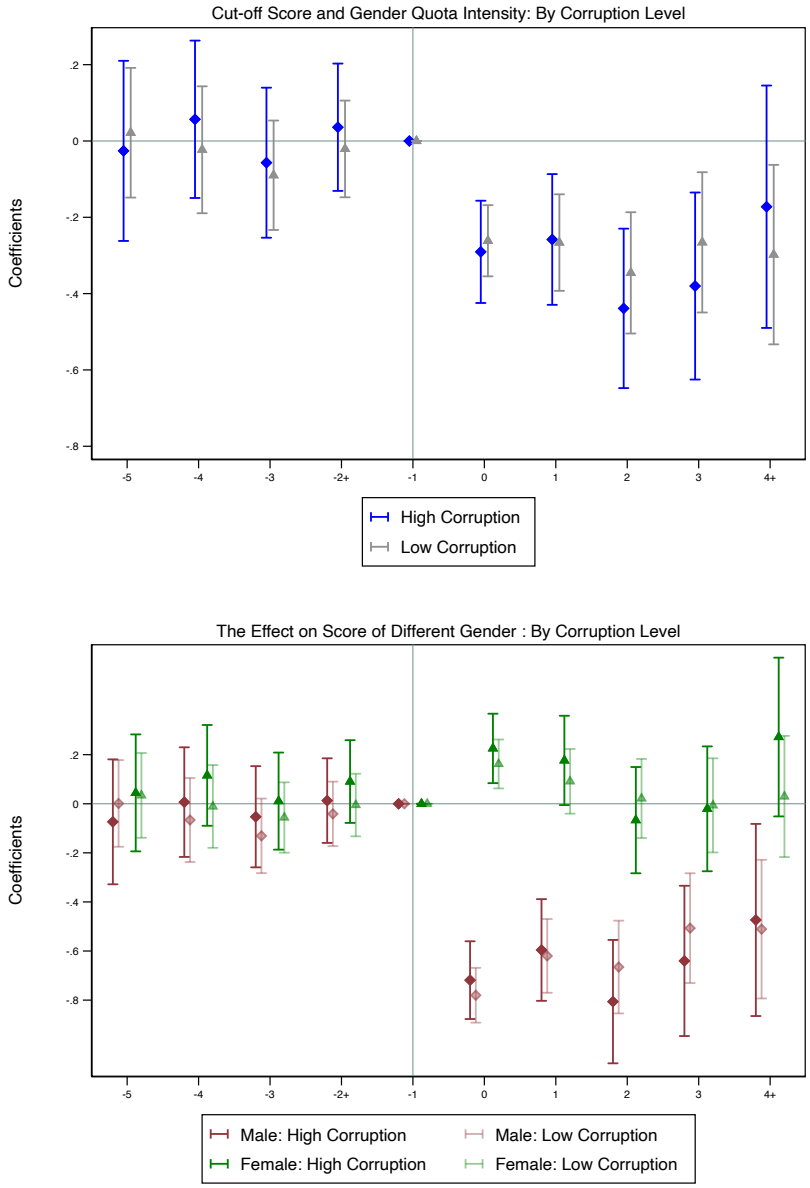
Notes - This figure shows the effect of ever hiring on tax revenue. "Ever hiring" is defined as an indicator equal to 1 if a county has hired at least once from year t onward. Tax revenue is measured in 10,000 RMB and presented in logarithmic form. All other specifications follow the main regression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

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



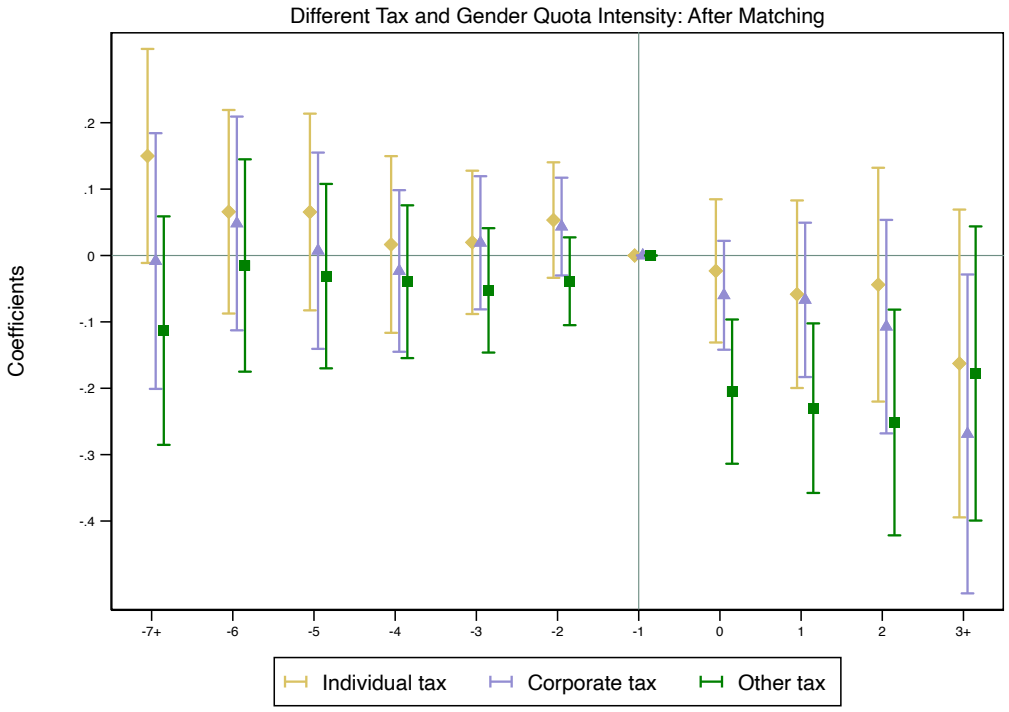
Notes - This figure shows the effect of gender quota intensity on tax revenue controlling spillover effect. Quota intensity is measured as the average share of gender-targeted positions across all treated years. Tax revenue is measured in 10,000 RMB and presented in logarithmic form. The figure displays 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



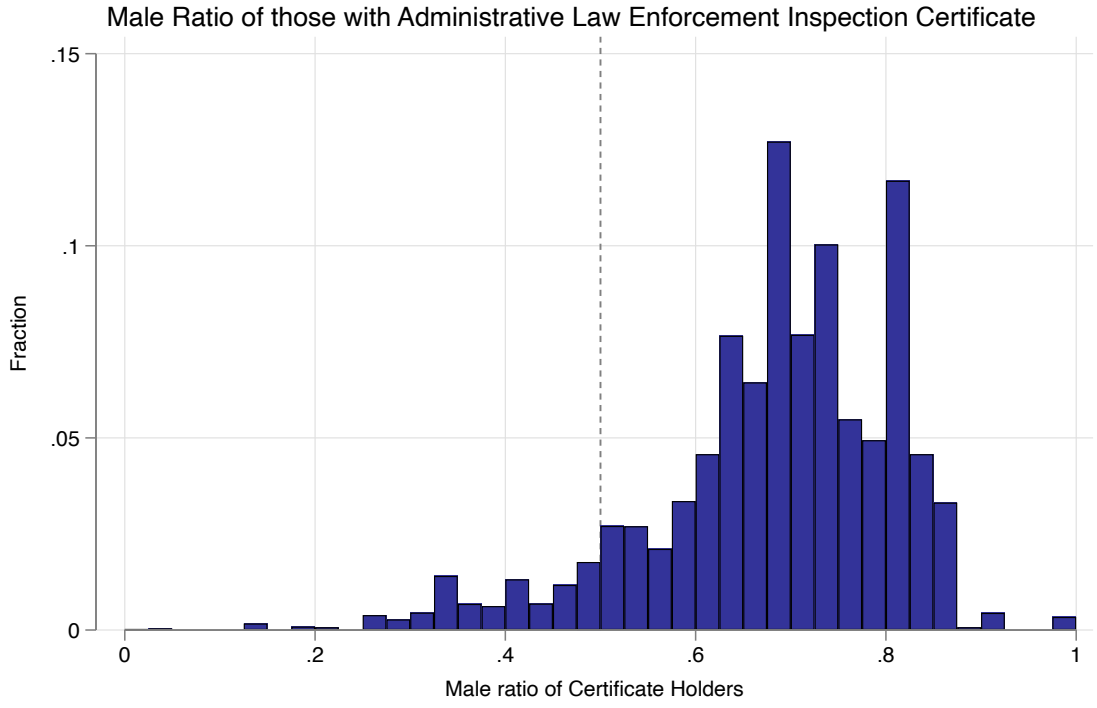
Notes - This figure shows the effect of the gender quota on cut-off scores, conditional on corruption levels. The top panel presents the effect on the county-level average cut-off score, while the bottom panel shows the effect by gender. Cut-off scores are standardized across counties. Corruption levels 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



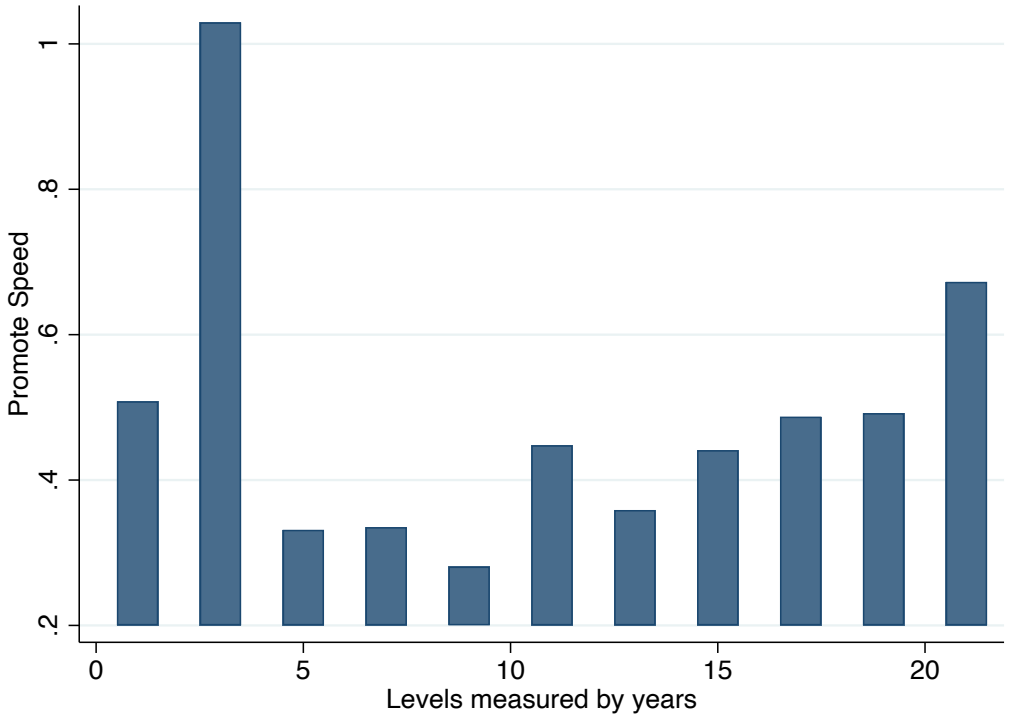
Notes - This figure shows the effect of gender quota intensity on different types of tax revenue. "Other tax" refers to tax revenue excluding individual income tax, corporate income tax, and value-added tax. Tax revenue is measured in 10,000 RMB and presented in logarithmic form. All other specifications 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 B: 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) 20km	(2) 50km	(3) 100km	(4) 200km	(5) 300km	(6) 400km	(7) 500km	(8) 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 controlling for 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) Score 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

	Tax revenue		Number of administrative approvals	
	Famous Drinking Culture (1)	Not Famous Drinking Culture (2)	Famous Drinking Culture (3)	Not Famous Drinking Culture (4)
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 heterogeneity of drinking culture. 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 Female Ratio 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) Promote Speed	(2) Promote Speed	(3) Promote Speed	(4) Promote Speed
Male	-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 × Male		0.713***		
		(0.173)		
Male ratio in the same cohort except self × Male				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. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

B Appendix: Details for China's Tax System

Under China's current tax-sharing system, the taxes collected by county bureaus are divided into central, shared, and local revenues.[23] Although tax revenue is crucial for state capacity, the central government has shifted its fiscal focus in recent years, moving away from viewing tax revenue growth as the sole objective and instead emphasizing targeted support through tax reductions for high-tech enterprises and qualified small businesses.

As the frontier units, county-level bureaus contribute approximately 85% of government revenue, handling taxpayer identification, determining liabilities, ensuring compliance, and providing tax services to local businesses and individuals. In recent decades, taxes are primarily paid through checks, bank cards, and authorized electronic systems, while county-level tax bureaus are responsible for supervising withholding, enforcing compliance, ensuring timely payments, and managing audits, disputes, and refunds. In addition to tax collection, these bureaus are responsible for tax governance, including the registration, modification, or cancellation of tax records for businesses and individuals, as well as the regulation of invoices and the timely submission of tax returns and related documents. Furthermore, they conduct audits, verify tax compliance, examine financial and operational records, investigate taxable assets, and enforce tax laws through legally authorized inspection procedures.

The tasks within each county-level tax bureau can be categorized into administrative and operational functions, each assigned to distinct departments. Administrative departments, such as the Office, Human Resources and Education, Financial Management, and Disciplinary Inspection, are primarily responsible for daily administrative duties unrelated to direct tax collection. Operational departments directly involved in tax collection, governance, and investigation include the Tax Policy Department, Tax Collection and Governance

[23] Central revenues comprise import-stage VAT and consumption tax, customs duties, domestic consumption tax, vehicle purchase tax, vessel tonnage tax, offshore petroleum resource tax, and stamp duty on securities transactions. Shared revenues include domestic VAT (shared equally between central and local governments), corporate income tax, and individual income tax (with a 60:40 split between central and local governments, respectively). However, corporate income tax paid by certain large enterprises, such as the Industrial and Commercial Bank of China, is classified as central revenue and excluded from sharing. Local revenues include environmental protection tax, property tax, urban land use tax, urban maintenance and construction tax (excluding certain centrally paid portions), land appreciation tax, resource tax (excluding offshore petroleum), general stamp duty, vehicle and vessel tax, farmland occupation tax, deed tax, and tobacco leaf tax.

Department, Policy and Legal Department, Taxpayer Service Department, Revenue Accounting Department, Information Technology Department, Risk Management Department, and Non-Tax and Social Security Department.[24]

C Appendix: Details for Robustness

As a robustness check, I implement an alternative IV–DiD strategy, with the first-stage specification defined as follows:

$$TreatInt_{ct} = \alpha ShareEven_{ct} + \mu_t + \nu_c + \Omega_c \cdot t + \xi_{ct} \tag{5}$$

where $ShareEven_{ct}$ is the share of even-numbered vacancies in county c of year t. $TreatInt_{ct}$ measuring the mean share of positions requiring gender in county c from year t onwards. μ_t accounts for the year fixed effects, ν_c denotes the county fixed effects, and ξ_{ct} represents the error term. I also incorporated the city year trend $\Omega_c \cdot t$ to account for

[24](1) Tax Policy Department: Primarily responsible for understanding and managing the policies for various types of taxes, such as VAT, consumption tax, vehicle purchase tax, corporate income tax (former national taxes); personal income tax, urban maintenance construction tax, vehicle and vessel tax, property tax, land value-added tax, resource tax, deed tax, environmental protection tax (former local taxes). Departments may be subdivided by different tax categories, such as income tax, international taxation, goods and services tax, and property and behavior tax. This is a typical business department that requires expertise in tax policies. (2) Tax Collection and Management Department: Responsible for managing the tax collection process. This is a core business department with detailed and complex daily tasks. Emergency tasks are common, which may lead to frequent overtime. (3) Policy and Legal Department: At the General Administration, provincial, municipal, and some county-level tax bureaus, the internal audit and inspection departments may be further subdivided. This department is responsible for tax law-related work, such as managing legal advisors, public lawyers, and supervising the legality of tax-related actions. It is suitable for individuals with a legal background. (4) Taxpayer Service Department: As the name suggests, this department is responsible for services related to taxpayers, including tax law publicity, taxpayer guidance, and handling taxpayer complaints. The work is relatively fragmented, and there is frequent interaction with tax service halls. (5) Revenue Accounting Department: Responsible for organizing tax revenue, tax payment collection, and tax forecasting and analysis. The work is particularly busy during specific times, such as the end or beginning of the month, and involves reconciliation with the national treasury and coordination with the People's Bank of China. This often results in overtime. (6) Information Technology Department: Responsible for the informatization of the tax system, including the operation and maintenance of the "Golden Three" system. The department also handles computer repairs, video system debugging, and other daily IT tasks. The work is extensive and detailed, requiring proficiency in computer systems. (7) Risk Management Department: Responsible for handling various risk-related tasks assigned by higher-level authorities. In simple terms, if higher authorities identify potentially erroneous data through system scans, the risk management department must verify the situation. (8) Non-Tax and Social Security Department: This is a new department following the merger of national and local tax authorities. It is mainly responsible for the collection and management of 12 types of non-tax income and social security fees. As a new business area in the tax system, the department frequently interacts with social security departments.

city-level policies that may exhibit variations over time. In the second stage, I instrument for treatment using $ShareEven_{ct}$, as specified below:

$$Y_{ct} = \hat{\beta}\widehat{TreatInt}_{ct} + \delta_t + \gamma_c + \Theta_c \cdot t + \epsilon_{ct} \tag{6}$$

where Y_{ct} represents the outcome of interests within county c in year t, including female share of new hires, average standardized cut-off score, number of applicants, county-level tax revenue. $\widehat{TreatInt}_{ct}$ indicates the estimated treatment intensity from the first stage. All other variables and fixed effects follow the same specification as in equation 2.